NATIONAL CENTER FOR EDUCATION STATISTICS

Institute of Education Sciences

National Household Education Surveys Program of 2007

Data File User's Manual, Volume I

Study Overview and Methodology

National Household Education Surveys Program of 2007

Data File User's Manual, Volume I

Study Overview and Methodology

October 2008

Mary Hagedorn
Shelley Brock Roth
Kevin O'Donnell
Susan Smith
Westat

Gail Mulligan
National Center for Education Statistics

U.S. Department of Education
Margaret Spellings
Secretary

Institute of Education Sciences
Grover J. Whitehurst
Director

National Center for Education Statistics
Mark Schneider
Commissioner

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

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

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

> National Center for Education Statistics
> Institute of Education Sciences
> U.S. Department of Education
> 1990 K Street, NW
> Washington, DC 20006–5650

October 2008

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

Suggested Citation

Hagedorn, M., Roth, S.B., O'Donnell, K. Smith, S., and Mulligan, G. (2008). *National Household Education Surveys Program of 2007: Data File User's Manual, Volume I.* (NCES 2009-024). *National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.*

Contact:
Gail Mulligan
(202) 502-7491
(e-mail) nhes@ed.gov

ACKNOWLEDGMENTS

This page is intentionally blank.

TABLE OF CONTENTS

TABLE OF CONTENTS--CONTINUED

TABLE OF CONTENTS—CONTINUED

LIST OF TABLES

TABLE OF CONTENTS—CONTINUED

LIST OF TABLES

TABLE OF CONTENTS—CONTINUED

LIST OF FIGURES

This page is intentionally blank.

1. INTRODUCTION

The National Household Education Surveys Program (NHES) was developed by the National Center for Education Statistics (NCES) and incorporates random digit dial (RDD) telephone surveys of households in the United States. The RDD telephone data collection for the 2007 administration (NHES:2007) was conducted by Westat from January 2 through May 6, 2007. In addition, a nonresponse bias study involving both telephone and in-person data collection was conducted from January 2 through July 24, 2007. In the NHES:2007 Screener Survey, household members were enumerated and demographic and educational information that determined eligibility for the two distinct topical surveys was collected. The NHES:2007 surveys are as follows:

- The School Readiness Survey (SR-NHES:2007) gathered information on children's enrollment in preschool and other center-based care and education arrangements, plans for kindergarten enrollment, developmental accomplishments and difficulties including emerging literacy and numeracy skills, educational activities with family members including family reading, television viewing, and health and disability status. Parents of children in elementary school were also asked about early school experiences, television viewing, and family reading.

- The Parent and Family Involvement in Education Survey (PFI-NHES:2007) addressed school choice, homeschooling, family involvement with children's schools, school efforts to involve families, parent involvement with children's homework, tutoring, parent and family involvement in activities outside of school, and child health and disability status.

The populations of interest in the NHES:2007 surveys were:

- Children age 3 through age 6 who were not yet enrolled in kindergarten, of interest in the SR Survey; and

- Children and youth enrolled in kindergarten through 12th grade or homeschooled for these grades, of interest in the PFI Survey.

The *National Household Education Surveys Program of 2007: Data File User's Manual* provides documentation and guidance for users of the NHES:2007 public-use and restricted-use data files for the SR and PFI surveys. The manual is composed of three volumes. Information about the purpose of the study, the data collection instruments, the sample design, data collection, and data processing procedures is included in Volume I. The data collection instruments, a chart summarizing weighting and sample variance estimation variables for all NHES surveys, and tables comparing NHES:2007 estimates[1] to those of other surveys are contained in appendixes to Volume I. Volumes II and III each address the data files for the 2007 SR and PFI data files, respectively. They each contain a guide to the data files, a discussion of data considerations and anomalies and, in appendixes, the data file layouts, and derived variable specifications.

---

[1] All survey estimates are weighted estimates computed using the appropriate final survey weight (the variable FPWT for estimates from the PFI survey and the variable FSWT for estimates from the SR survey).

The data files contain the following:

- The SR-NHES:2007 file includes data from interviews completed with parents or guardians of 2,633 preschoolers.[2]

- The PFI-NHES:2007 file contains data from interviews completed with parents or guardians of 10,681 children in kindergarten through 12th grade, including 10,370 students enrolled in public or private schools and 311 homeschooled students.

The NHES data are subject to federal law on data confidentiality; see chapter 5, section 5.1, for more information.

1.1   Background of Study

NHES was developed by NCES to complement its institutional surveys. NHES is the principal mechanism for addressing topics that cannot be addressed in institutional data collections. By collecting data directly from households, NHES allows NCES to gather data on a wide range of issues, such as early childhood care and education, children's readiness for school, parent perceptions of school safety and discipline, before- and after-school activities of school-age children, participation in adult and continuing education, parent involvement in education, and civic involvement. NHES uses RDD and computer-assisted telephone interviews (CATI). The survey has been conducted by Westat in the winter and spring of 1991, 1993, 1995, 1996, 1999, 2001, 2003, 2005, and 2007. As shown in table 1-1, each administration has included more than one survey.

The first test of NHES was a large field test conducted in the fall of 1989. This effort, which included the screening of about 15,000 households, included surveys on the following two topics: school dropouts (interviews were conducted with adult household respondents and 14- to 21-year-old youths) and early childhood education (interviews were conducted with parents/guardians[3] of 3- to 5-year-olds). The design of the field test and the results of the field test data collection activities are described in an *Overview of the NHES Field Test* (Brick et al. 1992).

---

[2] Throughout this report, the subgroup of children age 3 through 6 not yet enrolled in kindergarten is referred to simply as "preschoolers." Some of these children were reported to be enrolled in school with a grade of N (preschool, nursery school, prekindergarten, or Head Start).

[3] Respondents need not have been parents or legal guardians. The household member most knowledgeable about the child's care and education was identified by the Screener respondent and selected to respond to the survey. The respondent was usually, but not always, a parent.

Table 1-1.   Surveys conducted under the National Household Education Surveys Program, by years administered: 1991 through 2007

Survey topics	NHES survey administration								
	1991	1993	1995	1996	1999[1]	2001	2003	2005	2007
Early childhood education/program participation	√		√		√	√		√	
Adult education	√		√		√	√	√	√	
School readiness		√			√				√
School safety and discipline		√							
Parent and family involvement in education				√	√		√		√
Civic involvement				√	√				
After-school programs and activities			√[2]		√	√[3]		√	
Household and library use					√				

[1] NHES:1999 was a special end-of-decade administration that measured key indicators from the surveys fielded during the 1990s.

[2] These items were only asked about children in first through third grades.

[3] The NHES:2001 survey about after-school programs and activities (ASPA) also included before-school programs.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Surveys Program (NHES), 1991, 1993, 1995, 1996, 1999, 2001, 2003, 2005, and 2007.

1.2   NHES Survey Topics

This section provides more detailed information on the topical areas addressed in the full-scale NHES administrations and the survey populations associated with each topic.

- **Early Childhood Education/Program Participation Surveys, ECE/ECPP** (1991, 1995, 2001, and 2005). The ECE/ECPP surveys provide cross-sectional, national estimates of children's participation in care by relatives and non-relatives in private homes and in center-based daycare or preschool programs (including Head Start and Early Head Start). Additional topics addressed in ECE/ECPP interviews have included family learning activities, out-of-pocket expenses for nonparental care, continuity of care, factors related to parental selection of care, parents' perceptions of care quality, delayed kindergarten entry and grade retention, child health and disability, and child, parent, and household characteristics.

  ECE/ECPP collections have included interviews with parents of children from age three through third grade (1991, n = 13,298 children, 1995, n = 14,064) and children from birth through age 6 and not yet in kindergarten (2001, n = 6,749 children; 2005, n = 7,209). Some ECE/ECPP questions were included in the Parent Survey of NHES:1999 (n = 6,939 infants, toddlers, and preschoolers).

- **Adult Education/Adult Education and Lifelong Learning Surveys, AE/AELL** (1991, 1995, 1999, 2001, and 2005). The Adult Education surveys provide cross-sectional, national estimates of educational participation in basic skills/GED courses, English as a Second Language (ESL) courses, college and university degree and certificate programs, vocational/technical diploma or degree programs, apprenticeships, work-related courses,

courses taken for personal development or personal interest, and informal learning. Additional topics covered in the AE surveys include the subject matter of courses or programs, course duration, out-of-pocket expenditures, location and sponsorship, employer support, interest in educational activities, and barriers to participation in educational activities. Information on adult and household characteristics was also collected.

Sample sizes for the Adult Education surveys have ranged from 6,697 to 19,722 non-institutionalized adults age 16 and older, not enrolled in 12th grade or below, and not on active duty in the U.S. armed forces.

- **Adult Education for Work-Related Reasons Survey, AEWR** (2003). The AEWR Survey provides cross-sectional, national estimates of participation in college and university degree or certificate programs taken for work-related reasons, vocational/technical diploma or degree programs taken for work-related reasons, apprenticeships, work-related courses, and informal learning activities related to a job or career. In NHES:2003, interviews were conducted with 12,725 noninstitutionalized adults age 16 and older, not enrolled in 12th grade or below, and not on active duty in the U.S. armed forces.

- **School Readiness Surveys, SR** (1993, 2007). The SR Survey provides cross-sectional, national estimates of children's developmental accomplishments and difficulties including emerging literacy and numeracy, center-based program participation, educational activities with family members, and health and nutrition. Parents of preschoolers were also asked about their perceptions of skills or knowledge required to start school. In SR-NHES:1993, parents of children in elementary school were also asked about their child's adjustment to school, including feedback from teachers, and early school experiences. Information about family stability and other risk factors was collected along with parent/guardian and household characteristics.

  The SR-NHES:1993 collection included interviews with parents of children ages 3 to 7 in second grade or below and children ages 8 and 9 who were in first or second grade (n = 10,888). Literacy and numeracy items from SR 1993 were asked in Parent-NHES:1999 (n = 3,631 preschoolers), ECPP-NHES:2001 (n = 3,150 preschoolers), and ECPP-NHES:2005 (n = 3,354 preschoolers).

  The SR-NHES:2007 collection, described further in section 1.3, provides current information on the developmental status and school preparedness of preschool children. In addition to providing cross-sectional, national estimates, SR-NHES:2007 provides the ability to measure change in the status of preschoolers over time.

- **School Safety and Discipline Survey, SSD** (1993). The School Safety and Discipline Survey provides national estimates related to the school learning environment, discipline policy, safety at school, victimization, availability and use of alcohol and drugs, and alcohol and drug education as reported by youth and their parents. Youth were also asked about peer norms for achievement and behavior in school and substance use. Child, parent/guardian and household characteristics were also collected. The SSD collection included interviews conducted with parents of 12,680 students in grades 3 through 12 and with 6,504 youth in grades 6 through 12.

- **Parent and Family Involvement in Education Surveys, PFI/CI** (1996) **and PFI (**2003, 2007). The PFI surveys addressed specific ways that families are involved in their children's school, or center-based early childhood program (1996 only), school practices to involve and support families, involvement with children's homework, and involvement in educational activities outside of school. Parents of homeschoolers were asked about their reasons for choosing homeschooling and resources they used in homeschooling. The interviews also included questions about child, parent, and household characteristics. PFI collections have included interviews with parents of children from age 3 through grade 12 (1996, n = 20,792) and parents of children in kindergarten through grade 12 (2003, n = 12,426). Some PFI questions were incorporated in the Parent Survey of NHES:1999 (n = 24,600).

  PFI-NHES:2007, described further in section 1.3, provides current information about parent and family involvement in school, involvement in homework, and involvement in educational activities outside of school. In addition to providing cross-sectional, national estimates, PFI-NHES:2007 provides the ability to measure change over time.

- **Civic Involvement: Parent and Family Involvement in Education and Civic Involvement in Education Survey, Youth Civic Involvement Survey, Adult Civic Involvement Survey, PFI/CI** (1996)**, YCI** (1996, 1999)**, ACI** (1996). The PFI/CI and YCI surveys focused on the civic involvement of students and their parents. These surveys provide an assessment of the ways that parents and other adults can socialize children for informed civic participation and of opportunities youth have for participation in community service. The Youth Civic Involvement Survey (YCI) collection (n = 8,043) and Youth-NHES:1999 (n = 7,913) provide national estimates for 6th- through 8th-graders and 9th-through 12th-graders. The Adult Civic Involvement Survey (1996, n = 2,250) included interviews that could be used to compare adults in households without children age 3 through grade 12 to adults in households with children in this age/grade range.

- **After-School Programs and Activities Surveys (ASPA)** (2001, 2005). The ASPA surveys provide information about children's participation in care by relatives or nonrelatives in private homes, in school- or center-based programs, and in after-school activities. Parents were also asked about children's self-care. Information about the specific activites in which children were engaged during the after-school hours was collected. In addition, parents were asked about continuity of care arrangements, parental perceptions of care quality, and reasons for choosing parental care. ASPA-NHES:2001 included questions about before-school programs as well as after-school programs. Information on child, parent/guardian, and household characteristics was collected.

  ASPA collections have included interviews with parents of sampled children enrolled in kindergarten through eighth grade (1999, n = 12,396 children; 2001, n = 9,583 children; and 2005, n = 11,684 children).

- **Household and Library Use Survey (HHL)** (1996). The HHL survey examined public library use by household members, including the ways that they used public libraries (e.g., borrowing books, attending lectures, attending story hours) and the purposes for using public libraries (e.g., for school assignments, enjoyment, work-related projects). Demographic and educational information also was collected about each household member. HHL-NHES:1996 provide cross-sectional, national estimates of household characteristics and library use for all households in the United States, as well as estimates by state. This survey was administered to the 55,708 households that completed screeners in 1996.

1.3   NHES:2007 Surveys

The two surveys that compose NHES:2007 focus on topics that have been addressed in previous NHES administrations. SR-NHES:2007 includes topics addressed in the SR survey of NHES:1993. In addition some items in the SR survey, including emerging literacy and numeracy, have been included in Parent-NHES:1999, ECPP-NHES:2001 and ECPP-NHES:2005.   PFI-NHES:2007 includes topics addressed in PFI/CI-NHES:1996, Parent-NHES:1999, and PFI-NHES:2003.

School Readiness Survey

SR-NHES:2007 collected information on early learning and readiness for entering school, specifically, participation in preschool or other types of center-based care and education, including Head Start, children's developmental accomplishments including literacy and numeracy skills, educational activities with family members, plans for kindergarten enrollment, and the role of the parent in preparing their child for kindergarten. The survey also addressed the amount and type of television viewing by preschoolers. In SR-NHES:2007, parents of 2,633 sampled children ages 3 to 6 responded to the SR survey. In addition, some School Readiness questions were asked of parents of children in kindergarten through second grade, for example, early school experiences and family reading. Parents of children in kindergarten through second grade also answered items of interest to school readiness researchers in the NHES:2007 Parent and Family Involvement in Education Survey; their responses are included in the PFI data file.

Parent and Family Involvement in Education Survey

PFI-NHES:2007 collected information on school choice, homeschooling, school characteristics (including school type, lowest and highest grades at the school, school religious affiliation, and whether the school was a magnet or charter school), student experiences in school, teacher feedback on the child's school performance and behavior, family involvement in school, family help with homework, family involvement in activities outside of school, and factors affecting family involvement. The PFI-NHES:2007 file contains data from interviews completed with parents of 10,681 sampled children in kindergarten through 12th grade, including 10,370 students enrolled in public or private schools and 311 homeschooled children.

NHES:2007 Survey Design Activities

The NHES:2007 topical surveys drew heavily upon design work that was conducted for prior NHES administrations. In addition to considering the NHES instruments used previously for the same survey topics, other survey design activities were undertaken in order to ensure that the data resulting from the NHES:2007 administration addressed emerging issues and those of concern to researchers and policymakers.

- Westat staff consulted with experts in academic and research institutions and government agencies to obtain their perspectives on the survey topics. The experts were asked to identify important research areas of interest and to comment on the relative priority of specific areas of survey content. Telephone conferences were held with 9 SR experts and 10 PFI experts.

- A Technical Review Panel (TRP) was established for each survey. Due to overlap in the interests of researchers in school readiness and parent and family involvement in education, TRP members for the two surveys met both jointly and separately to discuss survey content.

- Survey staff also examined extant surveys to assess the content areas addressed and the items used to measure survey concepts.

- Survey staff conducted reviews of the relevant literature, drawing upon professional journals, scholarly books, and government reports.

- A set of research questions was developed for each survey. These research questions identified the content areas to be addressed and provided a means to map the survey instruments to content areas to ensure sufficient coverage of important issues.

The draft survey instruments underwent cognitive testing to assess respondent comprehension of the questions, their knowledge of the information requested, and the sensitivity of survey items. In addition, the instruments were field tested by telephone to evaluate interview flow, administration time, areas of respondent confusion, and items that were difficult for respondents to answer. Each of these activities contributed to the development of the SR and PFI surveys.

