**High** ICPSR 36423

**School Longitudinal Study, 2009-2013 [United States]**

*United States Department of Education*

*Institute of Education Sciences*

*National Center for Education Statistics*

Base Year, First Year Follow-up and 2013 Update School and Student Documentation

Inter-university Consortium for Political and Social Research P.O. Box 1248 Ann Arbor, Michigan 48106 www.icpsr.umich.edu

**Terms of Use**

The terms of use for this study can be found at: http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/36423/terms

**Information about Copyrighted Content**

Some instruments administered as part of this study may contain in whole or substantially in part contents from copyrighted instruments. Reproductions of the instruments are provided as documentation for the analysis of the data associated with this collection. Restrictions on "fair use" apply to all copyrighted content. More information about the reproduction of copyrighted works by educators and librarians is available from the United States Copyright Office.

NOTICE WARNING CONCERNING COPYRIGHT RESTRICTIONS

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use," that user may be liable for copyright infringement.

**U.S. DEPART MENT O F EDUCATION NCES 2011-328**

High School Longitudinal Study of 2009 (HSLS:09)

Base-Year Data File Documentation

High School Longitudinal Study of 2009 (HSLS:09)

Base-Year Data File Documentation

**JULY 2011**

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

**Ying Jin Steve Leinwand** American Institutes for Research

**Laura LoGerfo** *Project Officer* National Center for Education Statistics

**NCES 2011-328 U.S. DEPARTMENT OF EDUCATION** ITUDINAL :09 STUD9 Y0:20H IGHSCHOOL HSLS:09 LONGITUDINAL STUD9 Y0:20science D

math

**U.S. Department of Education** Arne Duncan *Secretary*

**Institute of Education Sciences** John Q. Easton *Director*

**National Center for Education Statistics** Jack Buckley *Commissioner*

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

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

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

1990 K Street NW Washington, DC 20006-5651

July 2011

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

This publication is only available online. To download, view, and print the report as a PDF file, go to the NCES Publications and Products address shown above.

This report was prepared for the National Center for Education Statistics under Contract No. ED-04-CO-0036/0003 with RTI International. Mention of trade names, commercial products, or organizations does not imply endorsement by the U.S. Government.

**Suggested Citation**

Ingels, S.J., Pratt, D.J., Herget, D.R., Burns, L.J., Dever, J.A., Ottem, R., Rogers, J.E., Jin, Y., and Leinwand, S. (2011). *High School Longitudinal Study of 2009 (HSLS:09). Base-Year Data File Documentation* (NCES 2011-328). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved [date] from http://nces.ed.gov/pubsearch.

**Content Contact**

Laura LoGerfo (202) 502-7402 laura.logerfo@ed.gov

**Executive Summary**

**Design and Purposes of HSLS:09**

The High School Longitudinal Study of 2009 (HSLS:09) is the fifth in a series of National Center for Education Statistics (NCES) secondary longitudinal studies. All of the studies monitor the transition of national samples of young people from their high school experiences through their postsecondary years, including further education, participation in the work force, and the assumption of other adult roles.

The core research questions for HSLS:09 explore secondary to postsecondary transition plans and the evolution of those plans; the paths into and out of science, technology, engineering, and mathematics; and the educational and social experiences that affect these shifts.

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

The first follow-up of HSLS:09 will take place in the spring of 2012 when most sample members will be in the spring of their 11th grade. Dropouts and transfer students will be followed, as well as those who remain in the base-year school. A postsecondary update will take place in the summer of 2013, to learn about the cohort’s postsecondary plans and decisions. High school transcripts will be collected in the fall of 2013, and a second follow-up will take place a few years beyond high school graduation. Further information on study design and purposes can be found in chapter 1 of this document.

**Instrumentation**

Base-year instrument design for HSLS:09 was guided by a theoretical framework or conceptual model. The model takes the student as the fundamental unit of analysis and attempts to identify factors such as motivation, beliefs, and interests that lead to academic goal-setting and decision-making. It traces the many influences (including perceived opportunities, barriers, and costs) on students’ values and expectations that factor into their most basic education-related choices. The HSLS:09 design also acknowledges the importance of social context—families, teachers, peers, and the wider community—to students’ experiences.

The student questionnaire for in-school administration was, for the first time in the history of the study series, made electronic, as was the student assessment in algebraic reasoning. The contextual questionnaires as well—parent, teacher, school administrator, and counselor— were designed for web self-administration or computer-assisted telephone interview (CATI) administration by an interviewer. Computerization of the instruments was desired for its

*HSLS:09 Base-Year Data File Documentation iii*

*Executive Summary*

contribution to higher quality data, because online quality editing and routing through the questionnaires would reduce error. Computerization was also of benefit to the assessment, especially in the accurate assignment of second-stage forms (a two-stage adaptive test was employed).

**Student Questionnaire**. The content of the student questionnaire included both future locating and substantive questions. The questionnaire elicited demographic information (e.g., sex, race/ethnicity); language background; and school experiences in the current and previous school year (including mathematics and science experiences and course enrollment). It also inquired into constructs such as mathematics self-efficacy and identification and high school, postsecondary, and career plans, among other topics.

**Parent Questionnaire**. The parent questionnaire included locating and substantive items. Substantive items covered household members and their roles and characteristics; demographic data; information on immigration status and language use; socioeconomic status (education, occupation, income); the student’s educational history (including grade retention and change of schools); family interactions; parental involvement in the ninth-grader’s learning; and plans and preparations for postsecondary education.

**Teacher (Mathematics and Science) Questionnaire**. Teachers were selected by virtue of teaching an HSLS:09 student in science or mathematics. The teacher questionnaire collected background information about the respondent, including both demographic characteristics and educational and teaching history. Mathematics and science teachers were asked to evaluate their mathematics or science department and provide information at the classroom level. In part because of the fall timing of the survey (exposure to the student was comparatively low), teachers were not asked to supply ratings or evaluations of individual HSLS:09 students.

**School Administrator Questionnaire**. The school questionnaire allowed for two respondents: the factual information sections (1–4) could be delegated to a knowledgeable staff member, but the final section was to be completed only by the principal, because its content concerned the principal’s background and beliefs. The questionnaire elicited information about school characteristics; the student population; the school’s teachers; course offered; and the goals, beliefs, and background of the principal.

**Counselor Questionnaire**. The counselor questionnaire sought information about school programs and practices, especially as they related to activities to assist the transition of students into high school, student program or course assignment, and the various facets of counseling services. The bulk of questions inquired about staffing and practices (e.g., counselor certifications and caseloads, basis for assignment to students), programs (enrichment, services for struggling students, dropout prevention programs, and so on), and mathematics and science placement (placement criteria for both ninth-graders and upperclassmen in both mathematics and science).

*iv HSLS:09 Base-Year Data File Documentation*

*Executive Summary*

**Mathematics Assessment in Algebraic Reasoning**. The mathematics assessment was designed to provide a measure of student achievement in algebraic reasoning at two points in time (9th and 11th grade). The test framework was designed to assess a cross-section of understandings representative of the major domains of algebra and the key processes of algebra. The test and item specifications describe six domains of algebraic content and four algebraic processes:

• •

Algebraic Content Domains:

- - - - - -

The language of algebra

The language of algebra

Proportional relationships and change

Proportional relationships and change

Linear equations, inequalities, and functions

Linear equations, inequalities, and functions

Nonlinear equations, inequalities, and functions

Nonlinear equations, inequalities, and functions

Systems of equations

Systems of equations

Sequences and recursive relationships

Sequences and recursive relationships

Algebraic Processes:

*HSLS:09 Base-Year Data File Documentation v*

- - - -

Demonstrating algebraic skills

Using representations of algebraic ideas

Performing algebraic reasoning

Solving algebraic problems

The assessment was built as a two-stage test, with a router (completed by all students) and a second-stage assignment of one of three forms of variable difficulty.

**Sample Design**

In the base-year survey of HSLS:09, students were sampled through a two-stage process. First, stratified random sampling and school recruitment resulted in the identification of 1,889 eligible schools. A total of 944 of these schools participated in the study, resulting in a 55.5 percent (weighted) or 50.0 percent unweighted response rate. In the second stage of sampling, students were randomly sampled from school ninth-grade enrollment lists, with 25,206 eligible selections (or about 27 per school).

The target population at the school level was defined as regular public schools, including public charter schools, and private schools, in the 50 United States and the District of Columbia, providing instruction in both 9th and 11th grade. The target population of students was defined to include all ninth-grade students who attended the study-eligible schools in the fall 2009 term.

All students who met the target population definition were deemed eligible for the study. However, not all students were capable of completing a questionnaire or assessment. Students

*Executive Summary*

who, by virtue of language barriers or severe disabilities, were unable to participate directly in the study were retained in the sample and contextual data were sought for them. Their ability to complete the study instruments will be reassessed in the first follow-up. Of the 25,206 eligible students, 24,658 were classified as questionnaire-capable and 548 as questionnaire-incapable.

HSLS:09 school and student samples are nationally representative and also state- representative for a subset of 10 states. For most purposes, the student is the unit of analysis. Data at the school, classroom, or home level may be attached to the student record as contextual data. Several contextual respondent populations were sampled. The school’s head administrator comprises one such respondent group. The lead counselor (or most knowledgeable about the entering 9th-grade class) was identified (with the help of the school), and used as a source of school-level student contextual data. Mathematics and science teachers of HSLS:09 ninth- graders enrolled in the subject were asked to complete a teacher questionnaire. The final source of contextual data was the parent. The parent was self-selected, using the criterion that the responding parent should be the one most knowledgeable about the ninth-grader’s current situation.

**Results of School Recruitment and Data Collection**

Table ES-1 summarizes the results of school recruitment and instrument completion by each component.

**Table ES-1. Summary of HSLS:09 base-year response rates: 2009**

Instrument Selected Participated

*vi HSLS:09 Base-Year Data File Documentation*

Weighted percent

Unweighted percent School 1,889 944 55.5 50.0 School administrator1 944 888 94.9 94.1 School counselor1 944 852 91.3 90.3

Student questionnaire 2, 3 25,206 21,444 85.7 85.1 Student assessment2, 3 25,206 20,781 83.0 82.4 Parent questionnaire2 25,206 16,995 67.5 67.4 School administrator2 25,206 23,800 94.5 94.4 School counselor2 25,206 22,790 90.0 90.4

Teacher questionnaire

Mathematics teacher4 23,621 17,882 71.9 75.7 Science teacher5 22,597 16,269 70.2 72.0 1 Uses the school base weight. 2 Uses the student base weight. 3 Among questionnaire-capable students (n = 24,658), some 21,444 completed the student questionnaire, and 20,781 completed the mathematics assessment. Thus, 87.0 percent (unweighted) completed the student interview or 87.4 percent weighted. Likewise, 84.3 percent (unweighted) completed a mathematics assessment or 84.7 percent weighted. 4 Uses the student base weight. Results reflect students who were enrolled in a mathematics course. 5 Uses the student base weight. Results reflect students who were enrolled in a science course. NOTE: All percentages are based on the row under consideration. SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year.

*Executive Summary*

Overall, about half of eligible selected schools participated, for a realized sample of 944 schools. More than 21,000 students participated, or about 86 percent (weighted) of eligible selected fall ninth-graders.

Of the students participating, approximately 98 percent were surveyed in in-school sessions, and 2 percent outside school. In-school sessions were 90 minutes in length, with 15 minutes for instructions and setup, 35 minutes for the student questionnaire, and 40 minutes for the two-part, 40-question adaptive algebraic reasoning assessment.

Parent and school staff surveys (administrator, counselor, mathematics teacher, science teacher) were designed for computerized administration in either of two modes—web-based self- administration, or CATI (computerized interviewer-administration).

**Weighting**

Analytic weights are used in combination with software that accounts for HSLS:09 complex survey design to produce estimates for the target population, with appropriate standard errors. Five sets of analytic weights were computed for HSLS:09: a school-level weight, a student-level weight, two student-level weights associated with contextual data from science and mathematics courses, and a student-level weight for use with parent-supplied family and home contextual data.

The school-level weight can be used for school-level analyses involving the school administrator and counselor questionnaires. The student-level weight is used with student-level analyses. Because of the comparatively low unit response rates for parents and teachers, three special student weights—adjusted for parent, mathematics teacher, and science teacher nonresponse, respectively—were also produced. These weights presuppose that parents and teachers provide contextual data for participating students, and that the student is the unit of analysis. Variance estimation is provided through two means: BRR (Balanced Repeated Replication) provided on both public- and restricted-use files and a Taylor series linearization (available on the restricted-use file). The BRR approach to calculating HSLS:09 standard errors is recommended, although both methods give similar results.

**School Nonresponse Bias Analysis**

Nonresponse bias analyses were conducted to determine whether unit nonresponse from any of the five interview data sources (school, student, parent/guardian, mathematics teacher, and science teacher) significantly increased the estimated bias for a set of population estimates. Weighted response rates for mathematics and science teachers as well as parents of the sampled students fell below 85 percent and thus, per NCES standards, were subjected to the bias analysis procedures. The remaining sources were also included for a complete evaluation of HSLS:09.

*HSLS:09 Base-Year Data File Documentation vii*

*Executive Summary*

Characteristics associated with the school (e.g., school type, percent minority) and with the sampled student (e.g., race/ethnicity, sex) were used in the bias tests.

Findings were compared before and after the base weights were adjusted for nonresponse. For example, among the 60 tests conducted, 18.3 percent were identified as having a significant level of bias before the weights were adjusted. This amount falls to almost zero once the weights were adjusted. The proportion of significant bias tests was largest for the school analytic weights (20 percent); however, the median absolute relative bias was reduced by more than 6 percentage points.

Overall the unit nonresponse bias showed minimal levels of bias for estimates generated with the student weights, the home-life weights, and the mathematics enrollee weight. Non- negligible biases were linked to the school and science enrollee weights and were primarily a result of some domains with relatively small sample sizes.

**Imputation**

Imputation of values for missing items is also an important feature of the HSLS:09 data set. Despite the best efforts of data collectors, some questionnaire items remain unanswered. Completeness of some key student variables in HSLS:09 was also adversely affected by unit nonresponse at the parent level (e.g., family income, parental educational attainment, and occupation, all critical components of the socioeconomic status index), or, more rarely, the failure of questionnaire completers to complete an assessment.

Imputation addresses the problem of missing items. Advantages of using imputed values include the ability to use all study respondent records in an analysis (complete-case analysis), which affords more power for statistical tests. 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 likely less biased than those produced with the incomplete data file.

HSLS:09 variables in general did not suffer from high levels of item nonresponse. Nevertheless, a set of key analytic variables was identified for item imputation to facilitate complete-case analysis on data obtained from the participating ninth-grade students. Values were assigned in place of missing responses for 18 variables identified from the student and parent questionnaires through single-value imputation. Missing student ability estimates in mathematics (*theta*), the associated standard error of measurement (*sem*) for the *theta*, and SES values derived from parent responses were replaced with five values using a multiple imputation procedure. Regardless of the method, indicator variables (flags) were created to allow users to easily identify the imputed values.

**Disclosure Risk Analysis and Protections**

The disclosure treatment methods used to produce the HSLS:09 base-year data files include variable recoding, suppressing, and swapping. Some variables that had values with

*viii HSLS:09 Base-Year Data File Documentation*

*Executive Summary*

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. Thus, rare events or characteristics have been masked for certain variables.

Other variables were classified as high risk and were suppressed from the public-use file. The suppressing techniques included removing the response from the file (i.e., reset to a “suppressed” reserve code) or removing records entirely from the public-use file (e.g., student nonrespondents).

*HSLS:09 Base-Year Data File Documentation ix*

**Foreword**

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

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

Chapter 2 describes the base-year data collection instruments, including both the development and content of the student, parent, science and mathematics teacher, counselor, and school administrator questionnaires. Chapter 2 also provides information on the development of the direct algebra assessment and the scoring procedures and psychometric characteristics.

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

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

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

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

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

*HSLS:09 Base-Year Data File Documentation xi*

**Acknowledgments**

We thank all the individuals who participated in the High School Longitudinal Study of 2009 (HSLS:09), including the thousands of students, parents, and school personnel who generously gave their time and approval to provide the data that the study reports.

The authors of this report also thank the many individuals who assisted in the planning of HSLS:09. We are particularly indebted to the HSLS:09 base-year Technical Review Panel (TRP), which met for three 2-day meetings in which panelists reviewed plans for the study, helped refine them, and provided important suggestions to help guide development of the instrumentation. The following individuals serve as the nonfederal members of the HSLS:09 Technical Review Panel: Clifford Adelman, Kathy Borman, Daryl Chubin, Jeremy Finn, Eric Grodsky, Thomas Hoffer, Vinetta Jones, Donald Rock, James Rosenbaum, Russ Rumberger, Phillip Sadler, Sharon Senk, and Timothy Urdan.

For the mathematics assessment, John Dossey offered his expertise and advice in the creation of suitable new items, while Donald Rock, in addition to his TRP role, provided counsel on assessment design and development. Item development was significantly enhanced by the expertise of a Mathematics Advisory Committee, which comprised Hyman Bass, Katherine Halvorsen, Joan Leitzel, Mark Saul, and Ann Shannon.