1.4   Overview of NHES Design

The SR and PFI surveys were developed to provide reliable national estimates. The surveys were conducted simultaneously because of the high costs associated with screening large numbers of households in order to meet the sample size requirements for precise estimates. By addressing more than one topic in NHES:2007, the cost of screening households to find eligible household members could be partitioned over the surveys. This strategy is key to the NHES design.

Another feature of NHES, within-household sampling, was developed in response to concerns about the burden placed upon households in which more than one household member could be sampled or the same household member would be asked to respond to multiple surveys. A Screener Survey was used to collect information on household composition and interview eligibility. One preschooler (age 3 through 6 and not enrolled in kindergarten or higher grades) was selected for the SR survey in every screened household that included one or more preschoolers. In addition, one child or youth in kindergarten through grade 12 (or homeschooled for one of these grades) was selected for the PFI survey in every screened household that included one or more such children. (See chapter 3 for a detailed discussion of precision requirements and sampling procedures for NHES:2007.)

Even though sampling methods reduced the number of interviews per household, the length of the interview was considered to be an important factor in obtaining good response rates[4] and reliable estimates. Therefore, the number of items included in the NHES:2007 surveys was limited in order to help improve response rates and reduce the demands made on survey respondents. The overall average administration time for the Screener was 3.14 minutes. The average administration time for the SR interview was 22.5 minutes; for the PFI interview it was 27.7 minutes.

Because of the complexity of the instruments and the complex sampling techniques used, NHES:2007 was conducted using computer-assisted telephone interviewing (CATI) technology. Some of the advantages of CATI include improved project administration, online sampling and eligibility checks,

---

[4] Unless stated otherwise, all unit response rates are weighted rates computed using the appropriate base weight.

scheduling of interviews according to a priority scheme to improve response rates, managing data quality by controlling skip patterns and checking responses during the interview for range and consistency, and a "help" function for a limited number of items to assist interviewers in answering respondents' questions during the interview. Items within each of the NHES:2007 instruments were programmed so that the appropriate items appeared on the interviewers' computer screen according to the respondents' answers to previous questions.

Table 1-2 summarizes the number of completed interviews and gives weighted unit response rates and overall unit response rates for the Screener and the SR and PFI surveys. Table 1-3 gives unweighted unit response and overall unit response rates for the Screener and the SR and PFI surveys. Because refusal conversion procedures were not used for non-subsampled cases, their unweighted response rates are lower than they would have been if such procedures had been used. There is no reason to expect that the response rate for this group would have been measurably different from the response rate for those cases that were subject to refusal conversion. Therefore, assuming the response rates for these groups are equal, the response rate would be unaffected by inclusion or exclusion of these cases from the response rate calculation. In contrast, including them in the calculation of the unweighted rate would definitely result in a lower response rate since the calculation would not incorporate statistical adjustments that take the subsampling into account. More details on the computation of these rates, including a discussion of the uses of weighted and unweighted response rates, are given in chapter 4.

1.5   Flow of the Interviews

Figure 1-1 shows the flow of the NHES:2007 interviews. Each household contact began with a Screener to obtain information used to sample household members for extended interviews.

If the household contained any children age 3 to 6 and not yet in kindergarten, exactly one such child was selected as the subject of an SR interview. If the household contained any children enrolled in kindergarten through grade 12 (or homeschooled for these grades), exactly one such child was selected as the subject for a PFI interview. The respondent for an SR or PFI interview was the parent or guardian in the household most knowledgeable about the child's care and education. (See chapter 3 for additional details about the sample.)

Whenever possible, all interviews with household members were conducted during the same telephone call as the Screener. Followup calls were made to complete interviews that were not completed during the initial contact.

1.6   Contents of Manual

The chapters that follow in Volume I provide additional information about the survey instruments (chapter 2), the sample design and estimation procedures (chapter 3), data collection and response rates (chapter 4), and data preparation (chapter 5). Appendix A provides a copy of the Screener and the SR/PFI questionnaires. Appendix B contains a summary of weighting and sample variance estimation variables. Appendix C contains tables comparing NHES:2007 estimates to those of other surveys. Volumes II and III of the *NHES:2007 Data File User's Manual* provide information on the SR and PFI data files, respectively. Each contains a guide to the relevant data file and codebook, a discussion of data considerations and anomalies, and, in appendixes, the file layouts, derived variable specifications, and codebooks.

Table 1-2. Summary of completed interviews and weighted unit response and overall unit response rates, by survey: 2007

Interview type	Number of completed interviews	Unit response rate[1]	Overall unit response rate[2]
Screener	54,034	52.8	52.8
SR survey	2,633	77.0	40.7
PFI survey	10,681	74.1	39.1

[1] The unit response rate is the percentage of completed interviews for a specific stage of the survey (i.e., the Screener, SR, or PFI interview). It is a ratio of the number of completed interviews to the number of units (e.g., households and household members) sampled for the interviews. For many telephone numbers sampled for the Screener interview, no contact was ever made. Based on results of the vendor-assisted method calculations, 37.6 percent of these numbers were assumed to be residential and were added to the denominator for the calculation of the Screener unit response and overall unit response rates. Additionally, the Screener unit response rate accounts for the subsampling of cases for nonresponse followup, which is discussed further in section 4.1.1.

[2] The overall unit response rate indicates the percentage of possible interviews that have been completed, taking all sampling stages into account. The overall unit response rate and the unit response rate are identical for the first stage of sampling and interviewing (i.e., the Screener). For the SR or PFI surveys, the overall unit response rate is the product of the Screener unit response rate and the interview unit response rate (e.g., for the SR survey, the calculation for the overall unit response rate is 100 x (0.528 x 0.770) = 40.7).

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Readiness (SR) Survey of the National Household Education Surveys Program (NHES), 2007; and Parent and Family Involvement in Education (PFI) Survey of the NHES, 2007.

Table 1-3. Summary of completed interviews and unweighted unit response and overall unit response rates, by survey: 2007

Interview type	Number of completed interviews	Unweighted unit response rate[1]	Unweighted overall unit response rate[2]
Screener	54,034	53.2	53.2
SR survey	2,633	78.8	41.9
PFI survey	10,681	76.0	40.4

[1] The unit response rate is the percentage of completed interviews for a specific stage of the survey (i.e., the Screener, SR, or PFI interview). It is a ratio of the number of completed interviews to the number of units (e.g., households and household members) sampled for the interviews. For many telephone numbers sampled for the Screener interview, no contact was ever made. Based on results of the vendor-assisted calculations, 37.6 percent of these numbers were assumed to be residential and were added to the denominator for the calculation of the Screener unit response and overall unit response rates. Only cases subsampled for followup were included in the calculation of the unweighted rates.

[2] The overall unit response rate indicates the percentage of possible interviews that have been completed, taking all sampling stages into account. The overall unit response rate and the unit response rate are identical for the first stage of sampling and interviewing (i.e., the Screener). For the SR or PFI surveys, the overall unit response rate is the product of the Screener unit response rate and the interview unit response rate (e.g., for the SR survey, the calculation for the overall unit response rate is 100 x (0.532 x 0.788) = 41.9).

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Readiness (SR) Survey of the National Household Education Surveys Program (NHES), 2007; and Parent and Family Involvement in Education (PFI) Survey of the NHES, 2007.

Figure 1-1.   Flow of the NHES:2007 interviews
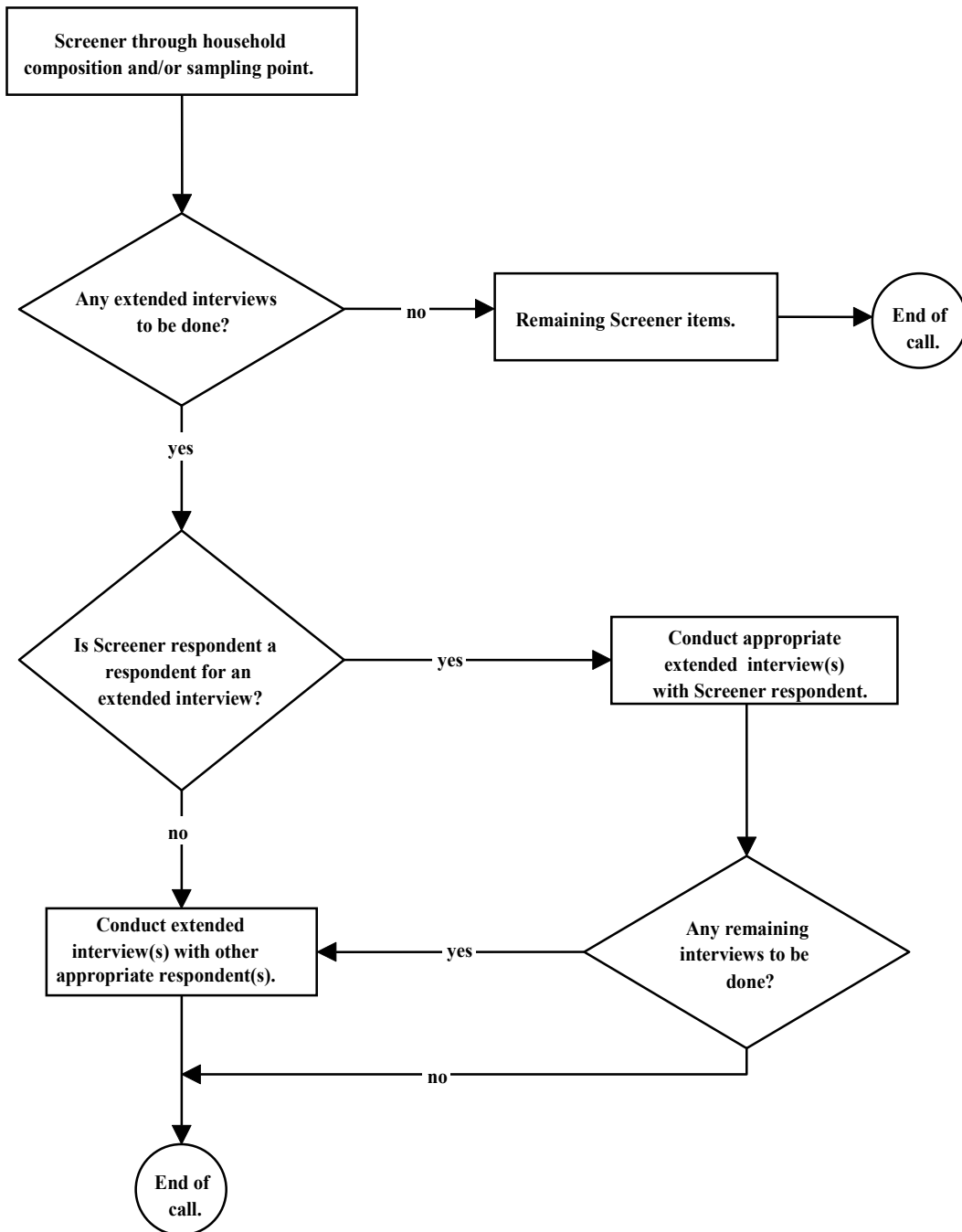
2. DESCRIPTION OF DATA COLLECTION INSTRUMENTS

The sections that follow describe the instruments used to collect data contained in the School Readiness Survey (SR) and Parent and Family Involvement in Education Survey (PFI) data files of the 2007 National Household Education Surveys Program (NHES:2007). A household Screener Survey was used to determine eligibility for the extended interviews, which were used to collect data on the topical areas of interest. The SR and PFI Surveys contain many common items, and a small subset of SR items was asked for children in kindergarten through second grade. In addition, it was expected that some children initially sampled as the subject of an SR interview would be found to actually be eligible for PFI instead (or vice versa) once the most knowledgeable parent was asked key eligibility questions. Taking these factors into consideration, the two surveys were combined into a single extended interview questionnaire for data collection purposes. Appendix A contains the NHES:2007 Screener Survey and the combined SR/PFI questionnaire.

2.1 The NHES:2007 Screener

The screening interview in NHES:2007 was used to determine whether sampled telephone numbers belonged to households, gather the information needed to sample children as SR or PFI interview subjects, select the appropriate parent/guardian respondent for extended interviews, and administer some household items for statistical purposes. The Screener was designed to accomplish these tasks efficiently, placing minimum burden on the respondent.

The first series of questions in the Screener determined whether the telephone number was residential and whether the person on the telephone was eligible to answer the questions. If it was determined that the telephone number was used for business only, the call was terminated. The survey continued for numbers that were for household use or for both household and business use.

If the person who answered the telephone was not a household member or was a household member under 18 years of age, an appropriate Screener respondent was requested. If no member of the household was 18 years old or older, a person designated as the male or female head of household was eligible to be the Screener respondent.

The order of subsequent Screener questions varied depending upon whether the household contained any members age 20 or younger. In households in which there were children or youth age 20 or younger (indicating that someone could potentially be eligible for an SR or PFI interview) the household members were enumerated (i.e., the name, age, and sex of each person were collected). To determine eligibility for SR and PFI, Screener questions asked whether household members ages 3 through 20 were attending school, whether children ages 4 through 18 were being homeschooled, and the grade or year of school in which they were enrolled. If a child/youth was sampled for an SR or PFI interview, the parent/guardian in the household who was most knowledgeable about his/her education was selected to be the respondent for the SR or PFI interview. In households not containing persons under the age of 21, additional household information (e.g., own/rent and other telephone numbers) was collected and was used to facilitate the development of weighting adjustments.

The Screener respondent was asked whether the home was owned or rented and whether there were other telephone numbers in the household for home use. This information was used for weighting and nonresponse adjustment. Then the Screener interview was terminated.

2.2   The NHES:2007 School Readiness Survey (SR)

The School Readiness Survey focused on the experiences and developmental accomplishments of children ages 3 to 6 as of December 31, 2006, who had not yet started kindergarten. Questions concerned enrollment in preschool, participation in early childhood programs and daycare centers, kindergarten plans, developmental accomplishments and difficulties (e.g., early literacy and numeracy skills), family learning activities, the child's television viewing, and parents' perceptions of their role in preparing their children to start school. In addition, the survey included questions about child characteristics, child health and disability, parent/guardian characteristics, and household characteristics. To avoid redundancy, household information was collected only during the first interview conducted in each household. Similarly, parent/guardian information was collected only once per household, unless sampled children in the same household had different parents.

The respondent for the SR interview was the parent or guardian in the household who was the most knowledgeable about the child's care and education. Typically, this was the mother of the child; however, the respondent could be a father, grandparent, another relative, or nonrelative designated as the most knowledgeable household member. For simplicity, when referring to the most knowledgeable respondent in the manual, this person will be called the parent.

Selected SR questions were also administered during PFI interviews with parents of children enrolled in kindergarten through second grade. Those items appear in the PFI data file as well.

2.3   The NHES:2007 Parent and Family Involvement in Education Survey (PFI)

The Parent and Family Involvement in Education Survey focused on children and youth enrolled in kindergarten through 12th grade who were age 20 or younger as of December 31, 2006, and those age 4 through age 18 who were homeschooled for kindergarten through 12th grade. In the PFI interview, subjects were routed to one of four questionnaire paths: elementary (kindergarten through 5th grade), middle school (6th through 8th grades), senior high school (9th through 12th grade), or home school (kindergarten through 12th grade). The elementary, middle school, and senior high school paths all included items about school choice, parent and family involvement at school, involvement in schoolwork, involvement and activities outside of school, factors affecting family involvement, and parents' communication with other parents. Parents of homeschoolers were asked a special set of questions about their reasons for homeschooling and the resources they use in doing so. For all questionnaire paths, parents were asked about child characteristics, the child's health and disability status, parent/guardian characteristics, and household characteristics. Again, to avoid redundancy, household information was collected only during the first interview conducted in each household. Similarly, parent/guardian information was collected only once per household, unless sampled children in the same household had different parents.

The respondent for the PFI interview was the parent or guardian in the household who was the most knowledgeable about the child's care and education. Typically, this was the mother of the child; however, the respondent could be a father, grandparent, another relative, or nonrelative designated as the most knowledgeable household member.

3. SAMPLE DESIGN AND IMPLEMENTATION

This chapter describes the sample design for the 2007 National Household Education Surveys Program (NHES:2007), including a number of special features of the design. Also presented are the procedures for weighting, variance estimation, and imputation for items that had missing values.

3.1   Precision Requirements for NHES:2007

The number of telephone numbers required for NHES:2007 was determined by the precision requirements for the estimates from the School Readiness Survey (SR) and the Parent and Family Involvement in Education Survey (PFI). The general precision requirement for both surveys was the ability to detect a 10 to 15 percent relative change for an estimate between 30 and 60 percent. However, because some key estimates (described below) were expected to fall outside the 30 to 60 percent range (e.g., literacy and numeracy in the SR survey), the sample size was increased above that needed for the general precision requirement to support estimation for those key estimates.

In NHES:2007, the overall screening sample was largely determined by the need to produce precise estimates of indicators for the populations covered by the SR and PFI surveys, particularly preschoolers (ages 3 to 6 and not yet in kindergarten).[5] For the SR survey, key sample size determinants were the requirements to detect changes in estimates of participation in center-based care and education arrangements, literacy skills, and numeracy skills, by single year of age and by race/ethnicity (White, non-Hispanic; Black, non-Hispanic; and Hispanic). The key estimates were selected to represent statistics that have been published from the SR-NHES:1993, Parent-NHES:1999, ECPP-NHES:2001, and ECPP-NHES:2005 surveys, and to reflect topics of interest to experts in the field. The subgroups were chosen because they are key subgroups used in analyses of NHES data for preschoolers. The key estimates considered in designing the sample for PFI were the percentage of children whose parents report that school practices[6] were done very well, and the percentage of children whose parents participated in six or more home learning activities;[7] the key analytic subgroups were race/ethnicity (the White, non-Hispanic; Black, non-Hispanic; and Hispanic subgroups), 2-year grade groups, parents' educational attainment (high school diploma or below, beyond high school diploma), school type (public, private), and school size (under 300; 300–599; 600–999; 1,000 or more). The key estimates were selected to represent statistics that have been published from the PFI/CI-NHES:1996, Parent-NHES:1999, and PFI-NHES:2003 surveys, to reflect topics of interest to experts in the field, and to include measures of both in-school and out-of-school involvement. The subgroups were chosen because they are key subgroups used in analyses of NHES data for school-age children. As a result, targets of about 3,790 completed SR interviews and 14,150 completed PFI interviews were established.

---

[5] Throughout this report, the subgroup of children age 3 through 6 not yet enrolled in kindergarten is referred to simply as "preschoolers." Some of these children were reported to be enrolled in school with a grade of N (preschool, nursery school, prekindergarten, or Head Start).

[6] The school practices considered were the following: School tells family how child is doing in school; school helps family understand child's development; school tells about chances to volunteer; school advises about home learning; and school gives information about community services.

[7] The home learning activities considered were the following: Telling the child a story; teaching the child letters, words, or numbers; teaching the child songs or music; working on arts or crafts with the child; taking the child along on errands; involving the child in household chores; taking the child to the library; taking the child to a play, concert, or other live show; taking the child to an art gallery, museum, or historical site; taking the child to a zoo or aquarium; talking with the child about his/her family history or ethnic heritage or related issues; or attending an event sponsored by a community, ethnic, or religious group.

Taking into account all stages of sampling and expected response, a goal of screening about 62,000 households was established. However, a lower than expected residency rate (i.e., the proportion of telephone numbers that are assigned to households) and a lower than expected unit response rate[8] caused a revision of the targets for screened households and numbers of completed extended interviews. The effect on the precision of the estimates was examined and found to be minimal. As shown in table 3-2, the final numbers of completed interviews were 54,034 Screeners, 2,633 SR interviews and 10,681 PFI interviews.

3.2   Sampling Households

Different methods have been developed over the years for selecting random samples of telephone households. Since NHES:1995, a list-assisted method, described by Casady and Lepkowski (1993), has been used for the NHES surveys. This method reduces the number of unproductive calls due to nonworking or nonresidential numbers (compared with simple random sampling of all numbers), produces a self-weighting sample, and results in a single-stage and unclustered sample. With the list-assisted method, an equal-probability random sample of telephone numbers is selected from all telephone numbers that are in 100-banks (numbers in a 100-bank have the same first 8 digits of the 10-digit telephone number) in which there is at least one residential telephone number listed in the white pages directory (the listed stratum). Both listed and unlisted telephone numbers are included in the listed stratum. Telephone numbers in 100-banks with no listed telephone numbers (the zero-listed stratum) are not sampled.

The sampling frame for the NHES:2007 RDD sample was MSG's Genesys frame of all telephone numbers in 100-banks with one or more telephone numbers listed in the white pages in the third quarter of 2006. MSG is a commercial firm that has produced samples of telephone numbers for previous NHES studies. A stratified list-assisted sample (described below) was used in order to support design goals for national-level and subdomain statistics for the SR and PFI surveys of NHES:2007.

Because NHES is a telephone survey, undercoverage bias resulting from differences between telephone and nontelephone households is a concern. Undercoverage bias is the average difference between the survey estimate and the population parameter being estimated that results from some members of the inference population being excluded from the sampling frame. For example, while NHES is conducted using a sample of telephone households, the inference population includes both telephone and nontelephone households, so undercoverage bias could result from the exclusion of persons in nontelephone households. Differences in telephone coverage rates, especially differential rates among population subgroups, such as those defined by region, age, race/ethnicity, and household composition, are of concern to telephone survey methodologists because they can introduce bias in the estimates.

The largest component of coverage bias in a telephone survey such as NHES is likely due to the prevalence of households without landline telephones and differences between such households and those with landline telephones. Blumberg, Luke, and Cynamon (2006) examined differences in characteristics among persons and households having no telephone service, cellular service only, and landline service (including both landline only, and landline and cellular). Although there are differences in landline coverage (e.g., young adults, adults in one-person households, and renters are less likely to have landline telephones), raking to population totals for these subgroups is used in NHES to statistically adjust for and

---

[8] The final unweighted residency and Screener unit response rates for NHES:2007 were 38 percent and 53 percent, respectively. These are lower than the expected rates of 45 percent and 64 percent, respectively.

reduce undercoverage bias. Various studies have been undertaken to examine the undercoverage bias for key subgroups in NHES. Brick, Burke, and West (1992) looked at undercoverage bias for 3- to 5-year-olds and 14- to 21-year-olds. Brick (1996) examined undercoverage bias for 0- to 2-year-olds and adults. Undercoverage bias for 3- to 7-year-olds was examined by Brick et al. (1997). Undercoverage bias for estimates of characteristics of households and for adults was investigated by Montaquila, Brick, and Brock (1997). Another potential source of undercoverage bias in telephone surveys that use the list-assisted method is the fact that not all telephone households are included in the sampling frame. Households in the zero-listed stratum have no chance of being included in the sample. Empirical findings were presented in Brick et al. (1995) to address the question of coverage bias associated with excluding the zero-listed stratum. Tucker, Lepkowski, and Piekarski (2002) found that about 3 percent of telephone households are in the zero-listed stratum. The results also indicate that households in the zero-listed stratum are not very different from households in the listed stratum. Because the proportion of telephone households that are in the zero-listed stratum is small and the persons living in these households are not very different from those living in households in the listed stratum, the bias resulting from excluding the zero-listed stratum is generally very small.

Results from these studies suggested that undercoverage bias was not a significant problem in NHES. However, since these studies were conducted (in the early- to mid-1990s), landline telephone coverage has changed considerably with the increasing prevalence of cell phone-only households (Blumberg, Luke, and Cynamon 2006). Thus, a re-examination of undercoverage bias was warranted. This was done as part of a large-scale bias study conducted in conjunction with NHES:2007. The results of this study are summarized in Section 4.4.

Another potential source of undercoverage bias in telephone surveys that use the list-assisted method is the fact that not all telephone households are included in the sampling frame. Households in the zero-listed stratum have no chance of being included in the sample. Empirical findings were presented in Brick et al. (1995) to address the question of coverage bias associated with excluding the zero-listed stratum. Tucker, Lepkowski, and Piekarski (2002) found that about 3 percent of telephone households are in the zero-listed stratum. The results also indicate that households in the zero-listed stratum are not very different from households in the listed stratum. Because the proportion of telephone households that are in the zero-listed stratum is small and the persons living in these households are not very different from those living in households in the listed stratum, the bias resulting from excluding the zero-listed stratum is generally very small.

In NHES:2007, a two-phase stratification was used to select telephone numbers in order to produce more reliable national estimates from the extended interviews for subdomains defined by race and ethnicity. In the first phase, a sample of 476,167 telephone numbers was drawn, with telephone numbers in areas with high percentages of Black or Hispanic residents sampled at higher rates than those in areas with low percentages of Black or Hispanic residents. The sampling frame used in the study contained the Census 2000 counts of persons in the area by race and ethnicity. A 100-bank was classified in the high minority concentration stratum if its population was either at least 20 percent Black or at least 20 percent Hispanic. The banks that did not meet this requirement were classified in the low minority concentration stratum. The sampling rate in the high minority concentration stratum was nearly twice that of the low minority stratum.

In the phase 2 sample, within each minority stratum, the sampled telephone numbers were classified as mailable or nonmailable according to whether they could be matched to a mailing address in the white pages telephone directory or from other databases. Mailable status was used because it has been

found to improve the efficiency of the sample by facilitating the oversampling of mailable numbers (which are more likely to be residential). Within each of the four strata defined by the combinations of minority concentration and mailable status, telephone numbers were subsampled at different rates. In the low minority stratum, telephone numbers in the mailable substratum were sampled at a rate about 42 percent higher than numbers in the nonmailable substratum; in the high minority stratum, telephone numbers in the mailable substratum were sampled at a rate about 38 percent higher than numbers in the nonmailable substratum. Within each of the four strata defined by the combinations of minority concentration and mailable status, telephone numbers were subsampled at different rates in order to attain the final phase 2 allocation.

In this manner, a phase 2 sample of 251,826 numbers was selected for NHES:2007,[9] and a reserve sample of 26,664 telephone numbers was also selected. Assuming that 45 percent of the sampled telephone numbers would belong to households and assuming a Screener unit response rate of 64 percent, it was expected that about 62,000 screening interviews would be completed. For example, in table 3-1, 29,192 Screeners were expected to be completed in stratum 1 (mailable, high minority). The number of completed Screeners in stratum 1 was calculated in the following manner: First, the final NHES:2007 phase 2 allocation to stratum 1 (74,480 telephone numbers) was multiplied by the expected residency rate (73 percent) to get the approximate number of residential telephone numbers (54,370). For the 60 percent of those residential numbers that were randomly designated to receive the standard protocol (see section 4.1 for details on the standard protocol), a 63 percent expected response rate was used to estimate the expected number of completed Screeners; for the remaining 40 percent, a 39 percent initial cooperation rate was used to estimate the expected number of completed Screeners.[10] These calculations result in a total of 29,192 expected completed Screeners[11] for stratum 1. However, after the release of the initial sample of 251,826 telephone numbers, it was determined that the residency and response rates were lower than expected. Thus, the entire reserve sample of 26,664 telephone numbers was released. The total number of telephone numbers released for the study was 278,490, including the 26,664 reserve telephone numbers. The Screener unit response rate was 53 percent, and the number of households with completed screening interviews was 54,034 (table 3-2).

---

[9] The sample of 251,826 was selected using different rates for four strata. These strata were defined using exchange level classification of minority status and the telephone number level of mailable status, as follows: mailable high minority, mailable low minority, non-mailable high minority, and non-mailable low minority. Subsampling rates for each stratum were determined by the target sample sizes. All mailable telephone numbers were retained in the subsample. Non-mailable telephone numbers were subsampled at rates of approximately 45 percent for high minority and 51 percent for low minority.

[10] See sections 4.1.1 and 4.1.5 for details on the subsampling of cases for nonresponse followup.

[11] The rates given in table 3-1 and the associated text have been rounded to whole numbers for presentation purposes. However, more significant digits were used in the actual calculations.

Table 3-1.   Expected number of completed screeners, by sampling stratum: 2007

Stratum	Final NHES:2007 phase 2 allocation	Expected residency rate (percent)	Expected Screener response rate (percent)	Expected initial cooperation rate (percent)	Expected number of completed Screeners
Total	251,826	†	†	†	62,000
1 (Mailable, High minority)	74,480	73	63	39	29,192
2 (Mailable, Low minority)	63,203	76	69	42	28,078
3 (Not mailable, High minority)	60,309	11	44	36	2,703
4 (Not mailable, Low minority)	53,834	9	46	36	2,026

† Not applicable.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Surveys Program (NHES), 2007.

3.3   Sampling Within Households

To limit burden on respondents, a within-household sampling scheme was developed to control the number of persons sampled for extended interviews in each household. As part of this sampling scheme, in all households with children and youth ages 20 or younger, children/youth were enumerated. Once the enumeration of the appropriate household members was completed in the Screener, the sampling of household members for the extended interviews was done by computer. The SR and/or PFI interviews were conducted with parents/guardians of sampled children ages 3 to 6 and not yet in kindergarten or above for SR, and children/youth in kindergarten through 12th grade with a maximum age of 20 for PFI. Up to one eligible child was selected for the SR survey and up to one eligible child was selected for the PFI survey. If there were two or more eligible children in the SR domain or the PFI domain, one was selected with equal probability. This sampling algorithm was designed to limit the amount of time required to conduct interviews with parents in households with a large number of eligible children. Table 3-2 gives the expected and actual overall unit response rates and numbers of completed interviews for the SR and PFI NHES:2007 surveys.

Estimates from the October 2003 Current Population Survey (CPS) indicated that about 32 percent of all households had at least one eligible child. Using the within-household sampling algorithm developed for NHES:2007, 62,000 screened households should have yielded a sample size of 16,845 children for PFI and 4,512 children for SR. Assuming an SR and a PFI interview unit response rate of 84 percent,[12] the expected number of completed PFI interviews was 14,150 and the expected number of completed SR interviews was 3,790 . During data collection, lower than expected response rates for both the Screener and extended interviews prompted a revision of the estimates of the number of completed interviews to 56,389 for the Screener, 10,822 for the PFI interview, and 3,365 for the SR interview. The actual number of completed PFI interviews was 10,681 and the actual number of completed SR interviews was 2,633 . The difference between the expected and actual numbers of completed interviews was mainly due to the completion of fewer Screeners than expected.[13]

---

[12]Expected response rates were based on actual experience in NHES:2005.

[13] The actual unit response rate for the SR survey was 77 percent and the actual unit response rate for the PFI survey was 74 percent, compared to the expected rate of 84 percent for both surveys.

Table 3-2. Expected and actual numbers of completed interviews and weighted overall unit response rates for the NHES:2007 Screener and extended interviews

	Expected			Actual	
Interview	Original expected number of completed interviews	Revised number of completed interviews	Original expected overall unit response rate (percent)	Number of completed interviews	Overall unit response rate (percent)
Screener	62,000	56,389	64	54,034	53
PFI survey	14,150	10,822	54	10,681	39
SR survey	3,790	3,365	54	2,633	41

NOTE: SR and PFI overall unit response rates are the products of the Screener unit response rate and the extended interview (SR or PFI) unit response rate.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Readiness (SR) Survey and Parent and Family Involvement in Education Survey of the National Household Education Surveys Program (NHES), 2007.

Although the sample yield for children/youth was lower than expected, the lower yield did not affect the ability to detect differences between key estimates from 2007 and key estimates from previous NHES surveys beyond the thresholds that were used to design the sample. To reiterate, the key statistics for SR were participation in center-based care and education arrangements, literacy skills, and numeracy skills, by single year of age and by race/ethnicity (White, non-Hispanic; Black, non-Hispanic; and Hispanic), and for PFI were parental participation in school and home learning activities and parental satisfaction with school practices, by race/ethnicity, by 2-year grade group, by parents' educational attainment, by school type, and by school size.

3.4 Weighting Procedures

The objective of NHES:2007 is to make inferences about the entire civilian, noninstitutionalized population for the domains of interest. Although only telephone households were sampled, the estimates were adjusted to totals of persons living in both telephone and nontelephone households derived from the October 2005 and March 2006 CPS files to achieve this goal. The March 2006 CPS weights were adjusted to population totals based on the 2000 Decennial Census. Any additional undercoverage in the census of special populations, such as the homeless, remains in the totals obtained from the CPS. The weighting procedures are described briefly below.

3.4.1 Household-Level Weights

The primary purpose of the Screener in NHES:2007 was to provide information required to assess the eligibility of household members for an extended interview. Household-level information that is of analytic interest was also collected during the extended interview. Since no data intended for analyses were collected at the household level only, household-level weights were calculated solely for use as a basis for computing person-level weights for the analysis of the extended interview data. In computing household weights, a household base weight was developed to account for the RDD sampling of telephone numbers, including the sampling rate differences by minority concentration stratum and mailable substratum. This weight was adjusted for subsampling of cases for nonresponse follow-up, for

Screener nonresponse, and for households that had more than one telephone number, hence more than one chance of being included in the sample. A CHAID analysis was run to identify characteristics most associated with Screener nonresponse.[14]   These characteristics, which were primarily geographic characteristics associated with the telephone exchange, were used to form the cells for nonresponse adjustment of the household weights. The final adjustment was a poststratification adjustment to the household weights.   The primary purpose of the poststratification adjustment was to account for undercoverage resulting from the sampling of telephone households only. Poststratification ensures that survey weights sum to known population totals. The characteristics used in poststratification were census region (Northeast/South/Midwest/West) and presence of children less than 18 years of age. Table 3-3 presents the control totals used for poststratifying the household-level weights. The variables used in poststratification were chosen to address differences in coverage rates with respect to region in which the household is located and presence of children in the household.