*xii HSLS:09 Base-Year Data File Documentation*

**Contents**

**Page Executive Summary ..................................................................................................................... iii**

**Foreword ....................................................................................................................................... xi Acknowledgments ....................................................................................................................... xii**

**List of Tables ............................................................................................................................. xvii List of Figures ........................................................................................................................ xxxvii**

**Chapter 1. Introduction ............................................................................................................... 1** 1.1 Overview of the Data File Documentation ...................................................................... 1 1.2 Historical Background ..................................................................................................... 2 1.2.1 NCES Secondary Longitudinal Studies Program ................................................ 2 1.2.2 National Longitudinal Study of the High School Class of 1972 ......................... 4 1.2.3 High School and Beyond ..................................................................................... 4 1.2.4 National Education Longitudinal Study of 1988 ................................................. 5 1.2.5 Education Longitudinal Study of 2002 ................................................................ 6 1.3 High School Longitudinal Study of 2009 ........................................................................ 6 1.3.1 Overview of the HSLS:09 Design and Objectives .............................................. 6 1.3.2 HSLS:09 Research and Policy Issues .................................................................. 8 1.3.3 HSLS:09 Analysis Files and Systems ................................................................ 10

**Chapter 2. Base-Year Instrumentation..................................................................................... 11** 2.1 Introduction .................................................................................................................... 11 2.1.1 Instrument Development Process and Procedures ............................................. 11 2.1.2 HSLS:09 Instrument Development Goals ......................................................... 13 2.2 Base-Year Questionnaires.............................................................................................. 14 2.2.1 Student ............................................................................................................... 14 2.2.2 Parent ................................................................................................................. 15 2.2.3 Teacher............................................................................................................... 17 2.2.4 School Administrator ......................................................................................... 18 2.2.5 Counselor ........................................................................................................... 20 2.2.6 Rules for Defining Completed Interviews ......................................................... 20 2.3 HSLS:09 Mathematics Assessment of Algebraic Reasoning ........................................ 22 2.3.2 Scoring Procedures ............................................................................................ 26 2.3.3 Score Descriptions and Summary Statistics ...................................................... 27 2.3.4 Psychometric Properties of the Test .................................................................. 32 2.4 Linkage With Prior NCES Studies ................................................................................ 33 2.4.1 Questionnaire Linkage With Prior NCES Studies ............................................. 33 2.4.2 Assessment Linkage With Prior NCES Studies ................................................ 34

*HSLS:09 Base-Year Data File Documentation xiii*

*Contents*

**Page**

**Chapter 3. Sample Design .......................................................................................................... 35** 3.1 Base-Year Sample Design Overview ............................................................................ 35 3.2 Selection of School Sample ........................................................................................... 35 3.2.1 Target Population ............................................................................................... 35 3.2.2 School Sampling Frame ..................................................................................... 36 3.2.3 First-Stage Sample Design................................................................................. 38 3.2.4 Augmented-Sample States ................................................................................. 40 3.2.5 School Sample Size ........................................................................................... 41 3.2.6 School Eligibility ............................................................................................... 42 3.3 Selection of Student Sample .......................................................................................... 43 3.3.1 Target Population ............................................................................................... 43 3.3.2 Student Sample Sizes ......................................................................................... 43 3.3.3 Student Sampling Frames .................................................................................. 46 3.3.4 Second-Stage Sample Design ............................................................................ 48 3.3.5 Student Eligibility and Exclusions ..................................................................... 49 3.4 Selection of Contextual Samples ................................................................................... 51 3.4.1 Administrator Survey ......................................................................................... 51 3.4.2 Counselor Survey ............................................................................................... 51 3.4.3 Science and Mathematics Teacher Surveys ....................................................... 51 3.4.4 Parent Survey ..................................................................................................... 53

**Chapter 4. Data Collection Methodology and Results ............................................................ 55** 4.1 Introduction .................................................................................................................... 55 4.2 Data Collection Methodology ........................................................................................ 56 4.2.1 School Recruitment Overview ........................................................................... 56 4.2.2 Student Data Collection ..................................................................................... 63 4.2.3 Parent Data Collection ....................................................................................... 70 4.2.4 Staff Data Collection ......................................................................................... 78 4.3 Data Collection Results: Response and Participation Rates .......................................... 81

**Chapter 5. Data Preparation and Processing ........................................................................... 91** 5.1 Overview of Systems Design, Development, and Testing ............................................ 91 5.2 Data Processing and File Preparation ............................................................................ 92 5.3 Data Cleaning and Editing for Web/CATI/PAPI .......................................................... 92 5.3.1 Teacher Data ...................................................................................................... 93 5.4 Coding, Upcoding, Recoding, and Adjudication ........................................................... 97 5.4.1 Major Field of Study Coding ............................................................................. 97 5.4.2 Occupation Coding .......................................................................................... 101 5.4.3 Student Job at Age 30 Coding ......................................................................... 105 5.4.4 Students’ Previous Schools .............................................................................. 106 5.4.5 Teachers’ Postsecondary Institutions............................................................... 107 5.5 Construction of Select Student, Teacher, School Counselor, and School

Administrator Scale Scores .......................................................................................... 107 5.5.1 Student ............................................................................................................. 108 5.5.2 Teacher............................................................................................................. 110

*xiv HSLS:09 Base-Year Data File Documentation*

*Contents*

**Page** 5.5.3 School Counselor ............................................................................................. 110 5.5.4 School Administrator ....................................................................................... 110

**Chapter 6. Analytic Weights, Variance Estimation, and Nonresponse Bias Analysis ....... 115** 6.1 Overview: General Approach to Weighting ................................................................ 115 6.2 Choosing an Analytic Weight ...................................................................................... 115 6.3 School Weights ............................................................................................................ 118 6.3.1 Base Weight ..................................................................................................... 118 6.3.2 Adjustment for Nonresponse ........................................................................... 119 6.3.3 Weight Calibration and Final Analytic Weight ............................................... 121 6.3.4 Balanced Repeated Replication Weights ......................................................... 123 6.4 Student Weights ........................................................................................................... 124 6.4.1 Base Weight ..................................................................................................... 124 6.4.2 Adjustments for Nonresponse .......................................................................... 124 6.4.3 Weight Calibration and Final Analytic Weight ............................................... 126 6.4.4 Balanced Repeated Replication Weights ......................................................... 127 6.5 Student-Level Contextual Analytic Weights ............................................................... 127 6.5.1 Administrator and Counselor Data .................................................................. 128 6.5.2 Science and Mathematics Course Enrollee Contextual Weights ..................... 128 6.5.3 Student Home-Life Contextual Weights.......................................................... 130 6.6 Variance Estimation ..................................................................................................... 131 6.6.1 Standard Errors ................................................................................................ 131 6.6.2 Design Effects .................................................................................................. 135 6.7 Unit Nonresponse Bias Analysis ................................................................................. 138 6.7.1 Test of Significant Nonresponse Bias .............................................................. 138 6.7.2 School Nonresponse Bias Analysis ................................................................. 139 6.7.3 Student-Level Nonresponse Bias Analysis ...................................................... 139 6.7.4 Student-Level Contextual Nonresponse Bias Analysis ................................... 140 6.8 Quality Control for the Weights .................................................................................. 140

**Chapter 7. Item Response, Imputation, and Disclosure Treatment..................................... 141** 7.1 Overview ...................................................................................................................... 141 7.2 Item Nonresponse Bias Analysis ................................................................................. 141 7.2.1 Estimating Item Nonresponse Bias .................................................................. 141 7.2.2 Item Response Rates ........................................................................................ 143 7.2.3 High Item-Nonresponse Items ......................................................................... 143 7.2.4 Summarized Results ........................................................................................ 155 7.3 Item Imputation............................................................................................................ 161 7.3.1 Single-Value Imputation .................................................................................. 162 7.3.2 Multiple Imputation ......................................................................................... 166 7.4 Disclosure Risk Analysis and Protections ................................................................... 169 7.4.1 Base-Year Data Products ................................................................................. 170 7.4.2 Recoding, Suppression, and Swapping ............................................................ 170

*HSLS:09 Base-Year Data File Documentation xv*

*Contents*

**Page**

**Chapter 8. Data File Structure and Contents ........................................................................ 173** 8.1 Base-Year eDAT and ECB DVD Data Structure ........................................................ 173 8.1.1 Overview .......................................................................................................... 173 8.1.2 Student File ...................................................................................................... 173 8.1.3 School File ....................................................................................................... 173 8.1.4 CCD, PSS, and Other Restricted-Use Linkages .............................................. 173 8.1.5 Reserve Codes.................................................................................................. 174 8.1.6 Education Data Analysis Tool and Electronic Codebook ............................... 174 8.2 Composite Variables .................................................................................................... 175 8.2.1 Naming Conventions ....................................................................................... 175

**References .................................................................................................................................. 177 Appendixes**

Appendix A. Base-Year Questionnaires Appendix B. HSLS:09 Mathematics Assessment Items: 2009 Appendix C. Glossary of Terms Appendix D. Details of School and Student Sampling Appendix E. Parental Passive and Active Consent Forms Appendix F. Documentation for Composite Variables Appendix G. Standard Errors and Design Effects Appendix H. Unit and Item Nonresponse Bias Analyses Appendix I. Imputation Details Appendix J. Socioeconomic Status Appendix K. Base-Year Codebooks Appendix L. Comparison of Public and Restricted-Use Files Appendix M. Variable List for the HSLS:09 Electronic Codebook (ECB)

*xvi HSLS:09 Base-Year Data File Documentation*

**List of Tables**

**Table Page**

ES-1 Summary of HSLS:09 base-year response rates: 2009 ................................................... iv 1 HSLS:09 Mathematics Assessment grade-9 main study design: 2009 ............................. 24 2 Number and percentage of HSLS:09 Mathematics Assessment test-takers by

form: 2009 ......................................................................................................................... 26 3 Various types of scores from HSLS:09 Mathematics Assessment, by variable:

2009 ................................................................................................................................... 29 4 HSLS:09 algebra probability of proficiency scores, by variable: 2009 ............................ 31 5 School sampling-frame eligibility status and number sampled by sampling

stratum ............................................................................................................................... 37 6 Mapping between HSLS:09 locale and the variables included on the NCES

sampling frame files .......................................................................................................... 39 7 Postsampling eligibility and response status for schools by sampling stratum ................ 42 8 Classification for schools identified as ineligible during recruitment phase .................... 43 9 Minimum respondent sample sizes from power calculations by school and student

characteristics .................................................................................................................... 44 10 Student enrollment list counts, total number sampled, and average sampled by

sample design characteristics ............................................................................................ 45 11 Mode of delivery for HSLS:09 sample school ninth-grade enrollment lists .................... 47 12 Distribution of HSLS:09 sampled students by study eligibility status ............................. 49 13 Distribution of HSLS:09 study-eligible students by capability ........................................ 50 14 Teachers identified for the HSLS:09 by subject area, school type, region, and

locale ................................................................................................................................. 53 15 Summary of HSLS:09 base-year response rates: 2009 ..................................................... 55 16 Start and end dates for major HSLS:09 activities: 2009 ................................................... 56 17 Participation and status of eligible schools: 2009 ............................................................. 61 18 Reasons for refusal: 2009 .................................................................................................. 61 19 Accommodations offered to schools: 2009 ....................................................................... 62 20 Component reductions for converted initial refusal schools: 2009 .................................. 63 21 Participating schools by consent type: 2009 ..................................................................... 65 22 Student response rate by consent type: 2009 .................................................................... 66 23 Student eligibility and questionnaire incapability rates: 2009 .......................................... 67 24 Accommodations for participating students: 2009 ........................................................... 67 25 School participation by test mode: 2009 ........................................................................... 68 26 Student completions by test mode: 2009 .......................................................................... 69

*HSLS:09 Base-Year Data File Documentation xvii*

*List of Tables*

**Table Page** 27 Summary of parent incentive experiment results: 2009 ................................................... 72 28 Parent interview mode, by interview form, language, and consent type: 2009 ................ 73 29 Parent cases requiring intensive tracing, by interview completeness, language, and

in-school consent type: 2009 ............................................................................................ 74 30 Locate and response rates for parent cases, by tracing efforts: 2009 ................................ 75 31 Number of calls to parent sample members, by interview mode, completeness,

language, consent type, and student response status: 2009 .............................................. 76 32 HSLS:09 school sample size and participation yield by type and locale ......................... 81 33 Student response rates by school type: 2009 .................................................................... 82 34 Student response rates by student characteristics: 2009 ................................................... 82 35 Student interview mode of response, by school type: 2009 .............................................. 83 36 Parent response rates by school type: 2009 ...................................................................... 83 37 Parent response rates by student characteristics: 2009 ..................................................... 84 38 Parent interview response phases, by interview form, language, and consent type:

2009 ................................................................................................................................... 85 39 Students’ mathematics teacher participation rates by school type: 2009 ......................... 85 40 Students’ mathematics teacher participation rates by student characteristics: 2009 ........ 86 41 Students’ science teacher participation rates by school type: 2009 .................................. 86 42 Students’ science teacher participation rates by student characteristics: 2009 ................. 87 43 Students’ school administrator participation rates by school type: 2009 ......................... 87 44 Students’ school administrator participation rates by student characteristics: 2009 ........ 88 45 Students’ school counselor participation rates by school type: 2009 ............................... 88 46 Students’ school counselor participation rates by student characteristics: 2009 .............. 89 47 School administrator and counselor response rates: 2009 ................................................ 89 48 Teacher linkage counts by mathematics and science ........................................................ 95 49 Course linkage counts by mathematics and science ......................................................... 96 50 Mathematics teacher link by mathematics course link ..................................................... 97 51 Science teacher link by science course link ...................................................................... 97 52 Major text strings: 2009 .................................................................................................... 99 53 Expert coder results for major recoding, by mode of administration: 2009 ................... 100 54 Expert coder results for major upcoding: 2009 ............................................................... 100 55 Results of adjudication of major upcoding: 2009 ........................................................... 101 56 Final major codes in data file: 20091 .............................................................................. 101 57 Occupation text strings : 2009 ........................................................................................ 103

*xviii HSLS:09 Base-Year Data File Documentation*

*List of Tables*

**Table Page** 58 Expert coder results for parent occupation recoding, by mode of administration:

2009 ................................................................................................................................. 103 59 Expert coder results for parent occupation upcoding: 2009 ........................................... 104 60 Results of adjudication parent occupation upcoding: 2009 ............................................ 104 61 Final occupation codes in data file1: 2009 ...................................................................... 105 62 Expert coder results for student job at age 30 upcoding: 2009 ....................................... 106 63 Results of secondary school coding: 2009 ...................................................................... 107 64 Summary information for student scales ........................................................................ 109 65 Summary information for teacher scales ........................................................................ 111 66 Summary information for school counselor scales ......................................................... 113 67 Summary information for school administrator scale ..................................................... 113 68 Summary of HSLS:09 base-year number and percent of student questionnaire

completers with contextual data: 2009 ........................................................................... 116 69 HSLS:09 analytic weights ............................................................................................... 116 70 School-level participation categories .............................................................................. 119 71 Weighted response rate and average nonresponse adjustment by school

characteristics .................................................................................................................. 121 72 Average calibration adjustments, weight sums, and unequal weighting effect by

school characteristics ...................................................................................................... 123 73 Study-eligible student participation categories ............................................................... 125 74 Average calibration adjustments, weight sums, and unequal weighting effect by

school and student characteristics ................................................................................... 127 75 Sample size and percentage of cases, by HSLS:09 respondent group ............................ 128 76 Summary statistics for HSLS:09 contextual analytic weights by school type ................ 130 77 Average design effects (*deff*) and root design effects (*deft*) for school

administrator and counselor data .................................................................................... 136 78 Average design effects (*deff*) and root design effects (*deft*) for student and parent

data .................................................................................................................................. 137 79 Summary statistics for unit nonresponse bias analyses by HSLS:09 analytic

weight .............................................................................................................................. 139 80 School-level questionnaire items with a weighted item response rate below

85 percent ........................................................................................................................ 144 81 Student-level questionnaire items with a weighted item response rate below

85 percent ........................................................................................................................ 148 82 Parent-level questionnaire items with a weighted item response rate below

85 percent ........................................................................................................................ 149

*HSLS:09 Base-Year Data File Documentation xix*

*List of Tables*

**Table Page** 83 Mathematics teacher-level questionnaire items with a weighted item response rate

below 85 percent ............................................................................................................. 153 84 Science teacher-level questionnaire items with a weighted item response rate

below 85 percent ............................................................................................................. 154 85 Frequency distribution of the estimated bias ratios by study instrument ........................ 156 86 Summary statistics for school-level item nonresponse bias analyses ............................. 157 87 Summary statistics for student-level item nonresponse bias analyses ............................ 158 88 Summary statistics for parent-level item nonresponse bias analyses ............................. 159 89 Summary statistics for mathematics teacher-level item nonresponse bias analyses ....... 160 90 Summary statistics for science teacher-level item nonresponse bias analyses ............... 161 91 Variables included in the single-value imputation and number and weighted

percent of items imputed by study instrument ................................................................ 163 92 Imputation order and imputation methods for variables requiring imputation by

study instrument .............................................................................................................. 165 93 Distribution of responding students by parent response status and by availability

of parent responses to calculate SES .............................................................................. 169 B-1. Item forms, item parameters, and their standard errors of the HSLS:09 Mathematics

Assessment items: 2009 ...................................................................................................B-3 B-2. Proportion correct for each of the HSLS:09 Mathematics Assessment items: 2009 ...... B-5 G-1. School-level standard errors and design effects—overall .............................................. G-4 G-2. School-level standard errors and design effects—Public schools .................................. G-5 G-3. School-level standard errors and design effects—Private schools ................................. G-6 G-4. School-level standard errors and design effects—Northeast schools ............................. G-7 G-5. School-level standard errors and design effects—Midwest schools ............................... G-8 G-6. School-level standard errors and design effects—South schools ................................... G-9 G-7. School-level standard errors and design effects—West schools .................................. G-10 G-8. School-level standard errors and design effects—City schools .................................... G-11 G-9. School-level standard errors and design effects—Suburban schools ........................... G-12 G-10. School-level standard errors and design effects—Town schools ................................. G-13 G-11. School-level standard errors and design effects—Rural schools .................................. G-14 G-12. Student standard errors and design effects—overall .................................................... G-15 G-13. Student standard errors and design effects—Public schools ........................................ G-17 G-14. Student standard errors and design effects—Private schools ....................................... G-19 G-15. Student standard errors and design effects—Northeast schools ................................... G-21 G-16. Student standard errors and design effects—Midwest schools..................................... G-23

*xx HSLS:09 Base-Year Data File Documentation*

*List of Tables*

**Table Page** G-17. Student standard errors and design effects—South schools ......................................... G-25 G-18. Student standard errors and design effects—West schools .......................................... G-27 G-19. Student standard errors and design effects—City schools ............................................ G-29 G-20. Student standard errors and design effects—Suburban schools ................................... G-31 G-21. Student standard errors and design effects—Town schools ......................................... G-33 G-22. Student standard errors and design effects—Rural schools .......................................... G-35 G-23. Student standard errors and design effects—Male students ......................................... G-37 G-24. Student standard errors and design effects—Female students ...................................... G-39 G-25. Student standard errors and design effects—Hispanic students ................................... G-41 G-26. Student standard errors and design effects—Asian students ........................................ G-43 G-27. Student standard errors and design effects—Black students ........................................ G-45 G-28. Student standard errors and design effects—White students ........................................ G-47 G-29. Student standard errors and design effects—Multiracial students ................................ G-49 G-30. Student standard errors and design effects—Low percentile SES students ................. G-51 G-31. Student standard errors and design effects—Middle percentile SES students ............. G-53 G-32. Student standard errors and design effects—High percentile SES students ................. G-55 G-33. Parent standard errors and design effects—overall ...................................................... G-57 G-34. Parent standard errors and design effects—Public schools .......................................... G-59 G-35. Parent standard errors and design effects—Private schools ......................................... G-61 G-36. Parent standard errors and design effects—Northeast schools ..................................... G-63 G-37. Parent standard errors and design effects—Midwest schools....................................... G-65 G-38. Parent standard errors and design effects—South schools ........................................... G-67 G-39. Parent standard errors and design effects—West schools ............................................ G-69 G-40. Parent standard errors and design effects—City schools .............................................. G-71 G-41. Parent standard errors and design effects—Suburban schools ..................................... G-73 G-42. Parent standard errors and design effects—Town schools ........................................... G-75 G-43. Parent standard errors and design effects—Rural schools ............................................ G-77 G-44. Parent standard errors and design effects—Male students ........................................... G-79 G-45. Parent standard errors and design effects—Female students ........................................ G-81 G-46. Parent standard errors and design effects—Hispanic students ..................................... G-83 G-47. Parent standard errors and design effects—Asian students .......................................... G-85 G-48. Parent standard errors and design effects—Black students .......................................... G-87 G-49. Parent standard errors and design effects—White students .......................................... G-89