Table 3-3.   Control totals for poststratifying the NHES:2007 household-level weights

Census region[1]	Children under 18 in household	Control total[2]
Total	—	114,510,050
Northeast	No	13,993,709
Northeast	Yes	7,137,051
South	No	27,173,229
South	Yes	14,638,867
Midwest	No	17,390,279
Midwest	Yes	8,981,331
West	No	15,731,203
West	Yes	9,464,380

[1]The following states and the District of Columbia are in each census region: Northeast: CT, MA, ME, NH, NJ, NY, PA, RI, VT; South: AL, AR, DC, DE, FL, GA, KY, LA, MD, MS, NC, OK, SC, TN, TX, VA, WV; Midwest: IA, IL, IN, KS, MI, MN, MO, ND, NE, OH, SD, WI; West: AK, AZ, CA, CO, HI, ID, MT, NV, NM, OR, UT, WA, WY.
[2]The control totals are numbers of households.
SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, March 2006.

3.4.2   Person-Level Weights

The next weighting procedures resulted in person-level weights (i.e., weights used to estimate the number of persons and to produce estimates of characteristics of persons).   The household-level weight was used as the base weight, and the weighting procedures included the adjustment of the estimates to independent totals from the CPS.

---

[14] Chi-Square Automatic Interaction Detection (CHAID) is a categorical search algorithm that identifies characteristics associated with response propensity.

Person Weights for the SR and PFI Interviews

As described in section 3.3, a sampling algorithm was used to limit the number of persons sampled in each household while maintaining the sampling rates required to attain the target sample sizes. The sampling was based on information collected in the Screener interview from the adult household member who responded to the Screener, and the eligibility of the sampled children was later verified or updated when the parent/guardian most knowledgeable about the child/youth responded to the SR or PFI interview. Because sampling eligibility was determined based on the data collected in the Screener, the weighting procedures were developed with possible misclassification of children according to grade—resulting in a change in the survey administered—taken into account so that the estimates would not incur bias due to misclassification.

The first step in developing the person weights for the SR and PFI surveys was to account for the probability of sampling the child/youth in the given household, which is determined by the number of children in each domain (SR or PFI) in the household. The application of this adjustment to the household weight created a person-level base weight for the SR and PFI interviews.

The next step involved adjusting the person-level base weight for nonresponse to the SR or PFI interview. In order to account for slight differences in response propensities, nonresponse adjustment cells were created using a combination of age and individual year of grade. Categories included unenrolled children ages 3 years to 6 years, children ages 3 years to 6 years who were enrolled in preschool, and each single grade for children enrolled in grades kindergarten through 12. Enrolled children with no grade equivalent were included in the cell containing the modal grade for their age; that is, they were assigned to the grade in which most children their age are enrolled. For each cell, the ratio of the weighted number of eligible sampled children to the weighted number of children for whom parents completed an interview was then computed. This ratio was multiplied by the person-level base weight to create the nonresponse-adjusted person-level SR or PFI interview weight.

The final stage of weighting for the SR and PFI interviews was a raking adjustment. Raking was proposed by Deming and Stephan (1940) as a way to ensure consistency between complete counts and sample data from the 1940 U.S. Census. The raking procedure typically improves the reliability of survey estimates and also corrects for the bias due to households or persons not covered by the survey (e.g., households without telephones and households with unlisted telephone numbers belonging to zero-listed telephone banks). The raking procedure is carried out in a sequence of adjustments. First, the weights are adjusted to sum to the totals on one marginal distribution (or dimension) and then the adjusted weights are further adjusted to sum to the totals on the second marginal distribution, and so on. One sequence of adjustments to the marginal distributions is known as a cycle or iteration. The procedure is repeated until convergence of weighted totals is achieved.

The raking procedure for the SR and PFI weights involved raking the nonresponse-adjusted person-level weights to national totals obtained using the percentage distributions from the October 2005 CPS and the total number of children from the March 2006 CPS. The October 2005 CPS contains variables not available on the March 2006 CPS, but the totals in the latter are more current and more reliable due to the augmented sample used for the March CPS. In the procedure used in NHES:2007, the control total for a raking cell is the proportion in that cell from the October 2005 CPS multiplied by the estimate of the total number of children from the March 2006 CPS. The three raking dimensions used for the SR and PFI interview weights were a cross between race/ethnicity of the child (Black, non-Hispanic/Hispanic/other) and household income categories ($10,000 or less/$10,001–$25,000/$25,001 or

more), a cross of census region (Northeast/South/Midwest/West) and urbanicity (urban/rural), and a cross of home tenure (rent/own or other) and enrollment status or grade of child (with those enrolled in school but having no grade equivalent assigned to the modal grade for their age). These raking dimensions were used because they include important analysis variables (e.g., grade) and characteristics that have been shown to be associated with telephone coverage (e.g., race/ethnicity). Tables 3-4 and 3-5 show the control totals used for raking the SR and PFI interview weights, respectively.

Table 3-4.   Control totals for raking the SR-NHES:2007 person-level interview weights

Total		8,734,486
Race/ethnicity of child	Household income	Control total[1]
Black, non-Hispanic	$10,000 or less	325,617
Black, non-Hispanic	$10,001-$25,000	320,691
Black, non-Hispanic	$25,001 or more	672,876
Hispanic	$10,000 or less	224,156
Hispanic	$10,001-$25,000	603,418
Hispanic	$25,001 or more	1,091,048
Other	$10,000 or less	278,601
Other	$10,001-$25,000	579,151
Other	$25,001 or more	4,638,928
Census region[2]	Urbanicity	Control total[1]
Northeast	Urban	1,196,428
Northeast	Rural	221,381
South	Urban	2,301,791
South	Rural	858,488
Midwest	Urban	1,539,271
Midwest	Rural	521,193
West	Urban	1,857,958
West	Rural	237,976
Home tenure	Age/grade of child	Control total[1]
Rent	Unenrolled	1,714,544
Rent	Preschool	1,350,853
Own or other	Unenrolled	2,429,177
Own or other	Preschool	3,239,912

[1] The control totals are numbers of persons.

[2] The following states and the District of Columbia are in each census region: Northeast: CT, MA, ME, NH, NJ, NY, PA, RI, VT; South: AL, AR, DC, DE, FL, GA, KY, LA, MD, MS, NC, OK, SC, TN, TX, VA, WV; Midwest: IA, IL, IN, KS, MI, MN, MO, ND, NE, OH, SD, WI; West: AK, AZ, CA, CO, HI, ID, MT, NV, NM, OR, UT, WA, WY.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, March 2006; October 2005.

Table 3-5. Control totals for raking the PFI-NHES:2007 person-level interview weights

Total		53,185,978

Race/ethnicity of child	Household income	Control total[1]
Black, non-Hispanic	$10,000 or less	1,672,661
Black, non-Hispanic	$10,001-$25,000	1,998,302
Black, non-Hispanic	$25,001 or more	4,226,716
Hispanic	$10,000 or less	952,408
Hispanic	$10,001-$25,000	2,735,705
Hispanic	$25,001 or more	6,240,396
Other	$10,000 or less	1,416,805
Other	$10,001-$25,000	3,411,705
Other	$25,001 or more	30,531,280

Census region[2]	Urbanicity	Control total[1]
Northeast	Urban	8,046,784
Northeast	Rural	1,488,933
South	Urban	13,986,891
South	Rural	5,216,625
Midwest	Urban	8,749,595
Midwest	Rural	2,962,588
West	Urban	11,288,657
West	Rural	1,445,905

Home tenure	Grade of child	Control total[1]
Rent	Transitional kindergarten/ kindergarten/pre-1st grade	1,186,672
Rent	1st grade	1,328,422
Rent	2nd grade	1,166,716
Rent	3rd grade	1,216,608
Rent	4th grade	1,082,817
Rent	5th grade	1,105,484
Rent	6th grade	1,045,527
Rent	7th grade	1,045,227
Rent	8th grade	1,124,203
Rent	9th grade	1,113,763
Rent	10th grade	1,081,231
Rent	11th grade	918,851
Rent	12th grade	757,865
Own or other	Transitional kindergarten/ kindergarten/pre-1st grade	2,715,226
Own or other	1st grade	2,806,353
Own or other	2nd grade	2,750,847
Own or other	3rd grade	2,698,323
Own or other	4th grade	2,767,402
Own or other	5th grade	2,941,790
Own or other	6th grade	3,007,403
Own or other	7th grade	3,097,426
Own or other	8th grade	3,115,756
Own or other	9th grade	3,157,928

See notes at end of table.

Table 3-5. **Control totals for raking the PFI-NHES:2007 person-level interview weights—Continued**

Home tenure—continued	Grade of child—continued	Control total[1]
Own or other	10th grade	3,283,568
Own or other	11th grade	3,496,583
Own or other	12th grade	3,173,987

[1] The control totals are numbers of persons.

[2] The following states and the District of Columbia are in each census region: Northeast: CT, MA, ME, NH, NJ, NY, PA, RI, VT; South: AL, AR, DC, DE, FL, GA, KY, LA, MD, MS, NC, OK, SC, TN, TX, VA, WV; Midwest: IA, IL, IN, KS, MI, MN, MO, ND, NE, OH, SD, WI; West: AK, AZ, CA, CO, HI, ID, MT, NV, NM, OR, UT, WA, WY.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, March 2006; October 2005.

3.5   Computing Sampling Errors

Sampling error, the difference between the estimate from a sample and the true population parameter, results when data are collected on a sample rather than the full population. In surveys with complex sample designs, such as NHES:2007, direct estimates of the sampling errors assuming a simple random sample will typically underestimate the variability in the estimates. The NHES:2007 sample design and estimation included procedures that deviate from the assumption of simple random sampling, such as oversampling in areas with higher concentrations of minorities, sampling persons within households with differential probabilities, and raking to control totals.

One method for computing sampling errors to reflect these aspects of the sample design and estimation is the replication method. Replication involves splitting the entire sample into a set of groups or replicates based on the actual sample design of the survey. The survey estimates can then be computed for each of the replicates by creating replicate weights that mimic the actual sample design and estimation procedures used in the full sample. The variation in the estimates computed from the replicate weights can then be used to estimate the sampling errors of the estimates from the full sample. Appendix B contains a summary of weighting and sample variance estimation variables for NHES:1991–2007.

A total of 80 replicates were defined for NHES:2007 based on the sampling of telephone numbers. This number was chosen to provide reliable estimates of sampling errors with reasonable data processing costs. The specific replication procedure used for NHES:2007 was a jackknife replication method (Wolter 1985). It involved dividing the sample into 80 random subsamples (replicates) for the computation of the replicate weights. Replicate weights were created for each of the 80 replicates using the same estimation procedures that were used for the full sample. These replicate weights are included in the SR file as FRWT1 through FRWT80, and in the PFI file as FPWT1 through FPWT80. The computation of the sampling errors using these replicate weights can be done easily using the Windows-based software packages WesVar Complex Samples Software, SUDAAN (Shah et al. 1995), Stata, or AM Statistical Software; in WesVar or SUDAAN, the replication method should be specified as JK1. The current version of WesVar Complex Samples (version 5) is available from Westat. Information can be obtained at http://www.westat.com/wesvar. A previous version of WesVar (version 4) is available free of charge at that Web site or by sending an e-mail message to wesvar@westat.com. Information on obtaining SUDAAN can be found at http://www.rti.org/sudaan, and the AM software is available at http://am.air.org. Information on Stata can be found at http://www.stata.com.

Another approach to the valid estimation of sampling errors for complex sample designs is to use a Taylor-series approximation to compute sampling errors. To produce standard errors using a Taylor-series program, such as SUDAAN or the survey data analysis procedures (PROC SURVEYMEANS and PROC SURVEYREG) in SAS version 9, two variables are required to identify the stratum and the primary sampling unit (PSU). The stratum-level variable is the indicator of the variance estimation stratum from which the unit (telephone number or sampled person) was selected. The PSU is an arbitrary numeric identification number for the unit within the stratum. The PSU and stratum variables appear on each of the extended interview files and are called RPSU and RSTRATUM on the SR file, and PPSU and PSTRATUM on the PFI file. These variables can be used in SUDAAN to produce standard errors by specifying that the design is a "with replacement" sample (DESIGN = WR) and that the sampling levels are given by the appropriate stratum and PSU variables. For example, for estimates from the PFI interview file, use PSTRATUM PPSU in the NEST statement. In the SAS survey procedures, the stratum and PSU variables are specified in the STRATA and CLUSTER statements, respectively. (Information on obtaining SAS version 9 can be found at http://www.sas.com.)

Stata, another software package that uses Taylor-series methods, also uses the PSU and stratum variables to define the units needed for computation. (Information on obtaining Stata is available at http://www.stata.com.) To specify the stratum, PSU, and weight variables in Stata use the svyset strata, svyset psu, and svyset pweight commands. For example, for estimates from the PFI interview file, use the following commands to specify these design parameters:

> svyset strata pstratum
> svyset psu ppsu
> svyset pweight fpwt

The full sample weight to be used for analysis of the SR interview file is FRWT and for the PFI interview file the full sample weight is FPWT.

Data users should be aware that the use of different approaches or software packages in the calculation of standard errors may result in slightly different standard errors. Estimates of standard errors computed using the replication method and the Taylor-series method are nearly always very similar, but not identical. For a discussion of this issue see Broene and Rust (2000).

3.6 Approximate Sampling Errors

Although calculating the sampling errors using the methods described above is recommended for many applications, simple approximations of the sampling errors may be valuable for some purposes. One such approximation is discussed next.

Most statistical software packages compute standard errors of the estimates based upon simple random sampling assumptions. The standard error from this type of statistical software can be adjusted for the complexity of the sample design to approximate the standard error of the estimate under the actual sample design used in the survey. For example, the variance of an estimated proportion in a simple random sample is the estimated proportion (p) times its complement (l-p) divided by the sample size (n). The standard error is the square root of this quantity. This estimate can be adjusted to more closely approximate the standard error for the estimates from NHES:2007.

A simple approximation of the impact of the sample design on the standard errors of the estimates that has proved useful in previous NHES surveys and in many other surveys is to adjust the simple random sample standard error estimate by the root design effect (DEFT). The DEFT is the ratio of the standard error of the estimate computed using the replication method discussed above to the standard error of the estimate under the assumptions of simple random sampling. An average DEFT is computed by estimating the DEFT for a number of estimates and then averaging. A standard error for an estimate can then be approximated by multiplying the simple random sample standard error estimate by the mean DEFT.

In complex sample designs, like NHES:2007, the DEFT is typically greater than 1 due to the clustering of the sample and the differential weights attached to the observations. In NHES:2007, both of these factors contributed to making the average DEFT greater than 1. (See appendix B for the recommended DEFT for each data file of NHES:1991–2007.)

The average DEFT computed for estimates in the SR and PFI surveys ranged from 1.4 to 1.6. For the SR file estimates, the average DEFT was 1.4 overall. For estimates by race/ethnicity, the average DEFT was 1.6 for non-Hispanic Blacks and 1.4 for the other race/ethnicity categories. For estimates by interview path, the average DEFT was 1.4 for unenrolled children and 1.5 for children enrolled in preschool (ALLGRADE = N). Therefore, a DEFT of **1.4** is recommended to approximate the standard error of overall estimates in the SR interview file. For estimates by race/ethnicity or by interview path, a DEFT of **1.4** is recommended, with the exception of estimates of non-Hispanic Blacks (**1.6**) and children enrolled in preschool (**1.5**).

For the PFI file estimates, the average DEFT was 1.4 overall. For estimates by interview path, the average DEFT was 1.5. For estimates by race/ethnicity, the average DEFT was 1.5 for non-Hispanic Blacks and 1.4 for the other race/ethnicity categories. Therefore, a DEFT of **1.4** is recommended to approximate the standard error of overall estimates in the PFI interview file. For estimates by interview path, a DEFT of **1.5** is recommended; and for estimates by race/ethnicity, a DEFT of **1.4** is recommended, with the exception of non-Hispanic Blacks (**1.5**).

As stated earlier, the average DEFT can be used to approximate the standard error for an estimate. An example of how to do this for a **percentage** estimate derived using a statistical package like SAS[15] or SPSS is as follows. If a weighted estimate of 23 percent is obtained for some characteristic in the PFI file (suppose that 23 percent of children visited a museum in the past month), then an approximate standard error can be developed in a few steps. First, obtain the simple random sample standard error for the estimate using the weighted estimate in the numerator and the unweighted sample size in the denominator: the standard error for this 23 percent statistic would be 0.41 percent (the square root of (23 x 77)/10,681, where the weighted estimate (p) is 23 percent, 77 is 100 minus the estimated percent (1-p), and the unweighted sample size (n) is 10,681). The approximate standard error of the estimate from NHES:2007 is this quantity (the simple random sample standard error) multiplied by the DEFT for the PFI file estimates of 1.4. In this example, the estimated standard error would be 0.57 percent (1.4 x 0.41 percent).

The approximate standard error for a **mean** can be developed using a related procedure. The three steps required to do so are demonstrated using an example from the PFI file. First, the mean is estimated using the full sample weight and a standard statistical package like SAS or SPSS. Second, the simple random sample standard error is obtained through a similar, but unweighted, analysis. Third, the

---

[15] Here, the reference to "SAS" applies to SAS version 6.12 or earlier versions, or the non-survey procedures in SAS versions 8 or 9.

standard error from the unweighted analysis is multiplied by the mean DEFT for the PFI file estimates of 1.4 to approximate the standard error of the estimate under the NHES:2007 design. For example, suppose the average number of times in this school year the parents/adult household members of children enrolled in grades kindergarten through 12 in regular school have gone to meetings or participated in activities at the child's school is 8.6 and the simple random sampling standard error (unweighted) is 0.11. Then, the approximate standard error for the estimate would be 1.4 x 0.11 = 0.15.