*HSLS:09 Base-Year Data File Documentation xxi*

*List of Tables*

**Table Page** G-50. Parent standard errors and design effects—Multiracial students .................................. G-91 G-51. Parent standard errors and design effects—Low percentile SES students ................... G-93 G-52. Parent standard errors and design effects—Middle percentile SES students ............... G-95 G-53. Parent standard errors and design effects—High percentile SES students ................... G-97 H-1. Unit nonresponse bias before and after adjustments were applied to the school

base weights for selected variables ................................................................................. H-4 H-2. Unit nonresponse bias before and after adjustments were applied to the student

base weights for selected variables ................................................................................. H-7 H-3. Unit nonresponse bias before and after adjustments were applied to the home-life

base weights for selected variables ............................................................................... H-11 H-4. Unit nonresponse bias before and after adjustments were applied to the science

course enrollee base weights for selected variables ...................................................... H-15 H-5. Unit nonresponse bias before and after adjustments were applied to the

mathematics course enrollee base weights for selected variables ................................ H-19 H-6. Comparison of item respondents and nonrespondents for A1VANDALISM

(Frequency of vandalism at this school) by select sample school characteristics using W1SCHOOL weight ........................................................................................... H-24 H-7. Comparison of item respondents and nonrespondents for A1MTHSTREQ (How

mathematics course(s) required for grad compare with state requirements) by select sample school characteristics, using W1SCHOOL weight ................................ H-25 H-8. Comparison of item respondents and nonrespondents for A1YRSHSTCHR

(Principal’s years of secondary teaching experience) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-26 H-9. Comparison of item respondents and nonrespondents for A1HRSTUDENT

(Hours/week spent meeting with students) by select sample school characteristics, using W1SCHOOL weight ........................................................................................... H-27 H-10. Comparison of item respondents and nonrespondents for A1HRPARENT

(Hours/week spent talking and meeting with parents) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-28 H-11. Comparison of item respondents and nonrespondents for A1HREXTMGMNT

(Hours/week spent on external school management) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-29 H-12. Comparison of item respondents and nonrespondents for A1HRMONITOR

(Hours/week spent monitoring hallways/campus/lunchroom) by select sample school characteristics, using W1SCHOOL weight ....................................................... H-30 H-13. Comparison of item respondents and nonrespondents for A1BAMAJ2 (Principals

major for bachelor’s degree 2-digit CIP code) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-31

*xxii HSLS:09 Base-Year Data File Documentation*

*List of Tables*

**Table Page** H-14. Comparison of item respondents and nonrespondents for A1HRINTMGMNT

(Hours/week spent on internal school management) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-32 H-15. Comparison of item respondents and nonrespondents for A1HRTEACHERS

(Hours/week spent working with teachers on instructional issues) by select sample school characteristics, using W1SCHOOL weight ....................................................... H-33 H-16. Comparison of item respondents and nonrespondents for A1TRANSFRALT

(Percent of 08–09 students transferred out to an alternative program/school) by select sample school characteristics, using W1SCHOOL weight ................................ H-34 H-17. Comparison of item respondents and nonrespondents for A1VBLOCKMINS

(Length of block-scheduled vocational/technical courses) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-35 H-18. Comparison of item respondents and nonrespondents for A1HRDISCIPLN

(Hours/week spent on student discipline/attendance) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-36 H-19. Comparison of item respondents and nonrespondents for A1BULLY (Frequency

of student bullying at this school) by select sample school characteristics, using W1SCHOOL weight ..................................................................................................... H-37 H-20. Comparison of item respondents and nonrespondents for A1CONFLICT

(Frequency of physical conflicts among students at this school) by select sample school characteristics, using W1SCHOOL weight ....................................................... H-38 H-21. Comparison of item respondents and nonrespondents for A1HRPAPERWK

(Hours/week spent on paperwork required by authorities) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-39 H-22. Comparison of item respondents and nonrespondents for A1CAPACITY (Percent

capacity to which school is filled) by select sample school characteristics, using W1SCHOOL weight ..................................................................................................... H-40 H-23. Comparison of item respondents and nonrespondents for A1RETURN09 (Percent

of 9th-graders enrolled in this school Sept 2008 returned Sept 2009) by select sample school characteristics, using W1SCHOOL weight........................................... H-41 H-24. Comparison of item respondents and nonrespondents for A1HSSUBJECT (Main

subject principal taught at high school level) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-42 H-25. Comparison of item respondents and nonrespondents for A1ADA (Average daily attendance percentage for students) by select sample school characteristics, using W1SCHOOL weight ..................................................................................................... H-43 H-26. Comparison of item respondents and nonrespondents for A1NOMTHO (School

offers no mathematics course through some other means) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-44

*HSLS:09 Base-Year Data File Documentation xxiii*

*List of Tables*

**Table Page** H-27. Comparison of item respondents and nonrespondents for A1OFFCLCAPIB

(School offers calculus IB through some other means) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-45 H-28. Comparison of item respondents and nonrespondents for A1STARTDEG

(Principal’s highest degree started but not completed (if any)) by select sample school characteristics, using W1SCHOOL weight ....................................................... H-46 H-29. Comparison of item respondents and nonrespondents for A1OFFMPSCIA

(School offers computer science AP (A) through some other means) by select sample school characteristics, using W1SCHOOL weight........................................... H-47 H-30. Comparison of item respondents and nonrespondents for A1OFFCMPSCIB

(School offers computer science AP (AB) through some other means) by select sample school characteristics, using W1SCHOOL weight........................................... H-48 H-31. Comparison of item respondents and nonrespondents for A1OFFANGEOM

(School offers analytic geometry through some other means) by select sample school characteristics, using W1SCHOOL weight ....................................................... H-49 H-32. Comparison of item respondents and nonrespondents for A1OFFCLCAPBC

(School offers calculus AP (BC) through some other means) by select sample school characteristics, using W1SCHOOL weight ....................................................... H-50 H-33. Comparison of item respondents and nonrespondents for A1NOSCIO (School

offers no science course through some other means) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-51 H-34. Comparison of item respondents and nonrespondents for A1OFFALG3 (School

offers algebra III through some other means) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-52 H-35. Comparison of item respondents and nonrespondents for A1OFFINTSCI2

(School offers integrated science II or above through some other means) by select sample school characteristics, using W1SCHOOL weight........................................... H-53 H-36. Comparison of item respondents and nonrespondents for A1OFFSTATSAP

(School offers statistics AP through some other means) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-54 H-37. Comparison of item respondents and nonrespondents for A1OFFENVAP (School

offers environmental science AP through some other means) by select sample school characteristics, using W1SCHOOL weight ....................................................... H-55 H-38. Comparison of item respondents and nonrespondents for A12YRDEGREE (Percent of 08–09 12th-graders who went on to 2-year institution) by select sample school characteristics, using W1SCHOOL weight........................................... H-56 H-39. Comparison of item respondents and nonrespondents for A1OFFINTSCI1

(School offers integrated science I through some other means) by select sample school characteristics, using W1SCHOOL weight ....................................................... H-57

*xxiv HSLS:09 Base-Year Data File Documentation*

*List of Tables*

**Table Page** H-40. Comparison of item respondents and nonrespondents for A14YRDEGREE

(Percent of 08–09 12th-graders who went on 4-year degree-granting institution) by select sample school characteristics, using W1SCHOOL weight ........................... H-58 H-41. Comparison of item respondents and nonrespondents for A1OFFOTHPSCI

(School offers an other physical science through some other means) by select sample school characteristics, using W1SCHOOL weight........................................... H-59 H-42. Comparison of item respondents and nonrespondents for A1MILITARY (Percent

of 08–09 12th-graders who joined military) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-60 H-43. Comparison of item respondents and nonrespondents for A1OFFINTMTH2

(School offers integrated mathematics II or above through some other means) by select sample school characteristics, using W1SCHOOL weight ................................ H-61 H-44. Comparison of item respondents and nonrespondents for A1OFFSTATS (School

offers statistics or probability through some other means) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-62 H-45. Comparison of item respondents and nonrespondents for A1WORK (Percent of 08–09 12th-graders who entered the workforce) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-63 H-46. Comparison of item respondents and nonrespondents for A1OFFTECH (School

offers principles of technology through some other means) by select sample school characteristics, using W1SCHOOL weight ....................................................... H-64 H-47. Comparison of item respondents and nonrespondents for A1OFFINTMTH1

(School offers integrated mathematics I through some other means) by select sample school characteristics, using W1SCHOOL weight........................................... H-65 H-48. Comparison of item respondents and nonrespondents for A1DIDOTHER (Percent

of 08–09 12th-graders who did something else) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-66 H-49. Comparison of item respondents and nonrespondents for A1MSSUBJECT (Main

subject principal taught at middle school level) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-67 H-50. Comparison of item respondents and nonrespondents for A1OBLOCKMINS

(Length of other block-scheduled courses) by select sample school characteristics, using W1SCHOOL weight ........................................................................................... H-68 H-51. Comparison of item respondents and nonrespondents for A1OFFADVPHYS

(School offers advanced physics/phys II/AP/IB through some other means) by select sample school characteristics, using W1SCHOOL weight ................................ H-69 H-52. Comparison of item respondents and nonrespondents for A1OFFOTHESCI

(School offers an other Earth or environmental science through some other means) by select sample school characteristics, using W1SCHOOL weight ............... H-70

*HSLS:09 Base-Year Data File Documentation xxv*

*List of Tables*

**Table Page** H-53. Comparison of item respondents and nonrespondents for A1OFFGENSCI (School offers general science through some other means) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-71 H-54. Comparison of item respondents and nonrespondents for A1OFFOTHBIO

(School offers an other biological science through some other means) by select sample school characteristics, using W1SCHOOL weight........................................... H-72 H-55. Comparison of item respondents and nonrespondents for A1OFFLSCI (School

offers life science through some other means) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-73 H-56. Comparison of item respondents and nonrespondents for A1OFFCMPSCI (School offers computer science through some other means) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-74 H-57. Comparison of item respondents and nonrespondents for A1OFFALGP1P2

(School offers algebra I part 1 and part 2 through some other means) by select sample school characteristics, using W1SCHOOL weight........................................... H-75 H-58. Comparison of item respondents and nonrespondents for A1CHOICEOTHR

(School participates in another public school choice program) by select sample school characteristics, using W1SCHOOL weight ....................................................... H-76 H-59. Comparison of item respondents and nonrespondents for A1CHOICEIN (Students

can enroll in school or another school within district) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-77 H-60. Comparison of item respondents and nonrespondents for A1CHOICEOUT

(Students can enroll in public school in another district at no tuition cost) by select sample school characteristics, using W1SCHOOL weight........................................... H-78 H-61. Comparison of item respondents and nonrespondents for A1CHOICEPRIV

(Students can enroll in a private school using state/district funds) by select sample school characteristics, using W1SCHOOL weight ....................................................... H-79 H-62. Comparison of item respondents and nonrespondents for A1CHOICESCH

(Students from other districts can enroll in school at no tuition cost) by select sample school characteristics, using W1SCHOOL weight........................................... H-80 H-63. Comparison of item respondents and nonrespondents for A1OFFRMTH (School

offers review or remedial mathematics through some other means) by select sample school characteristics, using W1SCHOOL weight........................................... H-81 H-64. Comparison of item respondents and nonrespondents for A1FILLMTH (Ease of

filling high school mathematics teaching vacancies) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-82 H-65. Comparison of item respondents and nonrespondents for A1OFFADVCHEM (School offers advanced chemistry/chem II/AP/IB thru some other means) by select sample school characteristics, using W1SCHOOL weight ................................ H-83

*xxvi HSLS:09 Base-Year Data File Documentation*

*List of Tables*

**Table Page** H-66. Comparison of item respondents and nonrespondents for A1OFFADVBIO

(School offers advanced biology/bio II/AP/IB through some other means) by select sample school characteristics, using W1SCHOOL weight ................................ H-84 H-67. Comparison of item respondents and nonrespondents for A1OFFERTHSCI

(School offers Earth science through some other means) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-85 H-68. Comparison of item respondents and nonrespondents for A1HROTH (Hours/week

spent on other activities) by select sample school characteristics, using W1SCHOOL weight ..................................................................................................... H-86 H-69. Comparison of item respondents and nonrespondents for A1OFFENVSCI (School

offers environmental science through some other means) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-87 H-70. Comparison of item respondents and nonrespondents for A1OFFCLCAPAB

(School offers calculus AP (AB) through some other means) by select sample school characteristics, using W1SCHOOL weight ....................................................... H-88 H-71. Comparison of item respondents and nonrespondents for A1FILLSCI (Ease of filling high school science teaching vacancies) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-89 H-72. Comparison of item respondents and nonrespondents for A1OFFPREALG

(School offers pre-algebra through some other means) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-90 H-73. Comparison of item respondents and nonrespondents for A1OFFTRIG (School offers trigonometry through some other means) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-91 H-74. Comparison of item respondents and nonrespondents for A1AYPYR (Year of

AYP improvement as of 09-10 school year) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-92 H-75. Comparison of item respondents and nonrespondents for A1OFFANATOMY

(School offers anatomy or physiology through some other means) by select sample school characteristics, using W1SCHOOL weight........................................... H-93 H-76. Comparison of item respondents and nonrespondents for A1OFFCLC (School

offers calculus through some other means) by select sample school characteristics, using W1SCHOOL weight ........................................................................................... H-94 H-77. Comparison of item respondents and nonrespondents for A1OFFPHYSCI (School offers physical science through some other means) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-95 H-78. Comparison of item respondents and nonrespondents for A1OFFPHYS1 (School

offers physics I through some other means) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-96

*HSLS:09 Base-Year Data File Documentation xxvii*

*List of Tables*

**Table Page** H-79. Comparison of item respondents and nonrespondents for A1HRTEACHING

(Hours/week spent on principal’s own teaching assignments) by select sample school characteristics, using W1SCHOOL weight ....................................................... H-97 H-80. Comparison of item respondents and nonrespondents for A1OFFALG2 (School

offers algebra II through some other means) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-98 H-81. Comparison of item respondents and nonrespondents for A1OFFALG1 (School

offers algebra I through some other means) by select sample school characteristics, using W1SCHOOL weight .................................................................. H-99 H-82. Comparison of item respondents and nonrespondents for A1OFFGEOM (School

offers geometry through some other means) by select sample school characteristics, using W1SCHOOL weight ................................................................ H-100 H-83. Comparison of item respondents and nonrespondents for A1OFFCHEM1 (School

offers chemistry I through some other means) by select sample school characteristics, using W1SCHOOL weight ................................................................ H-101 H-84. Comparison of item respondents and nonrespondents for A1OFFBIO1 (School

offers biology I through some other means) by select sample school characteristics, using W1SCHOOL weight ................................................................ H-102 H-85. Comparison of item respondents and nonrespondents for S1ESTIN (Estimated

cost of 1-year tuition/fees at public 4-year college in students state) by select sample school characteristics, using W1STUDENT weight ...................................... H-103 H-86. Comparison of item respondents and nonrespondents for S1ESTCONF (How

confident student is in estimate given cost of public 4-year in-state college) by select sample school characteristics, using W1STUDENT weight ............................ H-104 H-87. Comparison of item respondents and nonrespondents for S1ESTFEE (Estimated

tuition/fees given for public 4-year in-state college includes room/board) by select sample school characteristics, using W1STUDENT weight ...................................... H-105 H-88. Comparison of item respondents and nonrespondents for S1ASIANOR (Student’s

Asian sub-group) by select sample school characteristics, using W1STUDENT weight .......................................................................................................................... H-106 H-89. Comparison of item respondents and nonrespondents for S1COSTIN (Cost of 1

year’s tuition/fees at specific 4-year in-state college) by select sample school characteristics, using W1STUDENT weight .............................................................. H-107 H-90. Comparison of item respondents and nonrespondents for S1COSTPRV (Cost of 1

year’s tuition/fees at specific private college) by select sample school characteristics, using W1STUDENT weight .............................................................. H-108 H-91. Comparison of item respondents and nonrespondents for S1FEEPRV (Cost of

tuition/fees given for private college includes room and board) by select sample school characteristics, using W1STUDENT weight ................................................... H-109

*xxviii HSLS:09 Base-Year Data File Documentation*

*List of Tables*

**Table Page** H-92. Comparison of item respondents and nonrespondents for S1FEEIN (Cost of tuition/fees given for 4-year in-state college includes room/board) by select sample school characteristics, using W1STUDENT weight ...................................... H-110 H-93. Comparison of item respondents and nonrespondents for S1COSTOUT (Cost of 1

year’s tuition/fees at specific 4-year out-of-state college) by select sample school characteristics, using W1STUDENT weight .............................................................. H-111 H-94. Comparison of item respondents and nonrespondents for S1FEEOUT (Cost

tuition/fee given for 4-year out-of-state college includes room/board) by select sample school characteristics, using W1STUDENT weight ...................................... H-112 H-95. Comparison of item respondents and nonrespondents for P1HHPARREL1 (First

resident parents relationship to 9th-grader) by select sample school characteristics, using W1PARENT weight .......................................................................................... H-113 H-96. Comparison of item respondents and nonrespondents for P1PUBPRV (Type of

postsecondary institution respondent thinks 9th-grader will attend) by select sample school characteristics, using W1PARENT weight ......................................... H-114 H-97. Comparison of item respondents and nonrespondents for P1ACCTPAY (Family

opened account(s) to save for 9th-graders college education) by select sample school characteristics, using W1PARENT weight ..................................................... H-115 H-98. Comparison of item respondents and nonrespondents for P1ENGLISH (English is

regularly spoken in home) by select sample school characteristics, using W1PARENT weight ................................................................................................... H-116 H-99. Comparison of item respondents and nonrespondents for P1COUNTRY2

(Country in which second resident parent was born) by select sample school characteristics, using W1PARENT weight ................................................................. H-117 H-100. Comparison of item respondents and nonrespondents for P1COUNTRY1 (Country in which first resident parent was born) by select sample school characteristics, using W1PARENT weight ................................................................. H-118 H-101. Comparison of item respondents and nonrespondents for P1CHINESE (Chinese

language regularly spoken in home) by select sample school characteristics, using W1PARENT weight ................................................................................................... H-119 H-102. Comparison of item respondents and nonrespondents for P1EUROLANG (Other

European language regularly spoken in home) by select sample school characteristics, using W1PARENT weight ................................................................. H-120 H-103. Comparison of item respondents and nonrespondents for P1FILIPINO (Filipino

language regularly spoken in home) by select sample school characteristics, using W1PARENT weight ................................................................................................... H-121 H-104. Comparison of item respondents and nonrespondents for P1MIDEAST (Middle

Eastern language regularly spoken in home) by select sample school characteristics, using W1PARENT weight ................................................................. H-122

*HSLS:09 Base-Year Data File Documentation xxix*

*List of Tables*

**Table Page** H-105. Comparison of item respondents and nonrespondents for P1OTHRASIAN (Other

Asian language regularly spoken in home) by select sample school characteristics, using W1PARENT weight .......................................................................................... H-123 H-106. Comparison of item respondents and nonrespondents for P1SASIAN (South

Asian language regularly spoken in home) by select sample school characteristics, using W1PARENT weight .......................................................................................... H-124 H-107. Comparison of item respondents and nonrespondents for P1SEASIAN (Southeast

Asian language regularly spoken in home) by select sample school characteristics, using W1PARENT weight .......................................................................................... H-125 H-108. Comparison of item respondents and nonrespondents for P1SPANISH (Spanish is

regularly spoken in home) by select sample school characteristics, using W1PARENT weight ................................................................................................... H-126 H-109. Comparison of item respondents and nonrespondents for P1OTHRLANG (Other

language regularly spoken in home) by select sample school characteristics, using W1PARENT weight ................................................................................................... H-127 H-110. Comparison of item respondents and nonrespondents for P1LANG9 (Language

9th-grader usually speaks to respondent in home) by select sample school characteristics, using W1PARENT weight ................................................................. H-128 H-111. Comparison of item respondents and nonrespondents for P1HISPOR2

(Spouse/partner/second resident parent is Mexican or other Hispanic/Latino) by select sample school characteristics, using W1PARENT weight ............................... H-129 H-112. Comparison of item respondents and nonrespondents for P1RSPLANG

(Language respondent usually speaks to 9th-grader in home) by select sample school characteristics, using W1PARENT weight ..................................................... H-130 H-113. Comparison of item respondents and nonrespondents for P1USYR1 (Year

respondent/first resident parent came to United States to stay) by select sample school characteristics, using W1PARENT weight ..................................................... H-131 H-114. Comparison of item respondents and nonrespondents for P1SAVEDPAY

(Amount currently set aside for 9th-graders future educational needs) by select sample school characteristics, using W1PARENT weight ......................................... H-132 H-115. Comparison of item respondents and nonrespondents for P1HISPOR1

(Respondent/first resident parent is Mexican or other Hispanic) by select sample school characteristics, using W1PARENT weight ..................................................... H-133 H-116. Comparison of item respondents and nonrespondents for P1INCOME (Household

income in 2007–continuous form) by select sample school characteristics, using W1PARENT weight ................................................................................................... H-134 H-117. Comparison of item respondents and nonrespondents for P1TUITION

(Respondent has info on cost of tuition/fees at specific public in-state institution) by select sample school characteristics, using W1PARENT weight .......................... H-135

*xxx HSLS:09 Base-Year Data File Documentation*

*List of Tables*

**Table Page** H-118. Comparison of item respondents and nonrespondents for P1DIFSCHLNG

(Difficulty joining in 9th-graders school events because speaks non-English) by select sample school characteristics, using W1PARENT weight ............................... H-136 H-119. Comparison of item respondents and nonrespondents for P1USYR2 (Year

spouse/partner/second resident parent came to United States to stay) by select sample school characteristics, using W1PARENT weight ......................................... H-137 H-120. Comparison of item respondents and nonrespondents for P1INOUTST (Whether

respondent thinks 9th-grader will attend in-state or out-of-state public institution) by select sample school characteristics, using W1PARENT weight .......................... H-138 H-121. Comparison of item respondents and nonrespondents for P1ESTIN (Estimate of cost of 1 years tuition/fees at public 4-year in-state institution) by select sample school characteristics, using W1PARENT weight ..................................................... H-139 H-122. Comparison of item respondents and nonrespondents for P1ESTCONF

(Confidence in estimate of 1 year’s cost for public 4-year in-state institution) by select sample school characteristics, using W1PARENT weight ............................... H-140 H-123. Comparison of item respondents and nonrespondents for P1ESTFEE (Estimate of

cost at public 4-year in-state institution includes room/board) by select sample school characteristics, using W1PARENT weight ..................................................... H-141 H-124. Comparison of item respondents and nonrespondents for P1HHPARREL2

(Second resident parents relationship to 9th-grader) by select sample school characteristics, using W1PARENT weight ................................................................. H-142 H-125. Comparison of item respondents and nonrespondents for P1REPEATGK (9th-

grader repeated kindergarten) by select sample school characteristics, using W1PARENT weight ................................................................................................... H-143 H-126. Comparison of item respondents and nonrespondents for P1REPEATG1 (9th-

grader repeated 1st grade) by select sample school characteristics, using W1PARENT weight ................................................................................................... H-144 H-127. Comparison of item respondents and nonrespondents for P1REPEATG9 (9th-

grader repeated 9th grade) by select sample school characteristics, using W1PARENT weight ................................................................................................... H-145 H-128. Comparison of item respondents and nonrespondents for P1USYR9 (Year 9th-

grader came to the United States to stay) by select sample school characteristics, using W1PARENT weight .......................................................................................... H-146 H-129. Comparison of item respondents and nonrespondents for P1USGRADE (Grade

level 9th-grader was placed in when started school in United States) by select sample school characteristics, using W1PARENT weight ......................................... H-147 H-130. Comparison of item respondents and nonrespondents for P1ELLNOW (Whether

9th-grader currently in English language learners program) by select sample school characteristics, using W1PARENT weight ..................................................... H-148

*HSLS:09 Base-Year Data File Documentation xxxi*

*List of Tables*

**Table Page** H-131. Comparison of item respondents and nonrespondents for P1COUNTRY9

(Country in which 9th-grader was born) by select sample school characteristics, using W1PARENT weight .......................................................................................... H-149 H-132. Comparison of item respondents and nonrespondents for P1HHOTHR (Where

9th-grader lives when not living with respondent) by select sample school characteristics, using W1PARENT weight ................................................................. H-150 H-133. Comparison of item respondents and nonrespondents for P1COSTIN (Cost of

tuition/fees at public 4-year in-state institution) by select sample school characteristics, using W1PARENT weight ................................................................. H-151 H-134. Comparison of item respondents and nonrespondents for P1FEEIN (Cost of

tuition/fees at public 4-year in-state institution includes room/board) by select sample school characteristics, using W1PARENT weight ......................................... H-152 H-135. Comparison of item respondents and nonrespondents for P1QHELP1 (9th-grader

helped respondent complete questionnaire) by select sample school characteristics, using W1PARENT weight ................................................................. H-153 H-136. Comparison of item respondents and nonrespondents for P1QHELP2 (Other family member helped respondent complete questionnaire) by select sample school characteristics, using W1PARENT weight ..................................................... H-154 H-137. Comparison of item respondents and nonrespondents for P1QHELP3

(Respondent’s friend helped respondent complete questionnaire) by select sample school characteristics, using W1PARENT weight ..................................................... H-155 H-138. Comparison of item respondents and nonrespondents for P1QHELP4 (Person

helped respondent complete questionnaire—other) by select sample school characteristics, using W1PARENT weight ................................................................. H-156 H-139. Comparison of item respondents and nonrespondents for P1ASIANOR2 (Asian origin of spouse/partner/ second resident parent) by select sample school characteristics, using W1PARENT weight ................................................................. H-157 H-140. Comparison of item respondents and nonrespondents for P1COSTPRV (Cost of tuition/fees at private 4-year in-state institution) by select sample school characteristics, using W1PARENT weight ................................................................. H-158 H-141. Comparison of item respondents and nonrespondents for P1FEEPRV (Cost of tuition/fees at private 4-year in-state institution includes room/board) by select sample school characteristics, using W1PARENT weight ......................................... H-159 H-142. Comparison of item respondents and nonrespondents for P1ASIANOR1 (Asian

origin of respondent/first resident parent) by select sample school characteristics, using W1PARENT weight .......................................................................................... H-160 H-143. Comparison of item respondents and nonrespondents for P1SKIPGK 9th-grader

skipped kindergarten) by select sample school characteristics, using W1PARENT weight .......................................................................................................................... H-161

*xxxii HSLS:09 Base-Year Data File Documentation*

*List of Tables*

**Table Page** H-144. Comparison of item respondents and nonrespondents for P1SKIPG1 (9th-grader

skipped 1st grade:) by select sample school characteristics, using W1PARENT weight .......................................................................................................................... H-162 H-145. Comparison of item respondents and nonrespondents for P1SKIPG8 (9th-grader

skipped 8th grade) by select sample school characteristics, using W1PARENT weight .......................................................................................................................... H-163 H-146. Comparison of item respondents and nonrespondents for P1COSTOUT (Cost of

tuition/fees at private 4-year out-of-state institution) by select sample school characteristics, using W1PARENT weight ................................................................. H-164 H-147. Comparison of item respondents and nonrespondents for P1FEEOUT (Cost

tuition/fees at private 4-year out-of-state institution includes room/board) by select sample school characteristics, using W1PARENT weight ............................... H-165 H-148. Comparison of item respondents and nonrespondents for N1COURSE (Students

fall 2009 science course-categorized) by select sample school characteristics, using W1SCITCH weight ........................................................................................... H-166 H-149. Comparison of item respondents and nonrespondents for N1GROUP (Science

teacher has students work in small groups) by select sample school characteristics, using W1SCITCH weight ........................................................................................... H-167 H-150. Comparison of item respondents and nonrespondents for N1EVIDENCE (Science

teachers emphasis on evaluating arguments based on evidence) by select sample school characteristics, using W1SCITCH weight ....................................................... H-168 H-151. Comparison of item respondents and nonrespondents for N1ACHIEVE

(Achievement of students in science course compared with average ninth-grader) by select sample school characteristics, using W1SCITCH weight ........................... H-169 H-152. Comparison of item respondents and nonrespondents for N1INTEREST (Science

teachers emphasis on increasing students interest in science) by select sample school characteristics, using W1SCITCH weight ....................................................... H-170 H-153. Comparison of item respondents and nonrespondents for N1TERMS (Science

teachers emphasis on important science terms/facts) by select sample school characteristics, using W1SCITCH weight .................................................................. H-171 H-154. Comparison of item respondents and nonrespondents for N1SKILLS (Science

teachers emphasis on science process/inquiry skills) by select sample school characteristics, using W1SCITCH weight .................................................................. H-172 H-155. Comparison of item respondents and nonrespondents for N1PREPARE (Science

teachers emphasis on preparation for further science study) by select sample school characteristics, using W1SCITCH weight ....................................................... H-173 H-156. Comparison of item respondents and nonrespondents for N1CONCEPTS (Science

teachers emphasis on teaching basic science concepts) by select sample school characteristics, using W1SCITCH weight .................................................................. H-174

*HSLS:09 Base-Year Data File Documentation xxxiii*

*List of Tables*

**Table Page** H-157. Comparison of item respondents and nonrespondents for N1UNPREPPCT

(Percentage of students in science course that are unprepared) by select sample school characteristics, using W1SCITCH weight ....................................................... H-175 H-158. Comparison of item respondents and nonrespondents for N1TEST (Science teachers emphasis on standardized test preparation) by select sample school characteristics, using W1SCITCH weight .................................................................. H-176 H-159. Comparison of item respondents and nonrespondents for N1HISTORY (Science

teachers emphasis on history/nature of science) by select sample school characteristics, using W1SCITCH weight .................................................................. H-177 H-160. Comparison of item respondents and nonrespondents for N1IDEAS (Science

teachers emphasis on effectively communicating science ideas) by select sample school characteristics, using W1SCITCH weight ....................................................... H-178 H-161. Comparison of item respondents and nonrespondents for N1BUSINESS (Science

teachers emphasis on business/industry applications of science) by select sample school characteristics, using W1SCITCH weight ....................................................... H-179 H-162. Comparison of item respondents and nonrespondents for N1SOCIETY (Science

teachers emphasis on relationship between science and tech and society) by select sample school characteristics, using W1SCITCH weight .......................................... H-180 H-163. Comparison of item respondents and nonrespondents for N1ASSIGN (How

science teacher assigns students to small groups) by select sample school characteristics, using W1SCITCH weight .................................................................. H-181 H-164. Comparison of item respondents and nonrespondents for M1ENGCOMP (How

teacher compares boys and girls English or language arts abilities) by select sample school characteristics, using W1MATHTCH weight ..................................... H-182 H-165. Comparison of item respondents and nonrespondents for M1SCICOMP (How

teacher compares boys and girls science abilities) by select sample school characteristics, using W1MATHTCH weight............................................................. H-183 H-166. Comparison of item respondents and nonrespondents for M1COURSE (Students

fall 2009 mathematics course-categorized) by select sample school characteristics, using W1MATHTCH weight ..................................................................................... H-184 H-167. Comparison of item respondents and nonrespondents for M1ACHIEVE

(Achievement of students in mathematics course compared with average ninth- grader) by select sample school characteristics, using W1MATHTCH weight ......... H-185 H-168. Comparison of item respondents and nonrespondents for M1GROUP

(Mathematics teacher has students work in small groups) by select sample school characteristics, using W1MATHTCH weight............................................................. H-186 H-169. Comparison of item respondents and nonrespondents for M1UNPREPPCT

(Percentage of students in mathematics course that are unprepared) by select sample school characteristics, using W1MATHTCH weight ..................................... H-187

*xxxiv HSLS:09 Base-Year Data File Documentation*

*List of Tables*

**Table Page** H-170. Comparison of item respondents and nonrespondents for M1COMPUTE

(Mathematics teachers emphasis on speedy/accurate computations) by select sample school characteristics, using W1MATHTCH weight ..................................... H-188 H-171. Comparison of item respondents and nonrespondents for M1PREPARE

(Mathematics teachers emphasis on preparation for further mathematics study) by select sample school characteristics, using W1MATHTCH weight ........................... H-189 H-172. Comparison of item respondents and nonrespondents for M1IDEAS (Mathematics

teachers emphasis on connecting mathematics ideas) by select sample school characteristics, using W1MATHTCH weight............................................................. H-190 H-173. Comparison of item respondents and nonrespondents for M1REASON

(Mathematics teachers emphasis on reasoning mathematically) by select sample school characteristics, using W1MATHTCH weight ................................................. H-191 H-174. Comparison of item respondents and nonrespondents for M1BUSINESS

(Mathematics teachers emphasis on business/industry applications of mathematics) by select sample school characteristics, using W1MATHTCH weight .......................................................................................................................... H-192 H-175. Comparison of item respondents and nonrespondents for M1INTEREST

(Mathematics teachers emphasis on increasing students interest in mathematics) by select sample school characteristics, using W1MATHTCH weight ...................... H-193 H-176. Comparison of item respondents and nonrespondents for M1ALGORITHM

(Mathematics teachers emphasis on teaching mathematics algorithms/ procedures) by select sample school characteristics, using W1MATHTCH weight ...................... H-194 H-177. Comparison of item respondents and nonrespondents for M1COMPSKILLS

(Mathematics teachers emphasis on developing computational skills) by select sample school characteristics, using W1MATHTCH weight ..................................... H-195 H-178. Comparison of item respondents and nonrespondents for M1PROBLEM

(Mathematics teachers emphasis on developing problem solving skills) by select sample school characteristics, using W1MATHTCH weight ..................................... H-196 H-179. Comparison of item respondents and nonrespondents for M1TEST (Mathematics

teachers emphasis on standardized test preparation) by select sample school characteristics, using W1MATHTCH weight............................................................. H-197 H-180. Comparison of item respondents and nonrespondents for M1EXPLAIN

(Mathematics teachers emphasis on effectively explaining mathematics ideas) by select sample school characteristics, using W1MATHTCH weight ........................... H-198 H-181. Comparison of item respondents and nonrespondents for M1HISTORY

(Mathematics teachers emphasis on history and nature of mathematics) by select sample school characteristics, using W1MATHTCH weight ..................................... H-199 H-182. Comparison of item respondents and nonrespondents for M1LOGIC

(Mathematics teachers emphasis on logical structure of mathematics) by select sample school characteristics, using W1MATHTCH weight ..................................... H-200

*HSLS:09 Base-Year Data File Documentation xxxv*

*List of Tables*

**Table Page** H-183. Comparison of item respondents and nonrespondents for M1CONCEPTS

(Mathematics teachers emphasis on teaching mathematics concepts) by select sample school characteristics, using W1MATHTCH weight ..................................... H-201 H-184. Comparison of item respondents and nonrespondents for M1ASSIGN (How

mathematics teacher assigns students to small groups) by select sample school characteristics, using W1MATHTCH weight............................................................. H-202 I-1. Details of imputation procedures for variables requiring imputation ............................... I-4 I-2. Weighted distribution of imputed variables before and after imputation ......................... I-7 I-3. Covariates evaluated for inclusion in the *theta* and *sem* multiple imputation by

whether the variable was retained for the final model .................................................... I-15 J-1. Comparison of quintiles for X1SES with X1SES\_U ....................................................... J-6 J-2. Distribution of responding students by SES imputation group ....................................... J-8 J-3. Percent of cases with imputed SES component variables ............................................... J-9 J-4. Description of students within each imputation model.................................................. J-11 J-5. HSLS:09 variables evaluated for the SES multiple imputation models by whether

or not the variable was included in the respective model .............................................. J-12

*xxxvi HSLS:09 Base-Year Data File Documentation*

**List of Figures**

**Figure Page** 1 Longitudinal design for the NCES high school cohorts: 1972–2015 ................................. 3 2 Longitudinal design for the HSLS:09 ninth-grade cohort: 2009–21 .................................. 7 3 HSLS:09 base-year student survey conceptual map ......................................................... 11 4 Frequently used data collection acronyms: 2009 .............................................................. 56 5 Endorsing organizations: 2009 ......................................................................................... 57 6 Session administrator training agenda: 2009 .................................................................... 64 7 Example SAS-SUDAAN code to produce mean and linearization standard error ......... 134 8 Example SUDAAN code to produce mean and BRR standard error for a student-

level analysis ................................................................................................................... 134 9 Example STATA code to produce mean and linearization standard error ..................... 134 10 Example STATA code to produce mean and BRR standard error for a student-

level analysis ................................................................................................................... 134 J-1. Unweighted histograms for X1SES and X1SES\_U by school type .............................. J-12

*HSLS:09 Base-Year Data File Documentation xxxvii*

**Chapter 1. Introduction**

**1.1 Overview of the Data File Documentation**

This manual provides guidance and documentation for users of data from the base year of the High School Longitudinal Study of 2009 (HSLS:09). HSLS:09 is sponsored by the National Center for Education Statistics (NCES) of the Institute of Education Sciences, U.S. Department of Education, with additional support from the National Science Foundation. The base-year study was conducted through a contract to RTI International,1

a university-affiliated, nonprofit research organization in North Carolina, in collaboration with its subcontractors, the American Institutes for Research, Horizon Research, Windwalker, Research Support Services, and MPR Associates. This manual contains information about the purposes of the study, the survey instruments, the assessment, the sample design, and the data collection and data processing procedures. The manual provides guidance for understanding and using all components of the base-year study— student questionnaire and mathematics assessment data; questionnaire data from parents; and questionnaire data from mathematics and science teachers, school administrators, and counselors.