Users who wish to adjust the standard errors for estimates of **parameters in regression models** should follow a procedure similar to that discussed for means, above. Specifically, the estimates of the parameter in the model can be estimated using a weighted analysis in a standard statistical software package such as SAS[12] or SPSS. A similar, but unweighted, analysis will provide the simple random sample standard errors for these parameter estimates. The standard errors can then be multiplied by the DEFT to arrive at the adjusted standard error for the NHES:2007 design. For example, if a given parameter in a model involving items from the SR file has a weighted estimate of 2.33 and an unweighted simple random sample standard error of 0.45, then the adjusted standard error would be 1.4 x 0.45 = 0.63.

Alternatively, the final weight can be adjusted to reflect the DEFT before the parameter estimates are calculated in a standard statistical software package such as SAS or SPSS. To do this, first sum the values of the final weights for the sample of interest. For instance, for an analysis of all children enrolled in grades kindergarten through 12, sum the final weights for all 10,681 cases on the PFI file. Second, divide this sum by the number of cases to generate an average final weight. (In the above example, the number of cases is 10,681). Third, multiply the average final weight by the square of the DEFT for the population of interest. (In the above example, the average final weight would be multiplied by the square of 1.4, or 1.96.) Fourth, divide the final weight by the adjusted average weight and save the quotient as a new final weight. (In the above example, the new final weight is equal to the final weight divided by the product of 1.96 and the average final weight.) Finally, weight the analysis by this new final weight. The standard errors generated in the analysis will approximate the standard errors correctly adjusted for design effects.

It should be noted that direct computation of the standard errors is always recommended when the statistical significance of statements of difference would be affected by small differences in the estimated standard errors.

3.7   Imputation

In NHES:2007, as in most surveys, the responses to some data items are not obtained for all interviews. There are numerous reasons for item nonresponse. Some respondents do not know the answer to a question or do not wish to respond for other reasons. Some item nonresponse arises when an interview is interrupted and not continued later, leaving items at the end of the interview blank. Item nonresponse may also be encountered because responses provided by the respondent are not internally consistent, and this inconsistency is not discovered until after the interview is completed. In these cases, the items that were not internally consistent were set to missing.

For most of the data items collected in NHES:2007, the item response rate was very high. The median item response rate for items from the SR interview was 99.31 percent; and for the PFI interview it was 99.04 percent. There were 18 items in the SR file with item response rates of less than 90 percent, and 5 items in the PFI file. (Item response rates are discussed in more detail in chapter 4.) Despite the

high item response rate for cases for which an interview was completed, data items with missing data on the file were imputed. (In general, character string variables, such as countries of origin, languages, or "other/specify" responses were not imputed. School characteristics merged to the PFI data file from the Common Core of Data (CCD) and Private School Survey (PSS) files also were not imputed.) The imputations were done for two reasons. First, complete responses were needed for the variables used in developing the sampling weights. Second, users will be computing estimates employing a variety of methods and having complete data should aid their analysis.

A hot-deck procedure was used to impute missing responses (Kalton and Kasprzyk 1986). In this approach, the entire file was sorted into cells defined by characteristics of households or respondents that are likely to be associated with differences in response propensities. The variables used in the sorting also included any variables involved in the skip pattern for the item. Many of these sort order variables are not on the data files.

The standard set of sort order variables for the household-level items collected in the SR and PFI surveys consisted of:

- CENREG–the census region in which the household was located;

- HINCOME or HINCMRNG–household income category (specific or broad, respectively);

- KIDINHH–a variable derived specifically for imputation from the age of household members, indicating whether or not children under age 18 resided in the household. This variable was derived from Screener AGE; and

- HOWNHOME–whether the home was rented versus owned or other arrangement.

The standard sort order variables for the person-level items on the SR and PFI interview files were:

- ALLGRADR–a variable derived specifically for imputation that indicates the grade/grade equivalent of the sampled child;

- SEX–sex of the sampled child;

- PARGRADS–a variable derived specifically for imputation that indicates the highest education level attained by either parent in the household as less than high school credential, high school credential but no bachelor's degree, or college graduate. This variable was derived from MOMGRADE1, MOMGRADE2, MOMDIPL1, MOMDIPL2, DADGRADE1, DADGRADE2, DADDIPL1, and DADDIPL2; and

- HHPARNS–a variable derived specifically for imputation from HHMOM1, HHMOM2, HHDAD1, and HHDAD2 indicating whether there were two parents in the household or not.

All of the observations were sorted into cells defined by the responses to the sort variables, and then divided into two classes within the cell depending on whether or not the item was missing. For an observation with a missing value, a value from a randomly selected donor (i.e., an observation in the same cell but with the item completed) was used to replace the missing value. After the imputation was completed, edit programs were run to ensure the imputed responses did not violate skip patterns or edit rules. If any violations occurred, the program was adjusted and imputation was rerun, or if only a few cases were affected, they were manually imputed.

For some items, the missing values were imputed manually rather than using the hot-deck procedure. In NHES:2007, manual imputation was done (1) to impute certain person-level demographic characteristics; (2) to impute whether a child is homeschooled, attends regular school for some classes, and the number of hours attending regular school; (3) to impute variables that involved complex relationships that would have required extensive programming to impute using a hot-deck procedure; (4) to correct for a small number of inconsistent imputed values; and (5) to impute for a few cases when no donors with matching sort variable values could be found.

For manual imputation of the person-level demographic items and of the homeschooling items, the following three sort variables were used to ensure that all household members were grouped together: state, the three-digit ZIP Code (i.e., the first three digits of the ZIP code associated with the telephone exchange), and the person identification number.

After values had been imputed for all observations with missing values, the distribution of the item prior to imputation (i.e., the respondents' distribution) was compared to the post-imputation distributions of the imputed values alone and of the imputed values together with the observed values. The comparisons revealed similar item distributions pre- and post-imputation. This comparison is an important step in assessing the potential impact of item nonresponse bias and ensuring that the imputation procedure reduces this bias, particularly for items with relatively low item response rates (less than 90 percent).

For each data item for which any values were imputed, an imputation flag variable was created. If the response for the item was not imputed, the imputation flag was set equal to 0. If the response was imputed, the flag was set to a value of 1, 2, 3, or 4. The value of the imputation flag indicates the specific procedure used to impute the missing value. The assignment of these values is described below.

The procedure for hot-deck imputation only recognizes missing value codes as those that need to be replaced by imputed values. For NHES:2007, these missing codes were -7 = refused and -8 = don't know. Therefore, in some cases, variables that originally had values of -1 (inapplicable) had to be recoded to a missing value code (i.e., -9 = not ascertained) for some cases prior to being imputed using the standard hot-deck approach. This was done so that data were consistent with the skip patterns of the questionnaire. For these cases the imputation flag was set to 2. For example, in the PFI file, if the value of SCHOICE (PG2) equaled -8 for a child, then SDISRCT (PG3) was never asked and thus equaled -1 (inapplicable).   During the imputation process for this child, if SCHOICE was imputed to equal 2 (chosen), SDISRCT had to first be recoded from -1 (inapplicable) to -9 (not ascertained) before the imputation procedure would recognize SDISRCT as a variable that should be imputed to equal either 1 (school is in assigned school district) or 2 (school is not in assigned school district). In this case, the imputation flag for SDISRCT would be set to 2. If an item was imputed manually for any of the reasons described above, the flag was set to 3. The imputation flag was set to 4 if the reported value was "don't know" prior to imputation using the standard hot-deck approach. In all other cases in which an item was imputed, the imputation flag was set to 1.

The imputation flags were created to enable users to identify imputed values. Users can employ the imputation flag to delete the imputed values, use alternative imputation procedures, or account for the imputation in computation of the reliability of the estimates produced from the data set. For example, some users might wish to analyze the data with the missing values rather than the imputed values. If there is no imputation flag corresponding to the variable, no values for that variable were imputed. If the imputation flag corresponding to the variable is equal to 1, 2, 3, or 4, the user can replace the imputed

response with a missing value to accomplish this goal. This method could also be used to replace the imputed value with a value imputed by some user-defined imputation approach. Finally, if the user wishes to account for the fact that some of the data were imputed when computing sampling errors for the estimates, the missing values could be imputed using multiple imputation methods (Rubin 1987) or imputed so that the Rao and Shao (1992) variance procedures could be used.

This page is intentionally blank.

4. DATA COLLECTION METHODS AND RESPONSE RATES

4.1   Data Collection Procedures

The following sections discuss the procedures used in the data collection phase of the 2007 National Household Education Surveys Program (NHES:2007), including the use of computer-assisted telephone interviewing (CATI), staff training, interviewer assignments and contact procedures, and quality control.

4.1.1   Special Precollection Procedures

Before the beginning of data collection, special procedures were implemented to remove business and nonworking telephone numbers from the sample, and specific subsampling was done that reduced the number of telephone numbers from the full sample of 476,167 telephone numbers originally drawn to the final sample of 278,490 telephone numbers that was fielded. In addition, an advance mailing was conducted.

Identification of business and nonworking numbers. In NHES:2007, as in previous NHES administrations, procedures were used prior to data collection to reduce the number of unproductive calls. Prior to NHES:2001, Marketing Systems Group's (MSG's) Genesys ID process was used.[16] The Genesys ID process included tritone[17] checks for nonworking numbers and purging of listed business numbers (i.e., numbers listed in the yellow pages but not in the white pages). In NHES:2001 and NHES:2003, a more extensive procedure, the Genesys ID-PLUS process, was used prior to the field period. With the ID-PLUS utility, a telephone number was dialed by Genesys and allowed to ring up to two times (compared with one ring in the Genesys ID tritone test). If the telephone call was answered, a representative was available to speak to the respondent. In such cases, the representative attempted to ascertain whether the telephone number was a business number. In NHES:2005 and for NHES:2007, a more comprehensive prescreening procedure, the Genesys Comprehensive Sample Screening (Genesys-CSS) procedure, was used. Like the Genesys ID and ID-PLUS utilities, the Genesys-CSS utility also included the white and yellow pages matches. The primary differences between Genesys-CSS and the ID-PLUS procedure were enhanced identification of all types of wireless numbers and the predialing of numbers listed in the white pages.[18] With the Genesys-CSS utility, each telephone number was classified into one of the following categories:

LB (Listed Business)
UR (Unlisted Residence)
UB (Unlisted Business)
FM (Fax/Modem)
LA (Language Barrier)
NR (No Ring Back)

---

[16] MSG is the vendor that provides the sampling frame for the selection of telephone numbers. Genesys is the name of the system that generates the sampling frame.

[17] A tritone is the three-note sound heard when dialing a nonworking telephone number.

[18] With Genesys-CSS, all telephone numbers not identified as business numbers (including listed residential numbers) are dialed and allowed to ring up to two times, in order to identify business, cellular, and nonworking numbers. The dialing is done during the hours of 9 a.m. to 5 p.m. local time by specially trained agents. All calls are done in English.

NW (Nonworking)
BX (Blocked Exchanges)
PM (Privacy Manager[19])
WR (Wireless)
CP (Cell Phone)
DK (Undetermined: Residential/No Answer/Busy)

Telephone numbers identified by Genesys-CSS as LB, NW, WR, or CP, as well as UB telephone numbers for which no mailing address could be obtained were excluded from dialing. These exclusions amounted to 33 percent of the sample of telephone numbers. All telephone numbers that were not excluded from dialing as a result of the Genesys-CSS results were sent to up to two address vendors to obtain mailing addresses. A total of 40,382 of the 476,167 telephone numbers in the initial sample and 15,812 telephone numbers in the final sample of 278,490 were assigned a status of nonresidential as a result of the Genesys-CSS process and the process of matching to yellow pages and white pages listings.

Subsampling of telephone numbers. Two-phase stratification was used to select telephone numbers for the final NHES:2007 sample in order to produce more reliable national estimates. In the first phase, a sample of 476,167 telephone numbers was drawn, with telephone numbers in areas with high percentages of Black or Hispanic residents sampled at about twice the rate of those in areas with lower percentages[20] of Black or Hispanic residents. In the phase 2 sample, within each minority stratum, the sampled telephone numbers were stratified as mailable or nonmailable according to whether they could be matched to a mailing address. Within each of the four strata defined by the combinations of minority concentration and mailable status, telephone numbers were subsampled at different rates. Within each of the minority strata, telephone numbers in the mailable substratum were sampled at rates 42 percent higher than numbers in the nonmailable substratum. This process resulted in a sample of 278,490 telephone numbers for NHES:2007. Of this sample, 26,664 telephone numbers were set aside in a reserve sample to be fielded in case of lower-than-expected residency rates, response rates, or household members' eligibility rates.

Further subsampling of the original phase 2 sample was conducted for nonresponse followup. Prior to data collection, 60 percent of the original sample was designated for nonresponse followup,[21] including refusal conversion, a higher number of calls for noncontact cases, and a higher maximum call limit for telephone numbers at which contact with a household member had been made (this first 60 percent is called "wave 1"). The remaining 40 percent ("wave 2") of the original sample and the entire reserve sample were not subject to refusal conversion efforts and had a call limit of 14.

Advance mailing. The NHES:1996 field test showed that households receiving an advance letter were more likely to respond to the survey (Brick, Collins, and Chandler 1997). In an effort to increase Screener-level response, a mailing was planned for the households for which an address was obtained from either of two commercial firms. The advance letters explained the purpose of NHES:2007 and encouraged participation in the study. The letters were printed on National Center for Education Statistics (NCES) letterhead and were sent in U.S. Department of Education envelopes. Based on the findings of an experiment investigating the effectiveness of modest cash incentives (Brick et al. 2006), an incentive of $2 was included in each advance letter. In all, 152,261 telephone numbers were matched with

---

[19] Privacy Manager is a device that works with caller ID to screen and manage incoming calls.

[20] High minority areas were defined as having a population that was 20 percent Black or 20 percent Hispanic, based on figures from Census 2000.

[21] The reserve sample was selected first and set aside. Then a random 60 percent sample from the planned release was selected for nonresponse followup.

addresses; and all 152,261 telephone numbers with matched addresses were included in the final NHES:2007 sample. To coordinate the arrival of the letter with the initial call into the household, the mailing was conducted in two waves, one in late December 2006 and one in mid-January 2007, to correspond to the release of cases in the two waves of the sample. After the decision was made to release the reserve sample, cases with matched addresses were also sent an advance letter. Increased advance incentives of $5 and $10 were sent to random subsamples of the reserve sample in order to examine the effectiveness of those amounts in improving survey response.

Interactive Voice Response (IVR) Pre-Notification. It was not possible to send an advance letter to all cases in the RDD sample, because all telephone numbers could not be matched to addresses. In NHES:2007, an experiment was conducted to evaluate an Interactive Voice Response (IVR) system to deliver an advance announcement to a sample of telephone numbers, in order to ascertain its usefulness in notifying potential respondents of the survey and its effect on initial cooperation rates. This brief advance message introduced the study and its sponsorship, and informed respondents that they would be receiving a call from an interviewer. Those receiving the message had the option to press zero to speak with a staff member if they wished.

A total of 14,152 cases in wave 1 were designated to receive IVR call attempts. The Screener initial cooperation and refusal conversion rates for the cases in the IVR experiment were monitored on a weekly basis. With the experimental results indicating that IVR was having no effect on cooperation or refusal conversion, the decision was made to not use IVR pre-notification for the wave 2 sample. At the time of this decision (i.e., as of February 18, 2007), the initial cooperation rates for mailables and nonmailables in the wave 1 IVR treatment group were 39.3 percent and 26.8 percent, respectively, and these rates were 38.0 percent and 26.8 percent for mailables and nonmailables in the non-IVR treatment group. The refusal conversion rates at that time were 31.6 percent and 20.5 percent for mailables and nonmailables in the wave 1 IVR treatment group, and 32.0 percent and 23.9 percent for mailables and nonmailables in the wave 1 non-IVR treatment group.

4.1.2 CATI System Applications

The use of a CATI system for NHES:2007 included a number of applications that facilitated the implementation of the survey. Briefly, the most salient features of the CATI system for NHES:2007 were as follows:

- **Sampling:** The use of online sampling through CATI eliminated the need for separate screening and interviewing calls, reducing the cost and the burden on respondents.

- **Scheduling:** The CATI system was used to feed telephone numbers to the interviewers, maintain a schedule of callback appointments, and reschedule unsuccessful contact attempts to the appropriate day and time.

- **Skip patterns:** The CATI system was programmed to automatically guide interviewers through the complex skip patterns in the questionnaire, reducing the potential for interviewer error and shortening the questionnaire administration time. In addition, the CATI system was programmed so that questions about household members that were asked in one interview were not asked again in a subsequent interview. For example, if two sampled children had the same mother and father, the questions about the parents' characteristics were skipped in the second interview. Similarly, household characteristics questions were asked only in the first interview and not repeated in a second interview.