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

HSLS:09 base-year data have been made available for public users in two formats—via the eDAT (a web-based application on the NCES server) and through an electronic codebook (ECB) designed to be run in a Microsoft Windows environment on the user’s PC. In addition to the public-use ECB (NCES 2011-334), a restricted-use ECB (NCES 2011-333) is available to licensed users.

Chapter 1 addresses three main topics. First, it supplies an overview of the NCES education longitudinal studies program, thus situating HSLS:09 in the context of the earlier NCES high school cohorts studied in the 1970s, 1980s, 1990s, and 2000s. Second, it introduces HSLS:09 by delineating its principal objectives. Third, it provides an overview of the base-year study design. In subsequent chapters, these additional topics are addressed: instrumentation (chapter 2), sample design (chapter 3), data collection methods and results (chapter 4), data preparation and processing (chapter 5), weighting and estimation (chapter 6), item nonresponse and imputation (chapter 7), and data file structure and contents (chapter 8). Appendixes provide

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

*HSLS:09 Base-Year Data File Documentation 1*

*Chapter 1. Introduction*

additional information, including a hardcopy version of the questionnaires, technical detail concerning sample selection, codebooks for school- and student-level data, and a glossary of terms.

**1.2 Historical Background**

**1.2.1 NCES Secondary Longitudinal Studies Program**

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

NCES (and HSLS:09) are authorized by section 406(b) of the General Education Provision Act (20 U.S.C. 1221e) as amended by the Education Sciences Reform Act of 2002. The Education Sciences Reform Act of 2002 replaced the former Office of Educational Research and Improvement with the Institute of Education Sciences, in which NCES is now housed.

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

Figure 1 includes a temporal presentation of these five longitudinal education studies and highlights their component and comparison points for the time frame 1972–2015. (If HSLS:09 follows the precedent of NELS:88, the terminal interview will take place at around age 26 in the spring or summer of 2021, with postsecondary transcripts collected in the fall of 2021.)

*2 HSLS:09 Base-Year Data File Documentation*

*HSLS:09 Base-Year Data File Documentation 3*

**Figure 1. Longitudinal design for the NCES high school cohorts: 1972–2015**

12+14

**F5** 3212+13 3112+12

**PST** 3012+11

**PST** 2912+10

12+9

**F4** 282712+8

**F4-PST**

**F3-PST** 2612+7 12+6

**F4**

**F3** 252412+5 2312+4

**F3**

**F2-PST**

**F3** 2212+3 12+2

**F2**

**F1**

**F2**

**F2** 2021**HST**

**F2** 1912

**F1**

**BY-SA-P**

**F1-SA-A-D** 11**F3** 12+1

**F2-SA-P-T-A-HST- HSES-D**

**F1-SA-A- HST-D**

**CU-P/S-HST** 18**BY-SA**



**F1-SA-** 17109**BY-SA-P-A**

**F1-SA-T-A-HSES-D**

**BY-SA-P-T-A-L-F**

**P-T-A-C** 16

**BY-SA-P-T-A-C** 15814 7**BY-SA-P-T-A**

13**Year of Data Collection**

NLS:72 HS&B: 12th-grade cohort HS&B:10th-grade cohort NELS:88 ELS:2002  HSLS:09 NLS:72=National Longitudinal Study of the High School Class of 1972

BY=Base year data collection

P=Parent survey

PST=Postsecondary transcript HS&B=High School and Beyond: 1980

F1=1st follow-up data collection

T=Teacher survey

C=Counselor questionnaire NELS:88=National Education Longitudinal Study of 1988

F2=2nd follow-up data collection

A=Administrator survey

P/S=Parent or student ELS:2002=Education Longitudinal Study of 2002

F3=3rd follow-up data collection

L=Library/media center survey

HSES=HS effectiveness study HSLS:09=High School Longitudinal Study of 2009

F4=4th follow-up data collection

F=Facilities checklist

D=Dropout survey F5=5th follow-up data collection

CU=College update HST=High school transcript SA=Student assessment

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

*Chapter 1. Introduction*

*Chapter 1. Introduction*

**1.2.2 National Longitudinal Study of the High School Class of 1972**

The Secondary Longitudinal Studies program began about 40 years ago with the implementation of startup activities for NLS:72.2

NLS:72 was designed to provide longitudinal data for educational policymakers and researchers to link educational experiences in high school with important downstream outcomes such as labor market experiences and postsecondary education enrollment and attainment. With a national probability sample of 19,001 high school seniors from 1,061 public and religious and other private schools, the NLS:72 sample was representative of approximately 3 million high school seniors enrolled in 17,000 U.S. high schools during the spring of the 1971–72 school year. Each member of this cohort was asked to complete a student questionnaire and a cognitive test battery. In addition, administrators at the sample members’ schools were asked to supply information about the schools’ programs, resources, and grading systems, as well as survey data on each student. No parent survey was conducted. However, postsecondary education transcripts were collected in 1984 from the institutions attended by sample members. Five follow-up surveys were completed with this student cohort, with the final data collection taking place in 1986, when the sample members were 14 years removed from high school and approximately 32 years old.

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

**1.2.3 High School and Beyond**

The second in the series of NCES secondary longitudinal studies was launched in 1980. HS&B included one cohort of high school seniors comparable to the NLS:72 sample; however, the study also extended the age span and analytical range of NCES longitudinal studies by surveying a sample of high school sophomores. Base-year data collection took place in the spring term of the 1979–80 academic year with a two-stage probability sample. More than 1,000 schools served as the first-stage units, and 58,000 students within those schools were the second- stage units. Both cohorts of HS&B participants were resurveyed in 1982, 1984, and 1986; the sophomore group also was surveyed in 1992.3

In addition, to better understand the school and home contexts for the sample members, data were collected from teachers (a teacher comment form in the base year asked for teacher perceptions of HS&B sample members), principals, and a

2 For reports on the NLS:72 project, see Riccobono et al. (1981) and Tourangeau et al. (1987). 3 For a summation of the HS&B sophomore cohort study, see Zahs et al. (1995). For more information on HS&B in the high school years, with a focus on the sophomore cohort, see Jones et al. (1983). For further information on HS&B, see the NCES website: http://nces.ed.gov/surveys/hsb/.

*4 HSLS:09 Base-Year Data File Documentation*

*Chapter 1. Introduction*

subsample of parents. High school transcripts were collected for a subsample of sophomore cohort members. As in NLS:72, postsecondary transcripts were collected for both HS&B cohorts; however, the sophomore cohort transcripts cover a much longer time span (to 1993).

With the study design expanded to include a sophomore cohort, HS&B provided critical data on the relationships between early high school experiences and students’ subsequent educational experiences in high school. For the first time, national data were available that showed students’ academic growth over time and how family, community, school, and classroom factors promoted or inhibited student learning. Researchers were able to use data from the extensive battery of achievement tests within the longitudinal study to assess growth in knowledge and cognitive skills over time. Moreover, data were then available to analyze the school experiences of students who later dropped out of high school, and eventually, to investigate their later educational and occupational outcomes. These data became a rich resource for policymakers and researchers over the next decade and provided an empirical base to inform the debates of the educational reform movement that began in the early 1980s.

**1.2.4 National Education Longitudinal Study of 1988**

Much as NLS:72 captured a high school cohort of the 1970s and HS&B captured high school cohorts of the 1980s, NELS:88 was designed to study high school students of the 1990s— but with a baseline measure of their achievement and status, prior to their entry into high school. NELS:88 is an integrated system of data that tracked students from junior high or middle school through secondary and postsecondary education, labor market experiences, and marriage and family formation.

Data collection for NELS:88 was initiated with the eighth-grade class of 1988 in the spring term of the 1987–88 school year. Along with a student survey, NELS:88 included surveys of parents (base year and second follow-up), teachers (base year, first and second follow-ups), and school administrators (base year, first and second follow-ups). The cohort was also surveyed twice after their scheduled high school graduation, in 1994 and 2000.4

4 The entire compass of NELS:88, from its baseline through its final follow-up in 2000, is described in Curtin et al. (2002). Outcomes for the 1988 eighth-grade cohort in 2000 are reported in Ingels et al. (2002). Documentation of the NELS:88 assessment battery is found in Rock and Pollack (1995). The quality of NELS:88 data in the in-school rounds is examined in McLaughlin and Cohen (1997). The sample design is documented in Spencer et al. (1990). Eligibility and exclusion issues are addressed in Ingels (1996). NCES keeps an updated version of the NELS:88 bibliography on its website. The bibliography encompasses both project documentation and research articles, monographs, dissertations, and paper presentations employing NELS:88 data (see http://nces.ed.gov/surveys/nels88/Bibliography.asp).

*HSLS:09 Base-Year Data File Documentation 5*

High school transcripts were collected in the autumn of 1992 and postsecondary transcripts in the autumn of 2000. Through a process of sample freshening, NELS:88 offers three nationally representative cohorts of students: spring-term 8th-, 10th-, and 12th-graders.

*Chapter 1. Introduction*

**1.2.5 Education Longitudinal Study of 2002**

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

ELS:2002 gathers information at multiple levels. In the base year (2002), it obtained information not only from students, but also from students’ parents, teachers, and the administrators (principal and library media center director) of their schools. In the first follow-up (2004), the sample was freshened to represent the senior cohort of 2004 as well as the sophomore cohort of 2002, and high school transcripts were collected as were student questionnaires and tests and school administrator data.

In the second follow-up (2006), when most sample members had been out of high school for 2 years, computer-assisted student questionnaires were administered via the Web or telephone or in person, and data linkages and merges were added to the database, including Scholastic Assessment Test and ACT scores, General Educational Development scores, data from the Central Processing System, information from the Free Application for Federal Student Aid, and information from the National Student Loan Data System, including both federal loan and Pell grant information. A contract has been awarded for collection of third follow-up data in 2012.5

**1.3 High School Longitudinal Study of 2009**

**1.3.1 Overview of the HSLS:09 Design and Objectives**

The longitudinal design of HSLS:09 is set out in figure 2. The HSLS:09 base year took place in the fall term of the 2009–10 school year, with a randomly selected sample of fall-term 9th-graders in more than 900 public and private high schools with both a 9th and an 11th grade.6 Students took a mathematics assessment and survey online. Students’ parents, principals, and mathematics and science teachers as well as the school’s lead counselor completed surveys on the phone or on the Web.7

The first follow-up of HSLS:09 will take place in the spring of 2012 when most sample members will be in the spring of the 11th grade. A postsecondary update (or College Update) will take place in the summer of 2013, to find out about the cohort’s postsecondary plans and decisions. High school transcripts will be collected in the fall of 2013, and a second follow-up will take place in 2015, when most sample members will be 2 years beyond high school graduation.

5 ELS:2002 is documented in Ingels et al. (2007) (http://nces.ed.gov/pubsearch/pubsinfo.asp?/pubid=2008347). A bibliography is maintained on the NCES ELS:2002 website (http://nces.ed.gov/surveys/els2002/bibliography.asp). 6 Types of schools that were excluded from the sample based on the HSLS:09 eligibility definition are described as part of the discussion of the target population (see chapter 3, section 3.2.1). 7 However, an abbreviated paper-and-pencil questionnaire was used in some (779) parent interviews.

*6 HSLS:09 Base-Year Data File Documentation*

*Chapter 1. Introduction*

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

12+8

**3FU-PST**



26

12+7 2512+6 2412+5 2312+4 2212+3

**2FU**  2112+2 2012+1 1912 11

 **CU-P/S-HST** 181710**1FU-AS-P-A-C-D** 169158

**BY-AS-P-T-A-C** 714 132009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021

**Year of Data Collection**

 HSLS:09 HSLS:09=High School Longitudinal Study of 2009

A=Administrator survey BY=Base-year data collection

CU=College update F1=1st follow-up data collection

HST=High school transcript F2=2nd follow-up data collection

PST=Postsecondary transcript F3=3rd follow-up data collection

C=Counselor questionnaire SA=Student assessment (Algebraic Reasoning)

P/S=Parent or student P=Parent survey

D=Dropout survey T=Teacher survey

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

The core research questions for HSLS:09 explore secondary to postsecondary transition plans and the evolution of those plans; the paths into and out of science, technology, engineering, and mathematics; and the educational and social experiences that affect these shifts. (More will be said about research objectives in section 1.3.2 below.)

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

• use of a computer-administered assessment and student questionnaire in a school setting;

• an assessment that focuses on algebraic reasoning;

•

use of computerized (web/computer-assisted telephone interview) parent, teacher, administrator, and counselor questionnaires;

*HSLS:09 Base-Year Data File Documentation 7*

*Chapter 1. Introduction*

• • • • • • •

*8 HSLS:09 Base-Year Data File Documentation*

inclusion of a counselor survey to document school course and program assignment policies and procedures;

starting point in the fall of ninth grade, the traditional beginning of high school;

enhanced emphasis on the dynamics of educational and occupational decision- making;

enhanced emphasis on science, technology, engineering, and mathematics (STEM) trajectories;

follow-up in spring of 11th grade, including follow-up mathematics assessment;

concern with general trends in youth transition, not grade-based specific comparisons with prior spring cohorts of eighth-graders, sophomores, and seniors; and

linkage to selected state administrative data systems and augmentation of selected state public school samples to render them state-representative. At the same time, there are also major points of continuity with all or several of the past studies: • • • • • • • • •

commitment to collecting high school (grades 9–12) transcripts as in HS&B, NELS, and ELS;

a nationally representative school sample with an oversample of private schools and student numbers that are sufficient for subgroup reporting by major race/ethnicity categories, including Asians;

commitment to following the cohort beyond high school;

commitment to identifying and following high school dropouts;

contextual samples of parents as in HS&B, NELS, and ELS;

contextual samples of teachers as in HS&B, NELS, and ELS;

a school administrator survey as in HS&B, NELS, and ELS;

an ability-adaptive assessment battery as in NELS and ELS; and

production of a general purpose dataset that will support a broad range of descriptive and interpretive reporting.

**1.3.2 HSLS:09 Research and Policy Issues**

HSLS:09 provides a link to its predecessor longitudinal studies, which address many of the same issues of transition from high school to postsecondary education and the labor force. At the same time, HSLS:09 brings a new and special emphasis to the study of youth transition by exploring the path that leads students to pursue and persist in courses and careers in the fields of science, technology, engineering, and mathematics.

HSLS:09 measures mathematics achievement gains in the first 3 years of high school, but also will relate tested achievement to students’ choice, access, and persistence—both in

*Chapter 1. Introduction*

mathematics and science courses in high school, and thereafter in the science, technology, engineering, and mathematics pipelines in postsecondary education and in STEM careers. Indeed, the HSLS:09 mathematics assessment serves not just as an outcome measure, but also as a predictor of readiness to proceed into STEM courses and careers.

Additionally, HSLS:09 focuses on students’ decision-making processes. Generally, the study questions students on when, why, and how they make decisions about high school courses and postsecondary options, including what factors, from parental input to considerations of financial aid for postsecondary education, enter into these decisions. Questionnaires focus on factors that motivate students for STEM coursetaking and careers.

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

These patterns of transition cover individual and institutional characteristics. At the individual level the study will look into educational attainment and personal development. In response to policy and scientific issues, data will also be provided on the demographic and background correlates of educational outcomes. At the institutional level, HSLS:09 focuses on school effectiveness issues, including promotion, retention, and curriculum content, structure, and sequencing, especially as these affect students’ choice of, and assignment to, different mathematics and science courses and achievement in these two subject areas.

By collecting extensive information from students, parents, teachers, school counselors, school administrators, and school records, it will be possible to investigate the relationship between home and school factors and academic achievement, interests, and social development at this critical juncture. The school environment is captured primarily through student, teacher, and administrator reports. The extent to which schools are expected to provide special services to selected groups of students to compensate for limitations and poor performance (including special services to assist those lagging in their understanding of mathematics and science) can be examined. Base-year teachers reported on sampled students’ specific classroom environment and supplied information about their own background and training. Moreover, in the base-year and first follow-up parent surveys, the study provides a basis for examining policy issues related to parents’ role in the educational success of their children, including parents’ educational attainment expectations for their children, attitudes toward curricular and postsecondary educational choices, and the correlates of active parental involvement in their children’s educational experiences; these are among the many questions HSLS:09 can address about the home education support system and its interaction with the student and the school.

*HSLS:09 Base-Year Data File Documentation 9*

*Chapter 1. Introduction*

Additionally, because the survey focuses on 9th-graders, it will permit the identification and study of high school dropouts and underwrite trend comparisons with dropouts identified and surveyed in HS&B, NELS:88, and ELS:2002—but especially NELS:88, because both HSLS:09 and NELS:88 allow “early” dropouts (prior to spring of 10th grade) to be identified and studied as well as “late” dropouts in the last 2 years of high school.

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

• • • • • • •

academic (especially in mathematics), social, and interpersonal growth;

transitions from high school to postsecondary education, and from school to work;

students’ choices about, access to, and persistence in mathematics and science courses, majors, and STEM careers;

the characteristics of high schools and postsecondary institutions and their impact on student outcomes;

baccalaureate and sub-baccalaureate attainment;

family formation, including marriage and family development, and how prior experiences in and out of school relate to these decisions, and how marital and parental status affect educational choice, persistence, and attainment; and

the contexts of education, including how minority and at-risk status is associated with education and labor market outcomes.

*10 HSLS:09 Base-Year Data File Documentation*

**1.3.3 HSLS:09 Analysis Files and Systems**

HSLS:09 base-year data are available in two distinct applications: a restricted-use (NCES 2011-333) and a public-use (NCES 2011-334) electronic codebook housed on a DVD; and an online Education Data Application Tool for public use data. Details of file structure and contents across these applications are supplied in chapter 8.

**Chapter 2. Base-Year Instrumentation**

**2.1 Introduction**

**2.1.1 Instrument Development Process and Procedures**

Instrument design for the High School Longitudinal Study of 2009 (HSLS:09) was guided by a theoretical framework or conceptual model. This model (figure 3) takes 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. The study design also reflects the interaction between students and their families, teachers, peers, and community.

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

Social Context and Interpersonal Influences

Student Background and Previous Experiences

• Goals

• Belonging

• Motivation

• Identity

• Utility Value

• Perceived Opportunities and Barriers

• Costs

• Attributions

• Self-Concept

• Self-Efficacy

• Deterrents and Negative Experiences

Value

Expectancy

**Decision**

• Math-Science 9th-Grade Transition

Coursetaking HSLS:09 BY

• Engagement Interview

• Time Use

• College

• Occupations & Careers

Math & Science Classroom Environment

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

The theoretical framework or conceptual model served as the starting point for identifying constructs to be measured. From this framework, broad research domains were identified as relevant, and from each domain, key constructs were drawn. Items that could best measure the constructs were subsequently sought and selected. For example, “student background/ demographic characteristics” constitutes a research domain. Nested within it is the construct of English-language status, which in turn is tapped by specific items (e.g., items asking about whether a language other than English is spoken in the home). It should be noted that many of the constructs are subject-specific (mathematics or science)—for example, mathematics

*HSLS:09 Base-Year Data File Documentation 11*

*Chapter 2. Base-Year Instrumentation*

(and science) identity, mathematics (and science) utility, mathematics (and science) self- efficacy—and employ multiple items (to support reliable measurement).

Guided by the framework, the development and review process for each questionnaire consisted of the following steps:

1. *Literature review*. The research literature was consulted to help to flesh out the

framework by developing it at the level of specific constructs, and where possible, items or clusters of items that were intended to measure the construct. Past questionnaires from the National Center for Education Statistics (NCES) Secondary Longitudinal Studies study series were one source of items. However, many of HSLS:09’s themes were new to the study series. Indeed, considerable emphasis was placed on representing the recent relevant research literature at its broadest and deepest. Although some preference was given to items that had been used successfully on large national youth populations, and whose measurement properties were therefore well known, items used only on a small scale were also considered for the field test, as well as items written by the instrument development team to fill gaps in the available literature. In addition to field testing, new items were subject to cognitive interviews.

2. *Consultation*. The NCES project officer consulted with various federal government

offices and interest groups concerning data needs.

3. Circulating drafts of work in progress. Draft elements of the field test (and later, full-

scale) questionnaires—usually specific items, listed under and intended to measure a broader construct—were shared between the contractor teams for the separate questionnaires, and NCES and Education Statistical Services Institute (ESSI) staff, who took an active role in the development process. 4. *Technical review panel (TRP) review*. The HSLS:09 TRP, a specially appointed,

independent group of substantive, methodological, and technical experts, reviewed questionnaire content at each of its three meetings held under the base-year contract. The TRP met in November 2007 to review plans for the field test, including preliminary statements of questionnaire themes and constructs. The second meeting was held in January 2008, and drafts of the field test instruments were reviewed. The third meeting was held in January 2009, to review field test results and make recommendations for the main study questionnaires.8 5. *Writing of justifications for Office of Management and Budget (OMB) review*. For both the field test instruments, and later the main study instruments, a justification was written for the questionnaire items, noting issue areas, constructs to be measured within each, and the wording and response categories for the items that would be used to measure each construct. These draft questionnaires with justifications were submitted to the federal OMB for review and approval and subject to an ample public comment period. The questionnaires were revised based on OMB comments, and any questions from the public were addressed by the NCES project officer.

8 Minutes of the TRP meetings can be found in appendix D of the HSLS:09 base-year Field Test Report (Ingels et al. 2010).

*12 HSLS:09 Base-Year Data File Documentation*

*Chapter 2. Base-Year Instrumentation*

6. *Field testing and revision*. As noted above, the final step was revision of the instruments for the main study based on results from the field test, cognitive interviews, and OMB feedback. However, considerable hands-on testing of the programming logic for the questionnaires (and the computerized assessment) constituted the final step in developing a survey-ready instrument, after content approval from OMB. Specific items for the base-year mathematics assessment were reviewed by a mathematics advisory panel of mathematicians and mathematics educators (see section 2.3.1.1). Assessment items are not reviewed by OMB, nor were specific assessment items reviewed by the TRP. However, the larger assessment framework and goals and the assessment results (as seen in overall item statistics from the field test) were an integral element of the TRP deliberations.

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

**2.1.2 HSLS:09 Instrument Development Goals**

The primary research objectives of HSLS:09 are longitudinal in nature; therefore, the first priority for the study questionnaires was to select the items that would prove most useful in predicting outcomes as measured in future survey waves. The study of the transition through high school and out of high school to postsecondary education, the labor force, and, by degrees, adult status, is the major goal of all NCES high school longitudinal studies. To this goal HSLS:09 has added its special emphasis, on student choice behaviors, observed over time and studied in their school context, and on the science, technology, engineering, and mathematics (STEM) pipeline and its outcomes, both in educational and occupational terms. The innovation of starting the study at the very beginning of high school, fall of 9th grade, is another noteworthy element of the HSLS:09 design that differentiates it from preceding studies and that requires some differences of content.

However, instrument development goals of the study are reflected in technical innovations as well. Since the National Longitudinal Study of the High School Class of 1972 (NLS:72) in 1972, the entire suite of NCES secondary longitudinal studies has used paper-and- pencil methods for in-school data collection. It was a major goal of HSLS:09 that the student questionnaire and assessment—and the parent, teacher, administrator, and counselor surveys—

*HSLS:09 Base-Year Data File Documentation 13*

*Chapter 2. Base-Year Instrumentation*

should be computerized (a minor exception was a highly abbreviated paper version of the parent questionnaire, administered for nonresponse conversion). The advantages of an electronic questionnaire and assessment are readily stated.

The advantage for the mathematics assessment is that computerization of the first stage of the two-stage adaptive test allows for a more sophisticated routing that draws on IRT to base second-stage assignment of form on the total pattern of first-stage responses. Computerization also eliminates the possibility of test administrator error in scoring the first-stage test. The advantage for the student questionnaire is that an electronic instrument facilitates complex routing, and provides for online consistency checking. This is a profound break with the past, and will be even more important in the first follow-up, when, already, students are setting out on different paths that can be captured only with a more complex branching than would be feasible for a paper questionnaire. The instrument also prompts the respondent to correct errors and omissions, and supplies help text where needed. Electronic instruments accommodate variation in sequencing of questionnaire modules; this feature can be exploited to dilute and redistribute end-of-instrument nonresponse of the poorest readers. Finally, electronic instruments largely replace paper documents with their attendant security risks.

**2.2 Base-Year Questionnaires**

Content of the base-year questionnaires is summarized below. Hardcopy specifications of the electronic questionnaires appear as appendix A. Simplified hardcopy versions (lacking routing logic) can be viewed on the NCES HSLS:09 website (http://nces.ed.gov/surveys/hsls09/questionnaires.asp

**2.2.1 Student**

*14 HSLS:09 Base-Year Data File Documentation*

).

The student questionnaire was primarily self-administered using a computer during in- school sessions. If a student was unable to participate during the in-school sessions, a telephone interview was conducted using the same survey instrument with only the addition of interviewer instructions.

The student interview began and ended with questions that collected names, addresses, and phone numbers of people who would know how to locate the student for future rounds of the study. *Section A* collected this information for parents while *Section I* collected this information for a relative and a close friend*.*

The first substantive section of the student interview, *Section B*, asked for the 9th- grader’s demographic information including sex, race/ethnicity, and birth date. Students were also asked to indicate their native language; those who learned a foreign language first were asked how frequently they currently speak that language with their mother and friends.

The next section, *Section C*, collected information on the 9th-grader’s recent school experiences. Students were asked to indicate the school they attended in the previous school year

*Chapter 2. Base-Year Instrumentation*

(2008–09) and their grade level at that time. The 9th-graders also reported their involvement in various mathematics and science activities since the beginning of the previous school year. Finally, the students identified the mathematics and science courses they took in the 8th grade and the final grade they earned in each.

*Section D* gathered data on self-efficacy in mathematics and identification as a mathematics person. In addition, a series of questions was posed about the mathematics course the 9th-grader was taking in the fall of 2009 and the teacher of that course. The name of the teacher that the school linked to the student was preloaded into the questionnaire. The student could either confirm that the teacher listed was his or her teacher or type in the name of another mathematics teacher if the name provided was incorrect. *Section E* repeated all of the same questions as *Section D*, but pertained to science instead of mathematics.

*Section F* included questions on attitudes about school, mathematics, and science. Other questions focused on whom the student talks to about education, career plans, and personal problems; friends’ attitudes about school and related behaviors; and programs in which the student had participated such as Upward Bound or MESA (Mathematics, Engineering, Science Achievement). Students were also asked to compare and evaluate males’ and females’ ability in mathematics, science, and English and language arts. This question was repeated on the parent and teacher questionnaires.

*Section G* focused on high school, career, and college plans. Specifically, students were asked about their intentions to take further mathematics and science courses, if they had a career or college plan and who helped them create it, and their plans to take standardized college placement exams. In conclusion, they were asked how sure they were that they would graduate from high school.

The final substantive section, *Section H*, collected data on educational expectations, plans for the year after high school, college plans, estimates of the cost of college, and the student’s expected occupation at age 30.

Students were randomly assigned to one of two groups which determined the order in which these sections were administered. Half of the students completed the sections in alphabetical order from *Section A* to *Section I*. The other half were administered sections in the following order: A, B, C, E, D, H, G, F, I. *Sections F* and *H* were swapped to balance item nonresponse for students who were unable to complete the entire questionnaire in the full-length in-school session. Similarly, *Sections D* and *E* were reordered to ensure that when the in-school session was shortened roughly the same number of students would be administered the questions in each section.

**2.2.2 Parent**

Data collection staff asked that the parent or guardian most familiar with the 9th-grader’s school situation and experience complete the parent questionnaire. Guided by this definition of

*HSLS:09 Base-Year Data File Documentation 15*

*Chapter 2. Base-Year Instrumentation*

the preferred respondent, the parent identified either him- or herself as the survey respondent or another individual. In rare instances, a guardian such as a grandparent responded.

Parents had the option of self-administering the questionnaire via the web or completing a telephone interview. Some 60.5 percent of parent interviews were administered by interviewers on the telephone. When development of the English version was complete, the questions, response options, prompts for critical items, messages that warned of inconsistent or invalid responses, help text, and navigation buttons were translated into Spanish. Bilingual interviewers were trained to administer the Spanish version of the questionnaire over the telephone. They were able to toggle between the English and Spanish versions as needed.

There were seven sections of the parent interview. *Section A* collected information about the residents of the 9th-grader’s household including the presence of parents or guardians in the household, their relationship to the 9th-grader, and their marital status. The total number of adult residents and minor residents were also collected. The parent was also asked how much of the time the 9th-grader lived in their household and with whom and when he or she lived elsewhere. Finally, questions pertaining to siblings included the number of older siblings and whether any siblings had attended the 9th-grader’s high school within the last 5 years.

*Sections B* and *C* collected data on the parents or guardians in the household. Typically the respondent was a parent or a partner of a parent. In these cases, the first series of questions (P1 series) pertains to the respondent and the second series of questions (P2 series) pertains to the respondent’s resident spouse or partner, if applicable. In a small number of instances, the respondent was a guardian such as a grandparent or other adult relative. These respondents were asked if one or both of the 9th-grader’s parents (i.e., biological, adoptive, step- or foster parents) lived in the household. If neither parent lived in the household, the first series of questions referred to the respondent and the second series referred to his or her resident spouse or partner, if applicable. If one parent lived in the household, the first series of questions applied to the respondent and the second series applied to the resident parent (in this case, the respondent is P1 and the resident parent is P2). If both parents were living in the household, the first series of questions pertained to the first parent identified by the respondent and the second series pertained to the second parent. In this last very rare scenario, no data about the respondent’s education or occupation were collected, and the actual respondent would not be labeled as P1 or P2.

*Section B* collected data on race and ethnicity, immigration status, and language use. Race/ethnicity and immigration status were collected for both parents if there were two parents in the household. Parents were asked for the country in which the student was born, when he or she came to the United States if born elsewhere, and in what grade he or she was placed upon arrival. In addition, we learned whether the student had ever been or was currently enrolled in a program for English language learners.

*16 HSLS:09 Base-Year Data File Documentation*

*Chapter 2. Base-Year Instrumentation*

The next section, *Section C*, gathered information on the socioeconomic status of the 9th- grader’s parents. Each parent’s educational attainment, employment status, and current or most recent occupation was collected. Household income and home ownership were also ascertained.

*Section D* focused on the student’s educational history including skipping or repeating grades, changing schools, dropout episodes, and suspensions and expulsions. In addition, data were collected on disabilities, special education services, enrollment in honors courses, and the frequency of contact from the school about problematic behavior, attendance, or performance.

*Section E* measured parents’ involvement in the 9th-grader’s education and learning. Questions pertained to school selection, participation in school meetings and events, and helping with homework. In addition, parents were asked about activities the 9th-grader had engaged in outside of school and with a family member. Parents were also asked to compare and evaluate males’ and females’ ability in mathematics, science, and English and language arts.

Questions in *Section F* pertained to the 9th-grader’s plans and preparations for postsecondary education. Parents were asked how far in school they hoped their 9th-grader would go, how far they anticipated they would actually go, and if they had spoken with someone knowledgeable about the requirement for admission to a postsecondary institution. If postsecondary education was a goal, parents were asked further questions such as what type of postsecondary institution the 9th-grader was most likely to attend first, when this education would begin, how much they estimated a postsecondary education would cost, whether they planned to help pay for this education, and how they have prepared financially.

The final section of the interview, *Section G*, collected contact information for parents, relatives, and friends who can locate the 9th-grader in subsequent rounds of the study.

There were two abbreviated versions of the parent questionnaire. The primary abbreviated instrument, a modified version of the web/computer-assisted telephone interview (CATI) instrument, included a subset of the critical items from each section of the full-length interview. A secondary two-page paper instrument was used for the most reluctant or difficult-to- reach parents. This brief questionnaire asked how the respondent and another parent/guardian were related to the 9th-grader, data which help construct the family structure composite variable. It also collected data on parents’ education level, occupation, and income for constructing the socioeconomic status measure. See chapter 4 for further detail on parent data collection.

**2.2.3 Teacher**

All teachers who had an HSLS:09 student in his or her mathematics or science course were eligible for the teacher questionnaire. The school identified the teachers and courses in which an HSLS:09 student was enrolled. At the beginning of the questionnaire, teachers were presented with this list and asked to confirm each of the courses as one that they taught. The teacher would later be asked to report on each of the confirmed courses. If the teacher indicated that none of the listed courses were ones which he or she taught, he or she was routed to a screen

*HSLS:09 Base-Year Data File Documentation 17*

*Chapter 2. Base-Year Instrumentation*

which collected up to five course titles (a limit of five courses was set to avoid overburdening respondents).

After this introduction, the teacher questionnaire had four sections. *Section A* collected background information on the respondent, including demographic characteristics, educational history, certification, and teaching history. This section was completed by both mathematics and science teachers. The abbreviated teacher interview concluded at the end of *Section A*.

*Section B* was administered to mathematics teachers only. It asked respondents to evaluate mathematics teachers and the mathematics department in their school. It also asked these teachers how mathematics teaching assignments are made. A series of questions was asked about each course the teacher confirmed at the beginning of the interview. First, the teacher was asked to classify the course using a prescribed set of course titles (e.g., algebra I, geometry). Then the teacher assessed the achievement level and preparedness of students in the course and reported on the use of small groups in class and his or her emphasis on various course objectives. *Section C* included all of the aforementioned questions in *Section B*, but were asked of science teachers and pertained to science education in their school.

The final section of the teacher questionnaire, *Section D*, was administered to both mathematics and science teachers. It covered a range of topics including evaluations of the school’s principal and the school’s faculty. Other questions pertained to the prevalence of various problems at the school and limitations on their teaching. Respondents’ beliefs about the influence of a student’s home environment on their ability to be effective teachers were measured as were their beliefs about how males’ and females’ mathematics and science abilities compare. It should be remembered that the teacher data supply contextual information for students, who in turn constitute the unit of analysis. The teacher sample is not representative of teachers in the school. The design of this component does not provide a standalone analysis sample of teachers, but instead permits specific teacher characteristics and practices to be related directly to the learning context and educational outcomes of sampled students.

**2.2.4 School Administrator**

The school administrator questionnaire consisted of five sections. The first four asked factual questions about the school; it could be completed by the principal or another knowledgeable individual on the school’s staff. The school administrator was the only appropriate respondent for the final section, however, because it asked background and subjective questions. Different login credentials were issued to school administrators and their designees such that school administrators were able to access the entire questionnaire, while designees were able to access only the first four parts. In an effort to reduce the burden of reporting detailed statistics, respondents were instructed that informed estimates were acceptable.

*18 HSLS:09 Base-Year Data File Documentation*

*Chapter 2. Base-Year Instrumentation*

*Section A* collected data on a range of topics. Information on the school’s characteristics includes grade span, control (public or private), type (e.g., charter, magnet, single sex, religious), academic calendar, and course scheduling. This section also gathered information on average daily attendance, policy on informing parents of student absences, and transferring students to alternative schools. Another series of questions focused on schools identified as in need of improvement based on Adequate Yearly Progress requirements of No Child Left Behind. *Section A* concluded with questions about efforts the school had made to increase students’ interest in mathematics and science and to help struggling students.

*Section B* gathered information on the student body, including their racial makeup, the percentage of 9th-grade students who were repeating 9th grade, the percentage of the 2008–09 9th-grade class that returned to the school for the 2009–10 academic year, and the pursuits of the 2009 senior class. This section also determined the student enrollment expressed as a percentage of capacity (e.g., 110 percent filled) and the percentage of the student body enrolled in various programs such as a dropout prevention program or Advanced Placement courses.

*Section C* collected information on the school’s faculty, with particular emphasis on mathematics and science teachers. Respondents reported the number of full- and part-time teachers in mathematics, science, and all other subjects. The number of mathematics and science teachers certified by the state to teach in their respective subject areas was also collected. In addition, there was a series of questions about vacancies in the mathematics and science departments and efforts to fill them. The percentage of teachers absent on an average day was also collected.

*Section D* collected data on the mathematics and science curriculum. Requested information included the mathematics and science courses offered on- and off-site, whether completion of particular mathematics or sciences courses is required to graduate, whether these required courses are the same as or more advanced than state requirements, and whether students are placed in different algebra I courses based on ability.

The final section, *Section E*, included questions about the school administrator’s background and his or her evaluation of the school’s problems. Requested information included the administrator’s demographic characteristics, educational and occupational history including years of experience as a school administrator and teacher, and certification. The school administrator was also asked to report the number of work hours spent each week on various tasks and activities. Finally, the school administrator was asked to evaluate the school’s challenges.

The abbreviated version of the web/CATI instrument included all of the critical items in the full-length instrument. These questions were only drawn from the first three sections of the interview so a designee could complete the abbreviated instrument.

*HSLS:09 Base-Year Data File Documentation 19*

*Chapter 2. Base-Year Instrumentation*

**2.2.5 Counselor**

The counselor questionnaire had four sections. *Section A* collected the total number of full- and part-time counselors on staff, the number certified as high school counselors, and the average caseload per counselor. Other questions in the first section ascertained the way in which counselors are assigned to students, the goals emphasized by the counseling program, and how the respondent allocates his or her work hours to delivering various services. Additional areas of inquiry were how counselors and the school as a whole assisted 8th-grade students’ transition into 9th grade and the school’s use of career and education plans.

*Section B* focused on programs and services offered to students. Some of these questions pertained to enrichment courses, assistance for struggling students, dropout prevention programs, encouragement of the pursuit of mathematics and science education and employment, and assistance with the transition from high school to college or the workforce. Other topics included the use of mathematics competency tests and options for failing students.

*Section C* collected data on criteria used to place 9th-graders and upperclassmen in mathematics and science courses. *Section D,* the final section, collected background information on the school counselor including how he or she entered the counseling profession, how many years he or she had served as a counselor, and his or her educational history. Respondents were also asked for their evaluation of the school’s principal, teachers, and counselors.