- **Monitoring survey progress:** The CATI system was programmed to provide automatic status reports throughout data collection. This allowed ongoing monitoring of the survey's progress.

- **Online help:** The CATI system was programmed to provide an online help screen for selected items in the Screener and extended interviews, which were chosen based upon the observed use of help screens in prior survey administrations. These screens, which could be accessed with a keystroke, clarified terminology, explained the intent of questions, and helped the interviewer obtain correct information.

4.1.3   Interviewer Training

Interviewer training for NHES:2007 was conducted in three phases. In the first phase, new interviewing staff completed training in general interviewing techniques and the use of the CATI system. Those who successfully completed this training and experienced interviewers who had already conducted CATI surveys at Westat were then assigned to NHES project training.

NHES:2007 training began with a self-administered online training program using prepared scripts to introduce the study and the survey instruments. An introduction, screener scripts, and extended interview scripts were completed in a sequential format. Trainees also completed required exercises and tests. A learning management system was used to guide trainees through the self-administered modules and monitor their progress.

Following completion of the self-administered sessions, trainees were assigned to trainer-led sessions. Interviewing staff who worked at one of the Telephone Research Centers (TRC) attended trainer-led sessions at a TRC. Home-based interviewing staff participated in trainer-led group sessions using voice and computer linkages. The content of the training sessions was the same in the two formats. Upon completion of the trainer-led sessions, all trainees conducted role play interview scripts that were monitored by training staff prior to being scheduled for live data collection.

To keep pace with attrition and to meet the needs of interviewing, groups were scheduled for training beginning in mid-December 2006 and continuing into April 2007. In total, 402 interviewers were trained for the study.

The survey staff included 30 interviewers bilingual in English and Spanish. These interviewers received the same training in English as did all other interviewers. They were then trained to conduct the interviews in Spanish. All of the CATI screens were translated into Spanish, and these screens were available to bilingual interviewers at a keystroke, so they could interview in either English or Spanish when placing a call into a household.

4.1.4   Interviewing Procedures

The CATI system scheduled cases automatically, based on an algorithm that was customized for NHES:2007. The system assigned cases to interviewers in the following order of priority:

- Cases that had specific appointments;

- Cases that had resulted in busy signals 15 minutes earlier;

- Cases that had resulted in noncontact at a scheduled appointment time;

- Cases that had unspecified appointment/general callback times for the time period;

- Cases that had not been contacted on previous attempts and had not been attempted during the time period; and

- Cases that had not yet been called (initial cases).

Initial attempts to contact households and determine the presence of household members eligible for extended interviews were conducted in two groups separated by a one-week hold period: a group of four calls consisting of two evening calls, one daytime call, and one weekend call; and a group of three calls, consisting of two evening calls and a weekend call on a different day than the previous weekend call.

If contact had not been made with either a household member or an answering machine after these two sets of calls, the case was sent to a vendor for 14 additional calls to be made by predictive dialing.[22] Westat sent a total of 31,907 noncontact cases to the vendor for predictive dialing followup. Of those, 1,896 cases were returned to Westat following contact with a household member, and 455 of those cases, or approximately 1 percent of all cases sent for predictive dialing, completed a screener. This is similar to the 2 percent of no answer cases that completed the Screener in NHES:2005.

If contact had not been made with a household member but an answering machine had been reached, the cycles of four calls and three calls were repeated. After this, the no answer-answering machine cases were randomly subsampled to receive a total of 14, 21, or 28 call attempts.

Once a household member was contacted, up to 20 calls attempts were made to complete the screener with a household member in wave 1 of the sample, and up to 14 call attempts were made for wave 2 cases and reserve sample cases, except in the case of language problem or refusal cases, described below. Once a household member was sampled as the subject of an SR or PFI interview, up to 20 additional call attempts were made to complete the interview with the identified parent/guardian respondent.

Procedures for non-English speakers. NHES:2007 was conducted primarily in English, but provisions were made to interview persons who spoke only Spanish. As was noted earlier, the questionnaires were translated into Spanish, the Spanish versions of the CATI instruments were

---

[22] Predictive dialing is a process in which telephone numbers are automatically dialed and are routed to an attendant or operator when a telephone number is answered. The attendant identifies him or herself as an interviewer for the subcontractor and asks if the telephone number is for residential or business use. Calls resulting in no contact are not routed to an attendant or operator; they are automatically handled and classified as noncontact by a computer system.

programmed, and bilingual interviewers were trained to complete the interview in either English or Spanish.

When the person answering the telephone was not able to speak English, and the interviewer was not bilingual and was not able to identify an English-speaking household member, the interviewer coded the case as a "language problem" and further specified the case as either "hearing/speech problem," "Spanish," or "language other than English or Spanish." There were 1,109 Screeners that were classified by at least one interviewer as a hearing or speech problem; 289 of these cases (26 percent) were completed. In some cases, very experienced interviewers accustomed to handling difficult interviews were able to complete these cases with the respondent who had been coded as a hearing/speech problem; in other cases, another household member was contacted on a subsequent call.

Bilingual interviewers were the only ones who could access cases coded Spanish or another language for followup. If a bilingual interviewer encountered a Spanish-speaking respondent on an initial call into a household, the interviewer could immediately begin to conduct the interview in Spanish without ever coding the case as a language problem. A total of 5,771 Screeners were classified as Spanish-speaking by the first interviewer who made contact. About 47 percent of these cases were finalized as completes (n = 2,701), and about 30 percent were finalized as refusals (n = 1,751). About 7.4 percent of the 1,237 Screeners with respondents identified by the initial interviewer as speaking some language other than English or Spanish were completed (number completed = 92). About 65 percent of the households identified as non-English/non-Spanish were finalized as language problems (n = 800) and 25 percent were finalized as refusals (n=307).

4.1.5 Special Data Collection Procedures

Refusal conversion. Additional efforts to gain cooperation from households or individual respondents who had initially refused to complete an interview were also part of the data collection effort for NHES:2007. As discussed in chapter 3, 60 percent of the original sample was designated for refusal conversion. Unless an interviewer indicated that the initial refusal was "hostile" (e.g., profane or abusive), a refusal conversion attempt was made for each Screener or extended interview refusal. Cases classified as "hostile" were reviewed by a supervisor to determine whether another attempt should be made. For most of the field period, a 13-day hold was placed on initial refusals before a conversion attempt was made. This period was decreased near the end of data collection to facilitate survey closeout while maximizing response rates.

In order to increase the likelihood of successful refusal conversion attempts at the Screener level, a letter was sent to first refusal households in wave 1 for which an address had been obtained. Like the advance letters, these were printed on ED stationery and provided information about the study. The letters were sent by first class mail with $2. A total of 35,334[23] such letters were mailed, and 9,121 of these cases (27 percent of eligible telephone numbers in this group) were completed, versus 580 of the 4,097 cases to which a letter was not mailed (16 percent of eligible telephone numbers in this group). About 17 percent of all completed Screeners were conducted in those households that received first refusal conversion letters.

An additional refusal conversion attempt was made for wave 1 cases that had twice refused to participate in the Screener interview. The wave 1 cases included in this effort were those for which

---

[23] Eleven cases from wave 2 were also sent a letter in error; a total of 35,345 cases were sent a refusal conversion letter.

neither the first nor second refusal received a code of "hostile." Letters were printed on ED stationery and mailed via FedEx, or by Priority Mail if the address was a P.O. box or rural route. In all, 20,859 Screeners cases were sent a letter by FedEx or Priority Mail. This effort resulted in the completion of 3,530 Screeners (18 percent of the eligible telephone numbers that were sent a letter by Fed Ex or Priority Mail). An additional 217 Screeners were completed with nonmailable second refusals (10 percent of eligible telephone numbers). All Screener refusals were considered to be final if a third refusal was received.

Refusal conversion efforts were successful at the extended interview level as well; 91 SR and 471 PFI interviews were completed as a result of initial refusal conversion attempts (that is, conversion following a first refusal). An additional refusal conversion attempt was also made for extended interview cases for which two refusals had been received. One hundred thirteen PFI interviews were completed out of 1,106 refielded second refusal cases, and 20 SR interviews out of 234 refielded second refusal cases were completed. The total numbers of completed extended interviews resulting from both initial and second refusal conversion efforts were 111 for SR and 584 for PFI.

In summary, the refusal conversion activities for NHES:2007 were productive. Including both first and second conversion efforts, approximately 38 percent of the Screener refusal cases that were mailed a letter and approximately 22 percent of the cases called after an initial and second refusal but not mailed a letter (because the phone number was not matched to an address) were completed. Of the extended interviews released for a first refusal conversion attempt, 17 percent of SR cases and 19 percent of PFI cases were completed. In the following stage, 9 percent of SR cases and 10 percent of PFI cases refielded for a second refusal conversion attempt were completed.

"Maximum call" cases. Other efforts to increase the Screener unit response rate focused on "maximum call" cases, in which a person had answered the telephone on at least one of the initial call attempts, but the case was not completed after 9 call attempts. These cases were released for additional call attempts after all telephone numbers in the sample had been attempted. This effort resulted in the completion of 2,207 additional Screeners, or 4 percent of all completed Screeners. Nearly all wave 1 Screener cases that were finalized in maximum call status received 20 or more call attempts; nearly all wave 2 or reserve cases received 14 call attempts.

Answering machine cases. The first time that an interviewer reached an answering machine at a sampled telephone number, the CATI system displayed a message that was read by interviewers that gave the purpose of the call and the study sponsorship, told the person that an interviewer could call back at another time, and gave a toll-free telephone number for requesting information, scheduling an appointment, or completing the interview. If no contact with a household member was made after seven attempts, the case was held for a week and refielded for another cycle of seven attempts. After 14 attempts, a case was designated as NM, "no answer, answering machine" if the only contact had been with an answering machine. NM cases in wave 2 and the reserve sample received a maximum of 14 call attempts. Those in wave 1 received 21 or 28 call attempts. Of the 73,664 telephone numbers with at least one answering machine result, 22,282 (30 percent) completed a Screener. Among all completed screeners (54,034), there were 31,753 numbers (59 percent) that never had an answering machine result.

"No answer" calls. Numbers at which no contact with a person or answering machine was made during the seven initial contact attempts (NA, or "no answer" cases) have historically proved to be the least productive. After 7 attempts to reach a household member by the Westat TRC, all cases that were classified as NA were sent to a vendor for an additional 14 calls using predictive dialing methods. If

a household member answered as a result of the predictive dialing attempts, the case was sent back to the Westat TRC for additional followup by NHES-trained telephone interviewers. Westat sent 31,907 noncontact cases to the vendor for predictive dialing follow-up. Of those, 1,896 cases were returned to Westat following contact with a household member and of those, 455 cases, or approximately 1 percent of all cases sent for predictive dialing, were completed.

4.1.6   Data Collection Quality Control

Data collection quality control efforts began during the CATI development period. As the CATI system was programmed, extensive testing of the system was conducted. This testing included review by project research staff, telephone interviewing staff, data preparation staff, statistical staff, and the programmers themselves. The testing by staff members representing different aspects of the project was designed to ensure that the system was working properly from all of these perspectives. Two field tests were conducted prior to data collection to ensure that the CATI system was working properly and the timing and flow of the instruments was as expected. In the first field test, 55 SR interviews and 64 PFI interviews were completed. In the second field test, 154 SR interviews and 253 PFI interviews were completed.

Quality control activities continued during interviewer training and data collection. During interviewer training, interviewers paired with one another to conduct role-play interviews on telephones monitored by supervisors. When interviewers began actual data collection, they were monitored on an ongoing basis by telephone center supervisors. Project research staff also monitored the interviewers, especially during the beginning weeks of data collection. Data preparation staff reviewed the cases from the CATI system as they were completed and referred problems to the project staff for resolution. Interviewer memos were posted and distributed when any observations indicated that reminders to the interviewers were appropriate (e.g., to always confirm spellings of names). Additional training or coaching was provided to interviewers as necessary.

Throughout data collection, supervisors and telephone monitors (experienced telephone interviewers who were trained for monitoring) listened for about 15 minutes at a time to the interviewers from either a monitoring room or from a carrel on the floor of the telephone center. The monitors completed a special monitoring form that covered five major areas of telephone interviewing:

- Voice quality and reading skills;
- Listening, probing, and clarifying skills;
- Technical skills;
- Gaining respondent cooperation; and
- Interview management.

The monitors recorded their impressions of the interviewer's skills and abilities along with suggestions for improvement. Interviewers were individually coached by supervisors, and any who had exhibited difficulty were intensively monitored to make sure the difficulties were resolved. If the problems continued, then the interviewers were released from the NHES:2007 interviewing staff.

At least once a week, the CATI management system produced computer-generated reports that displayed unit response rates, refusal rates, and refusal conversion rates for each NHES:2007 interviewer. These reports assisted telephone center supervisors in identifying differences in interviewer performance. Supervisors relied on both monitoring sheets and standard reports to make staff assignments. For example, standard reports might have shown that some interviewers were more effective in refusal conversion and monitoring those interviewers could have revealed persons particularly skilled in gaining cooperation from the elderly who could be assigned to conduct refusal conversion on those cases.

4.2 Unit Response Rates in NHES:2007

A unit response rate is the ratio of the number of units with completed interviews (e.g., the units could be telephone numbers, households, or persons) to the number of units sampled and eligible for the interview. In some cases, these rates are easily defined and computed, while in other cases the denominator of the ratio must be estimated.

For reporting the results from NHES:2007, the overall unit response rate indicates the percentage of possible interviews that were completed taking all survey stages into account, while the unit response rate measures the percentage of interviews that were completed for a specific stage of the survey. Specifically, household members were identified for interviews in a two-stage process. Screener interviews were conducted to enumerate and sample household members, and then topical surveys were administered for the sampled members. If the first-stage Screener was not completed, no members could be sampled for other interviews. Under this design, the unit response rate for the first stage is the estimated percentage of households that completed the Screener. The unit response rate for the second stage (SR or PFI interviews) is the percentage of sampled persons that completed these extended interviews. The overall unit response rate is the product of the first- and second-stage unit response rates (i.e., the Screener unit response rate multiplied by the extended interview unit response rate)."

Unit response rates can be either unweighted or weighted. The unweighted rate, computed using the raw number of cases, provides a useful description of the success of the operational aspects of the survey. The weighted rate, computed by summing the weights (usually the reciprocals of the probability of selecting the units) for both the numerator and denominator, gives a better description of the success of the survey with respect to the population sampled, since the weights allow for inference of the sample data (including response status) to the population level. Both rates are usually similar unless the probabilities of selection and the unit response rates in the categories with different selection probabilities vary considerably. All of the unit response rates discussed below are weighted unless noted specifically in the text.

Unit response rates and overall unit response rates are identical for the first stage of sampling and interviewing (i.e., the Screener). The next section discusses the unit response rate for the Screener and provides a profile of the characteristics of the respondents. The discussion of unit response rates and overall unit response rates for SR and PFI interviews are given in the sections that follow.

4.2.1 Screener Unit Response Rate

The first panel of table 4-1 shows the disposition of the 278,490 telephone numbers that were fielded in NHES:2007. The three major categories of residential status are those identified as numbers for

residential households, those identified as nonresidential numbers (primarily nonworking and business telephone numbers), and those numbers that, despite numerous attempts, could not be classified as either residential or nonresidential. Calculation of unit response rates is complex because of the possible ways residential status can be assigned to this last group of numbers.

        As shown in the lower part of the table, the first weighted unit response rate of 52.8 percent for the Screener was calculated using a method where an outside vendor provided information about cases for which no contact was made. The vendor-assisted approach uses information from an outside vendor about cases for which no answer was obtained in the estimation of their residency rate. Because this approach uses direct information about likely residential status associated with the particular telephone number, this approach yields more accurate estimates of residency rates than the survival method that was previously used to estimate residency rates in NHES. To enable evaluation of the accuracy of the vendor-assisted method in classifying telephone numbers as residential, a subsample of cases with known residential status (based on the results of the NHES:2007 call attempts) was included in the set of numbers sent to the vendor for residential status identification. Among Screener cases resolved as residential (either respondents or nonrespondents) through the NHES:2007 data collection effort, 90 percent were identified by the vendor as residential. Among Screener cases resolved as nonresidential in NHES:2007 data collection, 82 percent were identified by the vendor as nonresidential. Estimates based on the vendor's results suggest that 37.6 percent of telephone numbers with undetermined residency status in NHES:2007 are residential. Therefore, the denominator of the vendor-assisted unit response rate is the weighted total number of residential telephone numbers plus the 37.6 percent of the weighted total of numbers with unknown residential status that are estimated to be residential. The numerator is the weighted number of telephone numbers for households that participated in the survey (i.e., that completed a Screener). Both the numerator and the denominator have been weighted by the probabilities of selecting the telephone numbers and weighted for the subsampling of nonrespondents and no-answer telephone numbers for extensive followup.

Table 4-1. **Number of telephone numbers dialed, by residential status and weighted and unweighted Screener unit response rates**

Screener response category	Number	Percentage of all numbers	Percentage of residential numbers
Total	278,490	100.0	
Identified as residential			100.0
Responded	54,034	19.4	46.9
Did not respond	61,139	22.0	53.1
Identified as nonresidential	142,254	51.1	†
Unknown residential status	21,063	7.6	†

Estimated Screener unit response rates[1]	Weighted rate (percent)	Unweighted rate (percent)
Vendor-assisted unit response rate	52.8	53.2
CASRO unit response rate	53.9	54.0
Conservative unit response rate	46.6	48.4
Liberal unit response rate	57.4	56.7

† Not applicable.

[1]All of the unit response rates use the weighted number of responding households (for weighted rates) or the unweighted number of responding households (for unweighted rates) as the numerator. The denominators vary but are all estimated totals. For the vendor-assisted method unit response rate, the proportion of unknown residential status numbers included in the denominator was estimated using information about the cases from an outside vendor. For the Council of American Survey Research Organizations (CASRO) unit response rate, the proportion of unknown residential status numbers included in the denominator was based upon the residency rate for the numbers with known residential status. For the conservative unit response rate, all of the unknown residential status numbers were included in the denominator. For the liberal unit response rate, none of the unknown residential status numbers were included in the denominator. For the unweighted rates, because refusal conversion procedures were not used for non-subsampled cases, these unweighted rates are lower than they would have been if such procedures had been used. There is no reason to expect that the response rate for this group would have been measurably different from the response rate for those cases that were subject to refusal conversion. Therefore, assuming the response rates for these groups are equal, the response rate would be unaffected by inclusion or exclusion of these cases from the response rate calculation.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Surveys Program, 2007.

Other estimates of the unit response rates were computed by allocating different proportions of the numbers with unknown residency status into the residential category in the denominator. (The footnote to table 4-1 explains four different schemes for estimating the unit response rate.) Based on the calculation of these different rates, it is reasonable to say that the Screener unit response rate is between 47 and 57 percent (the lower percentage assumes that all numbers of unknown status are residential while the higher percentage assumes that all numbers of unknown status are nonresidential). The variability in the estimates arises because it is not possible to identify precisely the residential status for each telephone number. The vendor-assisted method unit response rate is believed to be the most accurate because it uses data about the sampled telephone numbers in the estimation of the residency rate. Using this approach, the best estimate for the NHES:2007 Screener unit response rate is 53 percent.

The lower right part of table 4-1 also shows unweighted Screener unit response rates calculated using each of the approaches described earlier. If the raw count of telephone numbers had not been weighted, the Screener unit response rate using the vendor-assisted method would have been 53.2 percent.

Table 4-2 presents the Screener unit response rate by selected geographic area characteristics and characteristics of telephone numbers. These characteristics were considered because they are available for all telephone numbers and are sometimes associated with response propensity. The unit response rate was higher for telephone numbers with mailable addresses than for those without mailable addresses,[24] and was also higher for households where no answering machine message was left. The Screener unit response rate also varied by region of the country, with the highest unit response rates in the West North Central and East North Central divisions and the lowest unit response rates in the Pacific division. Areas with higher proportions of Whites generally had higher unit response rates than those with lower proportions of Whites, and areas with lower proportions of Hispanics, Blacks, and Asians had higher unit response rates than those with higher proportions in these subgroups. Areas with lower median home values generally had higher unit response rates than those with higher median home values. Areas with higher proportions of renters had lower response rates than those with lower proportions of renters. Areas in a county or not in a central city had significantly higher response rates than subcounties of an MSA or an MSA in its own county; similarly, areas in subcounties of an MSA or an MSA in its own county had significantly lower response rates than non MSAs.

---

[24] Mailable addresses received different levels of advanced incentives (see section 4.1.1). Unit response rates were lowest (32.4 percent) for wave 2 cases receiving the $2 advance incentive, higher (35.1 percent) for reserve cases receiving the $5 advance incentive, and highest (38.2 percent) for reserve cases receiving the $10 advance incentive. The wave 2 and reserve cases did not receive the followup treatment that the wave 1 cases received. In the calculation of all other response rates, an adjustment was made to account for the subsampling of cases for followup.

Table 4-2.   Number of telephone numbers dialed in the Screener, by response status, weighted unit response rate, and characteristic of the geographic area based on the telephone exchange

Characteristic	Total	Residential, responded	Residential, did not respond[1]	Non-residential	Unknown residential status	Estimated unit response rate (percent)[2]
Total	278,490	54,034	61,139	142,254	21,063	52.8
Mailable status						
Mailable address	152,261	51,323	56,312	30,239	14,387	54.4
No mailable address	126,229	2,711	4,827	112,015	6,676	43.3
Answering machine message indicator						
No message left	202,706	22,329	32,254	10,133	11,068	65.7
One or more messages left	75,784	31,705	28,885	132,121	9,995	42.3
Percent White						
Less than 30 percent	36,973	5,884	3,273	18,920	2,896	43.4
30 to 39 percent	16,950	2,619	3,583	9,364	1,384	46.1
40 to 59 percent	52,974	9,110	11,586	27,870	4,408	47.8
60 to 69 percent	34,345	6,420	7,530	17,661	2,734	49.7
70 to 79 percent	33,694	6,813	7,255	17,141	2,485	52.8
80 to 89 percent	33,382	6,826	7,165	16,814	2,577	53.6
90 percent or more	70,172	16,362	14,747	34,484	4,579	58.0
Percent Hispanic						
1st through 4th deciles	250,838	49,401	54,168	128,451	18,818	53.4
5th through 9th deciles	27,652	4,633	6,971	13,803	2,245	44.7
Median home value						
1st decile	26,441	5,154	5,366	14,601	1,320	57.6
2nd through 4th deciles	83,127	17,420	17,200	43,298	5,209	57.0
5th through 6th deciles	56,373	11,403	12,108	28,612	4,247	54.3
7th through 9th deciles	84,794	15,623	20,011	41,705	7,455	49.3
10th decile	27,755	4,431	6,454	14,038	2,832	44.2
Percent renters						
1st through 5th deciles	239,046	49,103	53,638	118,723	17,582	53.8
6th through 9th deciles	39,444	4,931	7,501	23,531	3,481	42.3
Percent college graduates						
Less than 20 percent	55,666	10,707	12,364	29,163	3,432	54.2
20 to 29 percent	104,188	21,090	23,485	52,213	7,400	54.0
30 percent or more	118,636	22,237	25,290	60,878	10,231	51.3
Percent Black						
1st through 5th deciles	259,388	50,993	57,170	131,405	19,820	53.0
6th through 9th deciles	19,102	3,041	3,969	10,849	1,243	48.3

See notes at end of table.

Table 4-2. **Number of telephone numbers dialed in the Screener, by response status, weighted unit response rate, and characteristic of the geographic area based on the telephone exchange—Continued**

Characteristic	Total	Residential, responded	Residential, did not respond[1]	Non-residential	Unknown residential status	Estimated unit response rate (percent)[2]
Percent Asian						
Less than 10 percent	211,010	42,948	45,760	107,609	14,693	54.7
10 to 19 percent	49,839	8,408	11,022	25,850	4,559	47.6
20 to 29 percent	10,739	1,691	2,550	5,435	1,063	44.4
30 percent or more	6,902	987	1,807	3,360	748	39.2
Metropolitan status						
In county in central city	108,958	19,091	23,123	58,323	8,421	50.6
In county not in central city	51,820	10,415	12,547	24,553	4,305	51.0
Subcounty of MSA or MSA its own county	68,464	13,706	15,640	33,528	5,590	52.2
Non-MSA	49,248	10,822	9,829	25,850	2,747	59.1
Median income						
1st through 4th deciles	107,159	20,103	22,318	57,789	6,949	54.3
5th and 6th deciles	56,590	11,268	12,852	28,269	4,201	53.6
7th decile or higher	114,741	22,663	25,969	56,196	9,913	54.5
Census region						
Northeast	49,177	9,210	11,580	24,349	4,038	49.9
Midwest	55,900	11,961	10,577	29,889	3,473	59.3
South	110,002	21,033	23,760	57,462	7,747	52.6
West	63,411	11,830	15,222	30,554	5,805	49.2
Census division						
New England	11,293	2,322	2,684	5,553	732	51.9
Middle Atlantic	37,884	6,888	8,896	18,796	3,304	49.1
East North Central	39,889	8,202	7,837	21,160	2,690	57.3
West North Central	16,011	3,759	2,740	8,729	783	64.0
South Atlantic	59,460	11,407	13,017	30,598	4,438	51.6
East South Central	16,046	3,362	3,417	8,425	842	56.7
West South Central	34,496	6,264	7,326	18,439	2,467	52.3
Mountain	18,085	3,658	3,729	9,368	1,330	55.8
Pacific	45,326	8,172	11,493	21,186	4,475	46.5

[1]The "residential, did not respond" counts include those nonrespondents not selected for extensive followup (i.e., the nonrespondents in wave 2).

[2] The estimated unit response rate is the vendor-assisted method unit response rate (i.e., the number of completed interviews divided by the sum of the number of completed interviews, nonresponses, and 37.6 percent of telephone numbers with an unknown residency status, weighted by the probability of selection).

NOTE: Deciles are any one of the numbers or values in a series dividing the distribution of the individuals in a series into 10 groups of equal frequency. The categories of each characteristic are based on a multivariate analysis for the nonresponse weighting adjustment, in which the sample was divided into subgroups with the most differential response rates.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Surveys Program, 2007.

4.2.2   Extended Interview Unit Response Rates

The number of persons enumerated and sampled, and those with completed interviews for each survey of NHES:2007, are given in table 4-3. Of the enumerated 4,030 children eligible for the SR interview, a sample of 3,706 children was selected; and of the enumerated 23,882 children eligible for the PFI interview, a sample of 14,021 children was selected. About 6.4 percent of the SR sampled children (n = 239) were classified as ineligible because they were enumerated in error (i.e., were not household members at the time of screening) or were not actually in the age and grade range eligible for the survey according to the reports of the SR interview respondents, and less than one percent of PFI sampled children (n = 92) were classified similarly. Completed SR interviews were obtained for 2,633 of the sample children for an estimated 77.0 percent single stage response rate and an overall response rate of 40.7 percent. PFI interviews were obtained for 10,681 of the sampled children for an estimated 74.1 percent single stage response rate and an overall response rate of 39.1 percent. The bulk of the unit nonresponse for the SR and PFI interviews was due to refusal of the parent/guardian to respond, 48.3 percent (n = 355) and 49.4 (n = 1,652) percent of nonresponse, respectively). Other reasons for SR and PFI interview unit nonresponse were inability to complete the interview with the parent/guardian despite many attempts (35.2 percent (n = 259) and 37.3 percent (n = 1,249) of nonresponse, respectively), language problems (2.0 percent (n = 15) of nonresponse for SR and 1.8 percent (n = 61) of nonresponse for PFI), and other miscellaneous reasons such as the parent/guardian being unavailable for an interview during the field period (14.4 percent (n = 106) of nonresponse for SR and 11.5 percent (n = 385) of nonresponse for PFI).

The unit response rates for the SR and PFI interviews can only be examined by variables available for both respondents and nonrespondents. For persons sampled for extended interviews in the NHES:2007 surveys, such variables are those available on the sampling frame and those available from the Screener. The variables shown for the SR interview are Census region (based on the telephone number) and grade/enrollment status. For the PFI interview the variables shown are Census region (based on the telephone number); grade or grade equivalent (collected in the Screener); and type of schooling (regular or home school, also collected in the Screener). Table 4-4 shows the number of sampled children by response status and unit response rate for each of the SR interview variables, and table 4-5 shows the number of sampled children by response status and unit response rate for each of the PFI interview variables. There is little variation in unit response rates by Census region, by grade for both SR and PFI, and by school type for PFI. The total number in the responded column matches the number of children sampled for each survey, regardless of how they completed. For SR, this includes 2,616 children sampled and completed as SR plus 116 children sampled as SR but completed as PFI, for a total of 2,732. For PFI, the number includes the 10,565 children sampled and completed as PFI plus the 17 children sampled as PFI but completed as SR, for a total of 10,582.

Table 4-3. **Number of enumerated children, completed interviews, and weighted unit response and overall unit response rates, by type of extended interview**

Type of interview	Number	Estimated unit response rate (percent)	Estimated overall unit response rate (percent)[1]
SR interview		77.0	40.7
Enumerated	4,030		
Sampled[2]	3,706		
Ineligible	239		
Did not respond	735		
Total complete	2,633		
Sampled as PFI, completed as SR	17		
Sampled as SR, completed as SR	2,616		
PFI interview		74.1	39.1
Enumerated	23,882		
Sampled[2]	14,021		
Ineligible	92		
Did not respond	3,347		
Total complete	10,681		
Sampled as SR, completed as PFI	116		
Sampled as PFI, completed as PFI	10,565		

[1]The estimated overall unit response rate is computed by multiplying the Screener unit response rate of 52.8 percent by the appropriate extended interview unit response rate.

[2]The number sampled for the SR interview includes the number sampled as SR, completed as SR (2,616); the number sampled as SR, completed as PFI (116); the number ineligible (239); and the number that did not respond (735). The number sampled for the PFI interview includes the number sampled as PFI, completed as PFI (10,565); the number sampled as PFI, completed as SR (17); the number ineligible (92); and the number that did not respond (3,347).

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Readiness (SR) and Parent and Family Involvement in Education (PFI) Survey of the National Household Education Surveys Program , 2007.

Table 4-4.   Number of sampled SR interviews, by response status and weighted unit response rates

SR interviews and status at sampling	Total	Responded[1]	Did not respond	Ineligible	Estimated unit response rate
Total	3,706	2,732	735	239	77.0
Census region					
Northeast	640	458	147	35	74.0
South	825	628	145	52	79.3
Midwest	1,323	955	267	101	75.9
West	918	691	176	51	78.5
Grade of child (Screener)					
Unenrolled/unknown/other	1,354	936	257	161	77.0
Preschooler	2,352	1,796	478	78	77.0

[1]Includes all completed interviews that were sampled for the SR interview, regardless of whether the interview completed was the SR interview or the PFI interview. This includes 2,616 children sampled and completed as SR plus 116 children sampled as SR but completed as PFI, for a total of 2,732.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Readiness (SR) Survey of the National Household Education Surveys Program , 2007.

Table 4-5. Number of sampled PFI interviews, by response status and weighted unit response rates

PFI interviews and status at sampling	Total	Responded[1]	Did not respond	Ineligible	Estimated unit response rate
Total	14,021	10,582	3,347	92	74.1
Census region					
Northeast	2,456	1,818	626	12	71.9
South	3,060	2,386	655	19	75.7
Midwest	5,176	3,849	1,294	33	73.1
West	3,329	2,529	772	28	75.6
Grade of child (Screener)					
Kindergarten	1,031	761	270	0	71.6
1st grade	1,093	795	296	2	72.5
2nd grade	1,025	762	259	4	73.5
3rd grade	942	700	240	2	71.8
4th grade	937	704	230	3	73.0
5th grade	967	732	232	3	73.5
6th grade	910	723	184	3	77.4
7th grade	965	765	196	4	77.9
8th grade	1,043	798	240	5	75.7
9th grade	1,102	820	274	8	72.2
10th grade	1,137	895	236	6	76.6
11th grade	1,213	920	289	4	73.8
12th grade	1,242	900	315	27	72.4
Other/unknown[2]	414	307	86	21	75.0
School (Screener)					
Regular school	13,493	10,192	3,235	66	78.2
Homeschool	422	343	74	5	74.0
Unknown	106	47	38	21	56.6

[1]Includes all completed interviews that were sampled for the PFI interview, regardless of whether the interview completed was the SR interview or the PFI interview. This includes the 10,565 children sampled and completed as PFI plus the 17 children sampled as PFI but completed as SR, for a total of 10,582.

[2]Other included ungraded and special education.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Parent and Family Involvement in Education (PFI) Survey of the National Household Education Surveys Program , 2007.

4.3   Item Response Rates

For most of the data items collected in the NHES:2007 surveys, the item response rates were very high. The tables in this section show the item response rates for a representative group of items from each interview. The items included were selected to represent key items considered in the sample design and to represent the range of item response rates. The number of cases for which each item was attempted and the percentage of cases for which a valid response was obtained are shown.

Tables 4-6 and 4-7 show the item response rates and total response rates for a representative group of items from the SR and PFI interviews, respectively. For the SR and PFI surveys, the median item response rates were 99.31 percent and 99.04 percent, respectively, and the median total response rates (the product of the item response rate and the overall unit response rate for the survey) were 40.42 and 38.72, respectively. For items that are asked only of a small subgroup of respondents, a small number of missing values could result in a low item response rate.

Most items on the SR public use data file have item response rates over 90 percent. The only items with item response rates of less than 90 percent are items that apply to only a small number of cases (MOMTYPE2, MOMNEW2, MOMLANG2, MOMSPEAK2, MOMBORN2, MHISPAN2, MWHITE2, MBLACK2, MAMIND2, MASIAN2, MPACI2, MRACEOTH2, MOMGRADE2, MOMWORK2, MOMHOURS2, MOMMTHS2, MOMENROL2, and MOMSTAT2).

Items with item response rates of less than 90 percent on the PFI public file include items that apply to only a small number of cases (MOMDIPL2, MOMVOTEC2, MOMSTAT2, MOMLEAVE2, and MOMLOOK2).

Table 4-6.   Item response rates and total response rates for selected items in the SR interview

Item	Number attempted	Item response rate	Total response rate[1]
Demographic characteristics			
Relationship to child–person 1	2,302	99.83	40.63
Language child speaks most at home	2,633	99.85	40.64
State, country, or territory child born in	2,633	99.85	40.64
Current school status			
Child enrolled/attending school	2,633	99.96	40.68
Child's grade in school	1,709	100.00	40.70
Child attending daycare, preschool, pre-K, or Head Start	2,633	99.73	40.59
When child expected to start Kindergarten	2,633	96.58	39.31
Developmental characteristics			
Child can identify red, yellow, blue and green	2,633	99.85	40.64
Child's method of holding a pencil	2,633	98.56	40.11
Child's ability to rhyme words	2,633	92.25	37.55
Family involvement outside of school			
Number of books child owns	2,633	98.21	39.97
Times read to child in past week	2,633	99.92	40.67
Child able to read story books on his/her own	2,633	99.66	40.56
Child reads words or pretends to read	337	98.81	40.22
Number of times per week for family dinner	2,633	99.92	40.67
Visited a library in the past month	2,633	99.85	40.64
Visited an art gallery, museum, or historical site in the last month	2,633	99.89	40.66
Number of hours per weekday child watches TV/videos	2,633	99.05	40.31
Number of hours per weekend day child watches TV/videos	2,553	99.29	40.41
Child watches Disney Channel at least once per week	2,504	98.08	38.02
Child watches a PBS channel at least once per week	2,504	98.08	38.02
Number of times per week adult in household reads	2,633	99.85	40.64
Role of parent in preparing child for school			
Importance of teaching child alphabet	2,633	99.89	40.66
Importance of teaching child to share	2,633	99.92	40.67
Importance of teaching child to read	2,633	99.77	40.61
Importance of teaching child numbers	2,633	99.89	40.66
Health and disability			
Child weighed less than 5.5 lbs. at birth	2,633	98.52	40.10
Doctor/health professional concerned about child's weight	2,633	99.92	40.67
Child has specific learning disability	2,633	99.81	40.62
Child has Pervasive Developmental Disorder (PDD)	2,633	99.77	40.61
Household characteristics			
Received WIC benefits	2,633	99.35	40.44
Received Food Stamps in past month	2,633	99.35	40.44
Received Section 8 housing assistance	2,633	99.28	40.41
Total household income range	2,633	89.67	36.50

[1]The total response rate for a given item is the product of the overall unit response rate for the survey and the item response rate for the item.

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Readiness (SR) Survey of the National Household Education Surveys Program, 2007.

Table 4-7.   Item response rates and total response rates for selected items in the PFI interview

Item	Number attempted	Item response rate	Total response rate[1]
Demographic characteristics			
Relationship to child–person 1	8,526	99.47	38.89
Language child speaks most at home	10,681	99.64	38.96
State, country, or territory child born in	10,681	99.63	38.96
Current school status			
Child enrolled/attending school	10,681	100.00	39.10
Child's grade in school	10,370	99.94	39.08
School characteristics			
Child attends public/private school	10,370	99.64	38.96
Allowed to choose school in any school district	1,453	97.87	38.27
Other schools considered for child	10,370	99.49	38.90
School ID for child's school provided	10,370	92.80	36.28
Student experiences			
Child's grades across all subjects	10,370	98.22	38.40
Child enrolled in advanced classes	3,543	97.66	38.19
Family/school involvement and school practices			
Attend general school meeting	10,370	99.63	38.96
Number of hours volunteering/fundraising	10,370	99.65	38.96
Family involvement in schoolwork			
How often homework done outside school	10,370	99.18	38.78
Family involvement outside of school			
Visited library in past month	10,681	99.59	38.94
Visited zoo/aquarium in past month	10,681	99.68	38.97
Health and disability			
Rating of child's health	10,681	99.72	38.99
Household worked with school to develop IEP	737	95.39	37.30
Parent/guardian characteristics			
Mother's marital status	10,287	99.24	38.80
Country where mother was born	10,287	99.27	38.81
Father has vocational/technical diploma	1,087	96.04	37.55
Household characteristics			
Received WIC benefits	10,681	99.05	38.73
Received Food Stamps in past month	10,681	99.03	38.72
Received Section 8 housing assistance	10,681	98.99	38.71
Total household income range	10,681	90.45	35.37

[1]The total response rate for a given item is the product of the overall unit response rate for the survey and the item response rate for the item.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Parent and Family Involvement in Education (PFI) Survey of the National Household Education Surveys Program, 2007.

4.4   Unit Nonresponse Bias

The estimates from the NHES:2007 surveys are subject to potential bias because of unit nonresponse to the Screener and the extended interview components. Generally speaking, the best approach to minimizing nonresponse bias is to plan and implement data collection procedures aimed at achieving high cooperation rates. For NHES:2007, such procedures included extensive training of the interviewers, advance mailings to the respondents, effective call scheduling strategies, and, where necessary, refusal conversion methods that included recontacting households by both telephone and mail if mailable addresses could be obtained, and monetary incentives. However, because some unit nonresponse occurs even with the best strategies, weighting adjustments are necessary to minimize potential unit nonresponse bias.

The term bias has a specific technical definition in this context. Bias is the expected difference between the estimate from the survey and the actual population value. For example, if all households were included in the survey (i.e., if a census was conducted rather than a sample survey), the difference between the estimate from the survey and the actual population value (which includes persons who did not respond to the survey) is the bias due to unit nonresponse. Since NHES is based on a sample, the bias is defined as the expected or average value of this difference over all possible samples.

Unit nonresponse bias, or the bias due to the failure of some persons or households in the sample to respond to the survey, can be substantial when two conditions hold. First, the differences between the characteristics of respondents and nonrespondents must be relatively large. For example, consider estimating the percentage of preschoolers who can recognize all the letters of the alphabet. If the percentage is nearly identical for both respondents and nonrespondents, then the unit nonresponse bias of the estimate will be negligible.

Second, the unit nonresponse rate must be relatively high. If the nonresponse rate is very low relative to the magnitude of the estimates, then the unit nonresponse bias in the estimates will be small, even if the differences in the characteristics between respondents and nonrespondents are relatively large. For example, if the unit nonresponse rate is only 2 percent, then estimates of totals that compose 20 or 30 percent of the population will not be greatly affected by nonresponse, even if the differences in these characteristics between respondents and nonrespondents are relatively large. It is important to realize that this condition requires the unit nonresponse rate to be large relative to the size of the estimates. If the estimate is for a small domain or subgroup, then even a relatively low overall rate of nonresponse can result in important biases if the differences between respondents and nonrespondents are large.

A special study was undertaken to examine nonresponse and the potential bias associated with unit nonresponse and noncoverage in NHES:2007. This study involved the selection of a sample of addresses in 30 areas around the country. Selected addresses were matched to telephone numbers where possible, and the matched cases were included with the RDD sample in data collection, receiving the same treatment as wave 1 of the RDD sample (see section 4.1 for details). In-person followup was conducted with those cases that were not matched with telephone numbers and matched cases that were not completed in telephone data collection.

Following up with cases with no matched telephone numbers allowed an investigation of noncoverage bias, since completed interviews were obtained for non-telephone households. Noncoverage bias was assessed by comparing estimates from persons in households with a landline telephone to

estimates from the full sample. Following up with nonrespondents from the telephone effort allowed an investigation of nonresponse bias. Nonresponse bias was assessed by comparing estimates from the full data collection effort (those completed by telephone or in-person followup) versus the reduced effort (those competed by telephone only).[25] An estimate was considered to have potential bias if it was both statistically significant and of substantive importance.[26] The comparison is limited because survey estimates for the households that did not respond to either the telephone or in-person effort are not available. The results of the bias study indicate that the NHES:2007 SR and PFI Survey estimates do not suffer from significant nonresponse bias, but the SR Survey estimate of the percentage of parents with education beyond a high school diploma includes a positive noncoverage bias.

The NHES:2007 estimates were produced using weights that were adjusted for nonresponse and calibrated to population totals—adjustments that are expected to reduce nonresponse and noncoverage bias. To evaluate potential bias in the estimates before weighting adjustments were applied, as well as the effect the weighting adjustments may have had on potential bias in the data, the RDD estimates generated using unadjusted weights were compared to the nonresponse-adjusted estimates and to the fully weighted (i.e., nonresponse-adjusted and raked) estimates. A comparison of estimates before and after the weighting adjustments indicated potential bias in several unadjusted SR Survey estimates that was reduced through the weighting process.[27] For the PFI Survey, although there were some differences in estimates of demographic characteristics, the estimates of key survey outcome variables at each stage of the weighting were comparable.

Tables 4-8 and 4-9 show estimates of overall bias for the SR and PFI Surveys, respectively. Overall bias was evaluated by comparing estimates from the bias study to those from the fully weighted RDD sample. Since no evidence of nonresponse bias was found, any differences can be thought to be attributable to noncoverage bias. As shown in table 4-8, the bias study sample yielded larger estimates of the percentage of preschoolers who can count to 20 or higher, the percentage of preschoolers whose speech is often understandable to a stranger, and the percentage of preschoolers who watch 2 or more hours of television in a typical weekday. There is no systematic relationship among these differences that would be indicative of bias, and it is likely that these differences were found mainly as a result of having examined so many characteristics.

There is a difference in the estimates of the percentage of preschoolers whose mothers are not in the labor force (36 percent from the RDD sample versus 26 percent from the bias study sample). This difference may be an indication of accessibility, with mothers who are not in the labor force being more available and more willing to complete the interview by telephone than mothers with other employment

---

[25] Cases with no matching telephone number could not be attempted in the initial telephone effort, but were attempted in-person. Thus, it should be noted that both the full effort and the reduced effort estimates included the nonmatched cases in order to eliminate the effect the nonmatched cases would have on estimates of nonresponse bias.

[26] All differences discussed in this section are significant at the 95 percent confidence level, based on a 2-sided *t*-test. While some relatively small differences (3 to 5 percentage points) might be statistically significant when sample sizes are large, the discussion is limited to differences that are potentially of substantive importance. Differences of substantive importance are defined as differences of 5 percentage points or more or relative differences of 3 or more (i.e., when one estimate is 3 or more times larger than the other). The Bias Study was designed to allow detection of a 5 percentage point difference in key statistics. For NHES, this is considered a meaningful threshold to use to identify which statistically significant differences are of substantive significance.

[27] Estimates after the weighting adjustments were lower than the unadjusted estimates for the proportion of preschoolers who participate in center-based care, recognize all colors, count to 20 or higher, and write their first name; who have parents who believe it is essential to prepare their child for kindergarten by teaching them the alphabet, numbers, and sharing; who have a family member that reads to them everyday in the past week; whose parents took three or more outings with them in the past month; who have household incomes above $50,000; and who have both a mother and father in the household. The final estimates were higher than the unadjusted estimates for the proportion of preschoolers who are 3 years-old, live in homes that are not owned, have parents with a high school diploma or below, are below the poverty threshold, have household incomes below $30,000, and have a mother only in the household.

status. One other curious difference is in the sex distribution of preschoolers; the bias study sample estimated 62 percent of preschoolers to be male, compared to 50 percent for the RDD sample and 52 percent for the CPS sample in October 2005.[28]

Table 4-9 shows there are no significant differences of substantive importance between the PFI Survey estimates for the bias study sample versus those for the RDD sample. However, there are some smaller differences of statistical significance. For instance, the estimated percentage of parents who report the school provides information very well about how to help the student plan for college or vocational school is 34 percent for the RDD sample, which is 13 percent lower than the bias study estimate. In addition, the estimated percentage of students in schools of size 1,000 or more is 12 percent lower for the RDD sample than the bias study sample (27 percent versus 31 percent).

In addition to the special bias study, the assessment of nonresponse bias also included a comparison of NHES:2007 estimates with estimates from the Current Population Survey and prior NHES collections. Those comparative tables appear in appendix C.

As presented in appendix C, selected estimates from the NHES:2007 surveys were compared to those from previous NHES cycles and from the Current Population Survey (CPS), which contained the same or comparable items. Most of the comparisons do not show statistically significant differences that are potentially of substantive importance (defined as differences of 5 percentage points or more).

The significant differences of 5 percentage points or more between SR-NHES:2007 and CPS:2005-2006 were observed in the estimates of Hispanics with a household income less than $15,000 (table C-6); and Whites with a household income from $30,001 to $50,000 and $50,000 or more (table C-6).

The observed substantive differences between SR-NHES:2007 and ECPP-NHES:2005 were in the estimates of the percentage of preschoolers' parents with a high school graduate education (table C-12); the percentage of preschoolers' parents with a high school graduate education for Whites and for Blacks (table C-13); and the percentage of preschoolers' parents with a graduate school education for Blacks (table C-13).

Differences between PFI-NHES:2007 and CPS:2005-2006 were observed in estimates of public school enrollment for kindergartners (table C-4); households earning more than $75,000 annually (table C-7); household income for Whites from $30,001 to $50,000 and $50,000 or more (table C-8); household income for other race/ethnicity groups from $30,001 to $50,000 (table C-8); both parents in the household (table C-18); and parent with a graduate school education (table C-18).

The observed differences between PFI-NHES:2007 and PFI-NHES:2003 were in estimates of the percentage of students attending schools with 300-599 students and 600-999 students (table C-17); the

---

[28] To evaluate the effect of the skewed sex distribution on the analysis of overall bias, the bias study weights were re-raked, including sex as an additional raking dimension. The analysis in Table 4-8 was then reproduced with the new weights. The conclusions remained the same, with a few exceptions. First, the difference between the RDD sample and bias study sample estimates of the percentage of preschoolers who can count to 20 or higher was no longer of substantive importance after re-raking. The estimated bias was -5.2 percentage points before re-raking and -4.8 percentage points after. Second, the difference in the percentage of preschoolers of age 4 years was statistically significant after re-raking. The estimated bias was 5.5 percentage points (not significant) before re-raking and 6.7 percentage points after. Finally, the difference in the percentage of preschoolers whose mother was looking for work increased from -4.5 percentage points (not of substantive importance) before re-raking to a statistically significant and substantive difference of -5.5 percentage points after.

percentage of students' parents with a graduate school education for Whites (table C-19); and blindness or another visual impairment[29] (table C-22).

---

[29] In PFI-NHES:2003, the question was phrased as "Has a health professional told you that your child has blindness or other visual impairment?", whereas in PFI-NHES:2007, the question was phrased as "Has a health professional told you that your child has blindness or another visual impairment *not corrected with glasses*?"

Table 4-8.   Estimates of overall bias for various characteristics from the School Readiness Survey: NHES:2007 RDD and NHES:2007 Bias Study

Characteristic	All RDD sample SR respondents			All Bias Study sample SR respondents			Bias[1]		Relative bias[2]
	Sample size	Estimate (percent)	Standard error	Sample size	Estimate (percent)	Standard error	Estimate (percent)	Standard error	Estimate (percent)
Participation in center-based care	1,759	55.3	0.89	177	51.1	2.29	4.2	2.46	8.2
Specific skills									
Recognizes all colors	2,265	82.5	1.02	249	85.6	2.34	-3.1	2.55	-3.6
Counts to 20 or higher	1,747	63.2	1.30	198	68.4	2.27	-5.2	2.61	-7.6
Recognizes all letters	891	31.8	1.19	93	29.4	2.89	2.4	3.12	8.2
Writes first name	1,709	59.8	1.27	193	60.4	3.08	-0.6	3.33	-0.9
Holds a pencil	2,298	86.8	0.95	252	86.7	1.90	#	2.12	0.1
Speech is often understandable to a stranger	2,170	81.1	1.20	249	87.3	2.17	-6.2	2.48	-7.1
Reads or pretends to read storybooks	2,582	98.0	0.31	289	99.2	0.45	-1.2	0.55	-1.2
Parents believe it is essential to do certain things to prepare child for kindergarten									
Teach child the alphabet	1,582	56.3	1.25	171	57.7	2.72	-1.4	2.99	-2.5
Teach child about sharing	1,732	61.8	1.37	177	57.8	2.97	4.0	3.27	6.9
Teach child to read	1,226	45.0	1.36	139	48.0	3.47	-3.0	3.73	-6.3
Teach child numbers	1,507	54.1	1.37	155	51.7	2.64	2.4	2.98	4.6
Show child how to hold a pencil	1,104	40.9	1.37	121	37.9	2.88	3.0	3.19	8.0
Family member read to child everyday in the past week	1,575	55.3	0.97	169	56.9	4.02	-1.6	4.14	-2.8

See notes at end of table.

Table 4-8. **Estimates of overall bias for various characteristics from the School Readiness Survey: NHES:