Data users are reminded that the head counselor at each school was asked to complete the questionnaire, reporting on the counseling services, program placement, and transitioning practices of their schools. Consequently, the respondents do not constitute a standalone nationally representative sample of high school counselors (or 9th-grade counselors). For this reason, the counselor-specific information in Section D should be viewed as methodological information about the HSLS:09 counselor sample, and not as the source of estimates of the characteristics of the population; that is, the data elicited by these questions cannot be generalized to the nation’s high school counselors. Nor are the counselors necessarily the current counselors of the 9th-graders in the HSLS:09 base-year sample (some of the counselors deal with entirely different grades).

**2.2.6 Rules for Defining Completed Interviews**

A completed case was defined as a respondent having reached a certain place in the questionnaire. However, it should be noted that because of the nature of the web survey, respondents had the ability to answer or skip any item. The completeness of data therefore varies across respondents. For this reason, in addition to requiring that a certain place in the questionnaire should be reached, it was also stipulated that a certain critical mass of questions (for all questionnaires, generally at least 15 items9

) should be answered. The point reached

9 In rare instances where information of key analytical value was provided, this criterion was relaxed.

*20 HSLS:09 Base-Year Data File Documentation*

*Chapter 2. Base-Year Instrumentation*

necessary for inclusion on the data file reflected a dual requirement—evidence of respondent seriousness in responding to the survey, and data of substantive value.

**Student.** The student interview comprised nine sections, two of which (A and I) do not appear on the data files:

A—Future locating information concerning parents B—Basic demographic information (age, sex, race/ethnicity, etc.)

C—Recent school experiences D—Mathematics self-efficacy and other social cognitive scales, teacher link

E—Science self-efficacy and other social cognitive scales, teacher link F—Attitudes about school, mathematics, and science G—High school, college, and career plans

H—Plans for the year after high school, perceptions of college costs I—Future locating information concerning relative or close friend

A student survey was counted as complete if the end of *Section C* was reached and a critical mass of questions (normally 15) was answered. However, rare cases that showed irrational patterns of response (akin to pattern marking on the assessment) were not counted as complete. **Parent.** The parent interview consisted of seven sections:

A—Family B—Family’s origin and language use

C—Family education and occupation D—Previous educational experiences

E—Parent’s involvement F—9th-grader’s future

G—Locating The abbreviated instrument included selected items from all sections.

Parents who did not complete the full interview were nevertheless deemed respondents if they reached all questions through immigration status in *Section B* (P1USYR2). However, cases were also counted as respondents (that is, as complete) if any valid interpretable socioeconomic status data (e.g., family income, parental educational attainment, or parental occupation) were provided (this occurred particularly in the parent pencil-and-paper interview).

*HSLS:09 Base-Year Data File Documentation 21*

*Chapter 2. Base-Year Instrumentation*

**Teacher.** The teacher survey consisted of four sections:

A—Background B—Mathematics department and instruction (mathematics teachers only)

C—Science department and instruction (science teachers only) D—Beliefs about teaching and school

The abbreviated interview comprised *Section A* only. Teachers who did not finish the questionnaire but provided educational histories with the exception of coursework in *Section A* and had data for at least 15 items were deemed respondents.

**Administrator.** The school administrator survey consisted of five sections:

A—School characteristics B—Student population

C—School’s teachers D—Courses offered

E—Goals and background The abbreviated instrument included selected items from *Sections A, B,* and *C*.

Respondents to the full-length or abbreviated interview who reached questions through A1REPEATG9, the last question in *Section B* of the abbreviated instrument and who answered at least 15 questions were deemed respondents.

**Counselor.** There was no abbreviated instrument for counselors. The counselor survey consisted of four sections:

A—Staffing and practices

B—Programs C—Mathematics and science placement

D—Opinions and background Counselors who finished *Section A* (that is, reached the last item) with responses for at least 15 items but did not complete the entire interview were deemed respondents.

**2.3 HSLS:09 Mathematics Assessment of Algebraic Reasoning**

This section describes the development and format of the HSLS:09 mathematics assessment of algebraic reasoning, the scoring procedures, and the types of scores used, along with summary statistics. The purpose of the HSLS:09 assessment battery is to provide a measure at two time points of student achievement in algebra for a cohort of grade 9 students—during the first part of their 9th-grade year (fall term of the 2009–10 school year) and again in spring 2012 when most of the cohort will be in the second semester of their 11th-grade year.

*22 HSLS:09 Base-Year Data File Documentation*

*Chapter 2. Base-Year Instrumentation*

**2.3.1.1 Mathematics Advisory Panel**

The initial draft of the algebraic reasoning framework and each of the proposed field-test items were developed by staff at the American Institutes of Research with support of and review by John Dossey, emeritus professor of mathematics at Illinois State University, who served as a project consultant. A Mathematics Advisory Panel reviewed, refined, and validated the framework and reviewed and approved each proposed item. The panel comprised the following individuals:

• • • • •

Hyman Bass, Professor of Mathematics, University of Michigan

Katherine Halvorsen, Professor of Mathematics and Statistics, Smith College

Joan Leitzel, President Emeritus, University of New Hampshire, and Professor of Mathematics (retired), Ohio State University

Mark Saul, Mathematics Teacher (retired), Bronxville High School, New York

Ann Shannon, Mathematics Education Consultant, Oakland, California

*HSLS:09 Base-Year Data File Documentation 23*

**2.3.1.2 Algebraic Reasoning Framework**

The item development process began with the development of a set of test and item specifications that described the importance of algebra and defined the domain of algebraic reasoning for the Mathematics Assessment of HSLS:09. This task entailed designing an assessment of student understanding, and growth in understanding, of key algebraic knowledge and skills in algebra as a measure of mathematical preparation for the study of science, preparation for further study within the mathematical sciences and statistics, and preparation for the requisite skills and expectations of the workplace. Accordingly, the framework was designed to assess a cross-section of understandings representative of the major domains of algebra and the key processes of algebra.

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

• •

Algebraic Content Domains:

− − − − − −

The language of algebra

The language of algebra

Proportional relationships and change

Proportional relationships and change

Linear equations, inequalities, and functions

Linear equations, inequalities, and functions

Nonlinear equations, inequalities, and functions

Nonlinear equations, inequalities, and functions

Systems of equations

Systems of equations

Sequences and recursive relationships

Sequences and recursive relationships

Algebraic Processes:

*Chapter 2. Base-Year Instrumentation*

− − − −

*24 HSLS:09 Base-Year Data File Documentation*

Demonstrating algebraic skills

Using representations of algebraic ideas

Performing algebraic reasoning

Solving algebraic problems

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

**2.3.1.3 Two-Stage Computer-Delivered Implementation**

The HSLS:09 mathematics assessment was administered by computer, using a two-stage design wherein each student completed a Stage 1 “router test” and then a Stage 2 test designated as “low,” “moderate,” or “high” that was assigned on the basis of Stage 1 performance. Table 1 shows this design:

• • • •

Each student took a common 15-item Stage 1 router test that consisted of 4 grade 9 items and 11 grades 9 and 11 items (current plans are to use some or all of these 11 items on the first follow-up grade 11 router).

On the basis of Stage 1 performance, each student was routed to a low, moderate, or high Stage 2 test, each consisting of 25 items drawn from the grade 9 and the grades 9 and 11 pools.

Students were only aware that they were taking a 40-item test.

For linking purposes, 12 items were common to both the high and moderate Stage 2 tests and 5 items were common to both the low and moderate Stage 2 tests (in addition, the 12 items common to both the high and moderate Stage 2 tests are expected to be used on the grade 11 test).

**Table 1. HSLS:09 Mathematics Assessment grade-9 main study design: 2009**

Items at stage 1

Stage 2 level

Items at stage 2 Items per student

Unique

Across grades Unique

Across grades

Across stages Total

Stage 1

Stage 2 4 11

High 13 12 0 40 15 25 Moderate 8 12 5 40 15 25 Low 20 0 5 40 15 25

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

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

*Chapter 2. Base-Year Instrumentation*

The 73 unique items comprising the Stage-1 router and Stage-2 test forms were selected from the field-test pool of 264 unique items, some designed for only grade 9 students, some for grades 9 and 11 students, and some for only grade 11 students. The selection of items was based on the following criteria:

• •

*HSLS:09 Base-Year Data File Documentation 25*

The entire pool of 73 items needed to represent a balance across the six content domains and the four algebraic processes.

The average difficulty of the 15 items allocated to the Stage 1 router test and to each set of 25 items on the Stage 2 tests was preset as follows on the basis of the difficulty parameter of the IRT model (i.e., b-parameter) obtained using the field-test data:

− − − −

Stage 1, router average difficulty = 0.6

Stage 2, low test average difficulty < –0.4

Stage 2, moderate test average difficulty = 0.6

Stage 2, high test average difficulty > 1.6

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

One item on the Stage 2 high test was eliminated from the analysis on the basis of very weak item statistics,10

leaving a pool of 72 items for scoring and analysis.

**2.3.1.4 Allocation of Second-Stage Forms**

A total of 20,781 students had complete assessment data. Table 2 shows the breakdown by form, and supplies number of students, and weighted and unweighted percent of students.11

10 Some 73 items were employed in the main study assessment, but one item was subsequently dropped, leaving a pool of 72 unique items. The dropped item (Q240) had poor item-total correlation (adjusted biserial correlation = 0.07) and the examination of the IRT item fit graphs showed that it did not fit the IRT model used in this study. 11 Throughout this document, weighted and unweighted estimates are reported. The unweighted estimates pertain to the sample and the weighted estimates to the specified target population. Weighted estimates reflect the fact that students (and schools) have different selection probabilities, hence their weights vary. For example, groups that were over-sampled relative to their proportion in the population (e.g., Asians, private school students) will generally have smaller weights (i.e., generalize to fewer members of the population).

*Chapter 2. Base-Year Instrumentation*

**Table 2. Number and percentage of HSLS:09 Mathematics Assessment test-takers by form:**

**2009**

Category Number

*26 HSLS:09 Base-Year Data File Documentation*

Percent Unweighted Weighted Total 20,781 100.0 100.0

Second-stage form1

Low 4,356 21.0 24.3 Moderate 10,070 48.6 48.8 High 6,283 30.3 26.9 1 Seventy-two students did not reach Stage 2 and were not included in the total for percentages. SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year.

**2.3.2 Scoring Procedures**

The assessment data were examined for possible indicators of lack of motivation to answer questions to the best of the student’s ability. Examples of possible indicators are missing responses and pattern marking (e.g., all answers were “A” or “ABCDABCDABCD...”). As a result, of the 20,956 students who took the assessment, 175 (< 1 percent) test records were discarded from the analysis sample for the following reasons:

• •

A total of 108 records were deleted for attempting (i.e., selecting one of the four response options) fewer than six items.

A total of 67 records was deleted for pattern marking (64 cases for selecting the same answer options to more than 10 consecutive items, 2 cases for having the repeating “ABCDABCDABCD...” pattern throughout Stage 2 test and most of the router test, and 1 case owing to other response pattern marking).

Classical item analyses were then conducted to provide information on item performance. The classical item statistics including p+ value, adjusted item-test biserial correlations, omit rate, distractor statistics, and Differential Item Functioning (DIF) statistics were computed and reviewed. One item was flagged for potential DIF but no bias was found after further expert review of the item wording and content. The p+ value for each of the items is presented in appendix B.

The scores used to describe students’ performance on the mathematics assessment are based on IRT12

(Hambleton and Swaminathan 1985). The IRT model uses patterns of correct, incorrect, and omitted responses to obtain ability estimates that are comparable across the low-, moderate-, and high-difficulty test forms. One of the assumptions under an IRT model is unidimensionality of the test items. To verify that the items met that assumption, confirmatory

12 Readers are reminded that technical terms are defined in a Glossary (appendix C).

*Chapter 2. Base-Year Instrumentation*

factor analysis (CFA) was conducted based on each test form.13

Specifically, the IRT three-parameter logistic (3PL) model was used to calibrate the test items and estimate a student’s ability. The 3PL model is a mathematical model for estimating the probability that a person will respond correctly to an item. This probability is given as a function of one parameter characterizing the proficiency of a given student and three parameters characterizing the properties of a given item—the item’s difficulty, discriminating ability, and a guessing factor. The IRT model accounts for the three characteristics of each test question in estimating a student’s ability. The item parameters for each of the items are presented in appendix B. BILOG-MG (Zimowski et al. 2003) was used in carrying out item calibration and student ability estimation. During item calibration, separate ability priors based on performance on the router test were used for each of the three sub-populations taking the different second- stage tests (i.e., low-, moderate- and high-forms). The Bayesian estimation procedure was applied in estimating student proficiency.

*HSLS:09 Base-Year Data File Documentation 27*

The model fit indices obtained from the CFA analyses suggested that the items were unidimensional within each form.

IRT scoring has several advantages over traditional raw number-right scoring. First, IRT uses the overall response pattern of right and wrong answers to estimate ability and therefore can account for the guessing factor—a low-ability student guessing several difficult items correctly. Specifically, if answers on several easy items are wrong, a correct difficult item is assumed, in effect, to have been guessed. Second, unlike in raw number-right scoring, where omitted (skipped) responses are treated as incorrect answers, IRT procedures treat the omitted responses as not administered and use the pattern of responses to estimate the probability of correct responses for all test questions. Therefore, omitted items are less likely to cause distortion of scores as long as enough items have been answered right and wrong to establish a consistent pattern. Finally, IRT scoring makes it possible to compare scores obtained from test forms of different difficulty, such as HSLS:09. The common items present in the routing test and in overlapping Stage 2 forms allow test scores to be placed on the same scale. Looking ahead to the plans for the HSLS:09 first follow-up survey, IRT procedures will be used to estimate longitudinal gains in achievement over time by using common items present in both the grade 9 and grade 11 forms.

**2.3.3 Score Descriptions and Summary Statistics**

Several different types of scores are used in HSLS:09 to describe students’ performance on algebra, all derived from the IRT model. Specifically, the IRT model uses information obtained from all students’ response patterns of right and wrong answers as well as characteristics of the assessment items to compute a student ability estimate, theta. This theta

13 It would be ideal to conduct the CFA based on the pool of all 72 items. However, because of the test design of this study, many item pairs had no common observations and therefore their covariance could not be computed. The resultant large number of missing covariances would lead to unreliable results if the CFA were based on the pool of all 72 items. Therefore, the CFA was conducted based on the data for each of the following three 40-item tests: Router + Low second-stage form; Router + Moderate second-stage form; and Router + High second-stage form.

*Chapter 2. Base-Year Instrumentation*

(ability) estimate is the basis for all other types of scores derived thereafter. On the data file, users will find the following scores:

• • • • •

*28 HSLS:09 Base-Year Data File Documentation*

Theta and the standard error of measurement of theta

Estimated number-right scores

Standardized scores (T-scores)

Quintile scores

Proficiency probability scores Details of the scores are described below. The choice of the most appropriate score for analysis purposes should be driven by the context in which it is to be used.

**2.3.3.1 Theta (Ability) Estimate and Standard Error of Measurement of Theta**

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

The theta ability scores provide a summary measure of achievement useful for correlational analysis with status variables, such as demographics, school type, or behavioral measures, and may be used in multivariate models as well. When longitudinal data become available with the HSLS:09 first follow-up, theta scores can also be used to measure achievement growth over time.

**2.3.3.2 Estimated Number-Right Scores**

The estimated number-right score is an overall, criterion-referenced measure of achievement at a point in time. The criterion is the set of skills defined by the HSLS:09 framework and represented by the 72 items in the HSLS:09 mathematics item pool. The estimated number-right score is an estimate of the number of items that students would have answered correctly had they responded to all 72 items in the item pool. The ability estimates and item parameters derived from the IRT calibration can be used to calculate each student’s probability of a correct answer for each item in the pool. These probabilities are summed over the total number of items in the item pool (72) to produce the IRT-estimated number-right score; therefore, the score has a potential range of 0 to 72. Table 3 presents the variable name, description, and summary statistics for the IRT-estimated number-right scores.

*Chapter 2. Base-Year Instrumentation*

**Table 3. Various types of scores from HSLS:09 Mathematics Assessment, by variable: 2009**

Variable Description

*HSLS:09 Base-Year Data File Documentation 29*

Weighted mean

Weighted standard

deviation Range X1TXMRTH HSLS:09 base-year mathematics theta score −0.07 0.97 −2.6–3.0 XTXMSCR HSLS:09 base-year mathematics IRT-

estimated number-right score 38.85 11.91 15.9–69.9 X1TXMTSCOR HSLS:09 base-year mathematics

standardized score (T-score) 50 10 24.3–81.8 X1TXMQUINT HSLS:09 base-year mathematics quintile — — 1–5

— Not available. SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year.

The IRT-estimated number-right scores are useful in identifying cross-sectional differences among subgroups in overall achievement level (see *HSLS:09 Base Year First Look Report* (Ingels et al. 2011) for an illustration of the cross-sectional use of a variety of mathematics scores). Similar to the theta ability scores above, they also provide a summary measure of achievement useful for correlational analysis with status variables, such as demographics, school type, or behavioral measures, and may be used in multivariate models as well.

When data are available from the HSLS:09 follow-up study, which is designed to be vertically linked to the base-year study, these scores, like the theta scores, may also be used as longitudinal measures of overall growth when an aggregated measure is preferred.14

**2.3.3.3 Standardized Scores (T-scores)**

When a disaggregated measure is desired, to measure and compare gains made at different points on the score scale (that is, to target a hierarchy of specific sets of skills), the probability of proficiency scores as discussed below may be preferred in longitudinal analysis.

The standardized scores (T-scores) provide a norm-referenced measurement of achievement, that is, an estimate of achievement relative to the HSLS:09 student population (i.e., fall 2009 grade 9 students) as a whole. They provide overall measures of status at a point in time compared with those of peers, as distinguished from the criterion-referenced scores, which represent status with respect to achievement on a particular criterion set of test items. The norm- referenced standardized scores do not answer the question, What skills do students have? but rather, How do they compare to their peers?

The standardized T-score is a transformation of the IRT theta (ability) estimate, rescaled to a familiar metric with a mean of 50 and a standard deviation of 10. The transformation

14 For examples of the use of an IRT-based score (estimated number-right) within similarly designed NCES longitudinal studies (ECLS-K and ELS:2002), see Guarino et al. (2006) and Bozick and Ingels (2008). The two NCES reports also illustrate both principal approaches to measuring achievement gain within a regression framework: use of gain scores as the dependent variable (Guarino et al.) versus use of follow-up scores as a covariate (Bozick and Ingels).

*Chapter 2. Base-Year Instrumentation*

facilitates comparisons in standard deviation units. For example, an individual with a T-score of 75 (or a subgroup with a mean of 75) has performed 2.5 standard deviations above the national average for 9th-graders, whereas a score of 40 corresponds to 1 standard deviation below the norm. These numbers do not indicate whether students have mastered a particular algebraic skill or concept, but rather what their standing is relative to that of others. The HSLS:09 T-scores are documented in table 3, which also presents the summary statistics of the other types of scores discussed in the sections below.

**2.3.3.4 Mathematics Quintile**

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

Mathematics quintiles are convenient normative scores for the user who wants to focus on an analysis of background or process variables separately for students at different achievement levels. For example, one might want to compare the school experiences or educational aspirations of students in the lowest quintile with those of students in the highest quintile group. Table 3 contains the variable name, description, mean, and ranges for the quintile scores.

**2.3.3.5 Probability of Proficiency Scores**

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

• • • •

**Level 1: Algebraic expressions.** Students able to answer questions such as these have an understanding of algebraic basics, including evaluating simple algebraic expressions and translating between verbal and symbolic representations of expressions.

**Level 2: Multiplicative and proportional thinking.** Students able to answer questions such as these have an understanding of proportions and multiplicative situations and can solve proportional situation word problems, find the percent of a number, and identify equivalent algebraic expressions for multiplicative situations.

**Level 3: Algebraic equivalents.** Students able to answer questions such as these have an understanding of algebraic equivalents and can link equivalent tabular and symbolic representations of linear equations, identify equivalent lines, and find the sum of variable expressions.

**Level 4: Systems of equations.** Students able to answer questions such as these have an understanding of systems of linear equations and can solve such systems

*30 HSLS:09 Base-Year Data File Documentation*

*Chapter 2. Base-Year Instrumentation*

algebraically and graphically and characterize the lines (parallel, intersecting, collinear) represented by a system of linear equations.

• **Level 5: Linear functions.** Students able to answer questions such as these have an understanding of linear functions, can find and use slopes and intercepts of lines, and can use functional notation. The levels are hierarchical in the sense that mastery of a higher level typically implies proficiency at the lower levels. The HSLS:09 proficiency probabilities are IRT-derived estimates and are computed using IRT-estimated item parameters. The probability of proficiency for a given student at a given level is calculated as the probability of getting correct at least three of the four items in a given cluster marking a proficiency level (the probability of a student getting at least three items correct out of four is expressed as the sum of (1) the probability of getting all four items correct and (2) the probability of getting any three items correct). Although clusters of four items anchor each mastery level, the probability of proficiency is a continuous score that does not depend on a student answering the actual items in each of the clusters but rather on the probability of a correct answer on these items given the overall pattern of response on the items completed.

Under the HSLS:09 two-stage adaptive assessment design, with different forms keyed to different ability levels, not all students received all items. Nevertheless, the IRT model permits proficiency probabilities to be estimated, even for those students who were not administered a particular proficiency/mastery cluster. The probability of proficiency scores are summarized in table 4.

**Table 4. HSLS:09 algebra probability of proficiency scores, by variable: 2009**

Weighted standard Variable Description Mathematical definition Range

deviation X1TXMPROF1 HSLS:09 proficiency

probability score: Level 1

*HSLS:09 Base-Year Data File Documentation 31*

Weighted mean Algebraic expressions

0–1 0.85 0.27 X1TXMPROF2 HSLS:09 proficiency

probability score: Level 2

Multiplicative and

proportional thinking 0–1 0.59 0.36 X1TXMPROF3 HSLS:09 proficiency

probability score: Level 3

Algebraic equivalents

0–1 0.41 0.36 X1TXMPROF4 HSLS:09 proficiency

probability score: Level 4

Systems of equations

0–1 0.19 0.23 X1TXMPROF5 HSLS:09 proficiency

probability score: Level 5

Linear functions

0–1 0.09 0.11

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

*Chapter 2. Base-Year Instrumentation*

Probability of proficiency scores may be used in a number of ways.15

**2.3.4 Psychometric Properties of the Test**

*32 HSLS:09 Base-Year Data File Documentation*

They may be used to locate the achievement of HSLS:09 sample members and subgroups at various behaviorally defined skill levels. The mean of a proficiency probability score aggregated over a subgroup of students is analogous to an estimate of the percentage of students in the subgroup who have displayed mastery of the particular skill. Because the range of the scores is 0 to 1, means can be expressed in percentage form. For example, the weighted mean for mastery of mathematics level 1 in HSLS:09 is 0.85, which is equivalent to saying that 85 percent of the grade 9 students had achieved mastery at this level (algebraic expressions). When HSLS:09 first follow-up (2012) data become available, the proficiency scores can be used to measure gain. The proficiency probabilities are particularly appropriate for relating specific processes to changes that occur at different points along the score scale. For example, two groups may have similar gains, but for one group, gain may take place at an upper skill level, and for the other, at a lower skill level. For those who gain at the higher skill level, there may be an association between their gains and curriculum exposure, such as taking advanced mathematics classes.

All items in the HSLS:09 mathematics assessment item pool were field tested. The field test was designed to provide information on item and test characteristics to ascertain the effectiveness of each item, develop a pool of main study items, and inform the placement of items on the main study test forms. Information about the psychometric properties of the items that were field tested, the setting of difficulty levels, differential item functioning, and the IRT scaling procedures are provided in the base-year field test report (Ingels et al. 2010).

The classical definition of reliability is the ratio of the true score variance to the observed score variance, which is the sum of the true scores variance and the error variance. In an IRT context, the true scores are the unobservable theta values that are estimated with a specified standard error from item response patterns. In HSLS:09, where Bayesian estimation procedures were applied, the estimate of the error variance was computed as the mean of the variances of the posterior distributions of ability for each test-taker in the sample. The true score variance is estimated by the variance of the Bayesian theta scores (ability estimates) in the whole sample (see Bock and Mislevy 1982 for more information on Bayesian estimation). The reliability is therefore the true score variance divided by the sum of the true score variance and the error variance (i.e., total variance). The IRT-estimated reliability of the HSLS:09 test was 0.92 after sample weights were applied. This reliability is a function of the variance of repeated estimates of the IRT ability parameter (within variance), compared with the variability of the sample as a whole. This 0.92 reliability applies to all scale scores derived from the IRT estimation including the probability of proficiency scores. Imputed test scores were not included in the reliability estimation.

15 See Bozick and Ingels (2008) for an illustration of the use of probability proficiencies in a similar longitudinal study, ELS:2002. For further discussion of the nonequivalence of scale score points and consequent need (if achievement gain is to be fully interpreted) for multiple criterion-referenced proficiency levels that mark distinct learning milestones, see Rock (2007).

*Chapter 2. Base-Year Instrumentation*

**2.4 Linkage With Prior NCES Studies**

**2.4.1 Questionnaire Linkage With Prior NCES Studies**

HSLS:09 data do not directly support certain kinds of cross-cohort comparison that were possible in earlier NCES Secondary Longitudinal Studies. Specifically, the study was not designed to facilitate intercohort time-lag comparisons. In this kind of comparison, same-grade persons of different cohorts are used to provide a time series for comparison—say, high school seniors in 1972, 1980, 1992, and 2004. Comparison is possible because each group is similarly defined (12th-graders) and because, by design, a core of questions has been repeated over time so that it is common to all the cohorts. Although in HSLS:09 there are some questionnaire items that are shared with some of the earlier studies, consistency of measures was not emphasized. (Nor is cross-cohort comparability a characteristic of the assessment. See section 2.4.2 below.) Moreover and more importantly, the in-school grade cohorts of HSLS:09—fall-term 9th-graders and spring-term 11th-graders—correspond to none of the prior cohorts, which represented spring-term 8th-, 10th-, or 12th-graders.

Nonetheless, three kinds of comparisons can be made between HSLS:09 and the prior secondary school cohorts: (1) the planned postsecondary measurement points are the same (2 years out of high school, and 8 years out of high school) across HSLS:09, ELS:2002, and NELS:88; (2) coursetaking can be compared between HSLS:09 and HS&B, NELS:88, and ELS:2002, based on the continuous data for grades 9 through 12 that are supplied by high school transcripts; and (3) because HSLS:09 models the same transition—from adolescence in the high school years to young adulthood, as marked by educational attainment, work and career, and family formation—the design answers the same basic questions as the predecessor studies. It supports longitudinal comparative analysis across the cohorts. All of the studies are based on essentially similar sample designs, and provide nationally representative data across public and private schools and support similar or the same race/ethnicity domains. Despite differences in emphasis, all of the studies draw content from the same or similar theoretical constructs (e.g., achievement growth, school effectiveness, social capital, social attainment, human capital). In essence, all of the studies including HSLS:09 address, in a manner inviting historical comparison, questions such as the following:

• • • • •

What steps do high school students take to attend a 4-year (or 2-year) college?

What are the medium-term outcomes of not completing high school in the traditional way (or at all)?

How, when, and why do students enroll in postsecondary education?

What kind of transition do the non–college-bound make into the labor market?

Did those high school students who expected to complete a postsecondary qualification actually do so?

*HSLS:09 Base-Year Data File Documentation 33*

*Chapter 2. Base-Year Instrumentation*

• What is the relationship between high school curriculum and experience and sub- baccalaureate and baccalaureate attainment?

• How has the percentage of recent graduates from a given cohort who enter the workforce in various fields changed over the past years?

**2.4.2 Assessment Linkage With Prior NCES Studies**

Differences in the content and scaling of the HSLS:09 and prior tests administered in the study series severely limit the possibility of comparisons. The HSLS:09 assessment measures a critical strand of mathematics—algebraic reasoning. Apart from a handful of National Assessment of Educational Progress (NAEP) items, there are no common items that link the HSLS:09 base year test to earlier mathematics assessments. In addition, the prior frameworks were different, and, in particular, broader, so it would not seem that the various tests measure precisely the same thing. Finally, the testing points—fall of 9th grade and spring of 11th grade— are not shared with the prior longitudinal studies, Program for International Student Assessment (PISA), or NAEP. Therefore, even a weak linkage, such as a concordance, would seem inadvisable to implement.

*34 HSLS:09 Base-Year Data File Documentation*

**Chapter 3. Sample Design**

**3.1 Base-Year Sample Design Overview**

Details of the complex design and resulting sample for the High School Longitudinal Study of 2009 (HSLS:09) base-year study are provided in this chapter. Section 3.2 pertains to the stratified random selection of schools; section 3.3 documents the selection of students within schools; and section 3.4 describes the selection of contextual samples.

**3.2 Selection of School Sample**

Survey responses and mathematics assessment scores for HSLS:09 were collected through a stratified, two-stage random sample design with primary sampling units defined as schools selected at the first stage and students randomly selected from the sampled schools within the second stage. A total of 944 schools out of 1,889 eligible schools participated in the base-year study resulting in a 55.5 percent weighted response rate (50.0 percent unweighted).16 The details are described in the following sections.

**3.2.1 Target Population**

The HSLS:09 base-year main study included one target population for each of the two sample design stages—schools and students within schools. The target population for schools, units selected in the first stage of sampling, was defined as regular public schools, including public charter schools, and private schools in the 50 United States and the District of Columbia providing instruction to students in both the 9th and 11th grades.

Schools excluded from this definition (*study-ineligible schools*) include those that met any of the following criteria:17

• • • • • •

Bureau of Indian Affairs (BIA) schools;

Special education schools for students with disabilities;

Career technical education (CTE) schools that do not enroll students directly;

Department of Defense (DoD) schools located outside the United States (OCONUS);

Schools without both a 9th and 11th grade;

Schools not in operation during the fall of 2009;

16 The American Association for Public Opinion Research (AAPOR) response rate definition RR6 (http://www.aapor.org/Standard\_Definitions/1818.htm) was used in the calculations. The weighted response rate includes weights defined by expression (2) multiplied by expression (3) in appendix D. Additional details are provided in chapter 6. 17 With this definition, approximately 8.9 percent of public-school ninth-grade students on the Common Core of Data were excluded from the HSLS:09 target population in the base-year study because they attended a study-ineligible school. The corresponding percentage excluded from the private-school student population was slightly lower at 5.3 percent.

*HSLS:09 Base-Year Data File Documentation NCES Report 35*

*Chapter 3. Sample Design*

• • • • •

*36 HSLS:09 Base-Year Data File Documentation*

Juvenile correction/detention facilities;

Other schools that address disciplinary issues but do not enroll students directly;

Ungraded schools (i.e., no metric to define students as being in the ninth grade);

Schools that only offer testing services for home-schooled students; and

Schools that do not require students to attend daily classes at their facility.

**3.2.2 School Sampling Frame**

The HSLS:09 sample schools were selected from two National Center for Education Statistics (NCES) files. The primary sample of regular public and public charter schools was selected from the 2005–06 Common Core of Data (CCD).18 The private schools were sampled from the 2005–06 Private School Universe Survey (PSS).19

Every attempt was made to identify and exclude study-ineligible schools using data on the NCES files prior to sampling. The following is a complete list of criteria used to exclude schools from the sampling frame and to exclude schools postsampling from the study:

• • • • • • • • *BIA schools*. These schools were located using Federal Information Processing Standards (FIPS) code = 59 (not an official U.S. FIPS code).

*Special education schools*. Schools were classified as ineligible for the study if the NCES school type indicator was set to “special education.” Additional schools were excluded if the school name included words such as “blind,” “unsighted,” “deaf,” or “impaired.”

*Ineligible CTE schools*. Public schools were excluded from sampling if the school type was set to “vocational” and total enrollment size listed on the CCD for the school was zero.

*OCONUS DoD schools*. These schools (Department of Defense schools outside the continental United States) were identified using FIPS code = 58 (not an official U.S. FIPS code).

*Schools without both a 9th and 11th grade*. Indicators to identify the lowest and highest grades of instruction were examined to identify schools without both 9th and 11th grades.

*Not in operation during the fall of 2009*. Closed public schools were identified using the operational status code on the CCD. Closed private schools could not be determined prior to sampling.

*Juvenile correction/detention facilities*. Schools with names containing the words “detention,” “correctional,” or “jail” were excluded from the sampling frame.

*Duplicates.* One record was randomly chosen for those few schools with multiple entries on the corresponding NCES file. Duplicates were identified using school

18 http://nces.ed.gov/ccd/ 19 http://nces.ed.gov/surveys/pss/

*Chapter 3. Sample Design*

name, location address, and administrator name in combination with information obtained from the Internet.

• *Ungraded schools*. If the lowest and highest grade indicators were both “UG” or “00,” the school was classified as ungraded.

If the ninth-grade enrollment count was missing, the information was imputed using the median enrollment count for the corresponding sampling stratum. Enrollment counts were imputed for 41 public school records (0.20 percent) and 237 private school records (3.21 percent) prior to sampling. Sampling frame counts (*schools on frame*) and the number of study-eligible schools (*eligible schools*) is provided in table 5 by school type, region, and locale.20

**Table 5. School sampling-frame eligibility status and number sampled by sampling stratum**

School sampling stratum

*HSLS:09 Base-Year Data File Documentation 37*

Schools on frame1

Eligible schools2 Sampled schools3

‘ n Percent4 ‘ n Percent5 ‘

n Percent6 Total 29,547 100.0 27,293 92.4 1,973 7.2

School type

Public 22,304 75.5 20,505 91.9 1,550 7.6 Private 7,243 24.5 6,788 93.7 423 6.2 Catholic 1,209 4.1 1,199 99.2 198 16.5 Other private 6,034 20.4 5,589 92.6 225 4.0

Region Northeast 4,536 15.4 4,395 96.9 357 8.1 Midwest 7,701 26.1 7,178 93.2 493 6.9 South 10,306 34.9 9,632 93.5 729 7.6 West 7,004 23.7 6,088 79.0 394 6.5

Locale City 7,384 25.0 6,787 91.9 667 9.8 Suburban 6,889 23.3 6,390 92.8 715 11.2 Town 4,323 14.6 3,868 89.5 204 5.3 Rural 10,951 37.1 10,248 93.6 387 3.8 1 Counts of schools listed in the table are from the 2007–08 CCD and 2007–08 PSS. These files were available from NCES at the time the school-level analysis weights were constructed (see chapter 6) and most closely reflect the target population under study. As discussed in section 3.2.2, the HSLS:09 school sample was randomly selected from the 2005–06 CCD and 2005–06 PSS, and supplemented with a sample of new schools from the 2006–07 CCD and 2007–08 PSS. 2 Some schools were classified as ineligible for the study based on sampling frame information. See the discussion at the beginning of section 3.2.2. 3 A large sample was selected for HSLS:09 to ensure a sufficient number of participating schools for the analytic objectives. As discussed at the end of section 3.2.5, only a portion of the sample was recruited for the study and some hold-sample cases were never released. 4 Unweighted percent is based on overall total within column. Percentages may not sum to 100 because of rounding. 5 Unweighted percent is based on the number listed on the sampling frame within each row of the table. 6 Unweighted percent is based on the eligible number of schools within each row of the table. NOTE: CCD = Common Core of Data. NCES = National Center for Education Statistics. PSS = Private School Universe Survey. SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year.

20 School locale is also referred to as “urbanicity”.

*Chapter 3. Sample Design*

Even though HSLS:09 was selected from the most recent NCES files available at the time of sampling (2005–06 CCD for public schools and 2005–06 PSS for private schools), the information contained in the lists was more than 2 years old. To maximize coverage of the intended target populations, random samples of new schools on the 2006–07 CCD and 2007–08 PSS were drawn after the start of school recruitment to supplement the original sample. New schools were identified by (1) eliminating known ineligibles from the new NCES files using the criteria listed above, and (2) merging the “cleaned” NCES files by the respective NCES IDs and separately by school name and location address. All new schools isolated with this process were again compared with the original sampling frames to ensure that they were not previously eligible for the study. Schools were classified as study-ineligible per information on the NCES files for both the original sample and new sample of schools and excluded from the sampling frame. Some sample schools were later reclassified as study-ineligible based on updated information obtained in the field during recruiting.

**3.2.3 First-Stage Sample Design**

A stratified probability proportional to size (PPS) sample of schools was selected for HSLS:09 (table 5) with the goal of producing national estimates on characteristics associated with, for example, high school success and family influences in education choices. Within each first-stage stratum, samples were selected using Chromy’s sequential probability with minimum replacement sampling algorithm (Chromy 1981). The composite measure of size (*mos*) used in the sampling procedure was calculated as a linear combination of student counts multiplied by the desired overall sampling rates within race/ethnicity. Details of the sample design are found in appendix D; the probabilities of selection are discussed in chapter 6 as they relate to the analysis weights. A total of 48 mutually exclusive first-stage sampling strata were created for HSLS:09. The strata were defined by cross-classification of three variables:

• • •

*38 HSLS:09 Base-Year Data File Documentation*

School type or sector (public, private–Catholic, private–other);

Region of the United States (Northeast, Midwest, South, West); and

Locale (city, suburban, town, rural). All study-eligible schools on the CCD were given a *school type* classification of public. A distinction between regular public and public charter schools was not made for the purposes of sampling. School type on the PSS was determined by whether the religious orientation/affiliation variable was set to “Roman Catholic.” All non-Catholic PSS private schools were classified in the private–other category.

Within school type, the eligible schools were classified into four *regions* of the United States for the second stratification variable. The following assignments were made based on the FIPS state code associated with the physical location of the school: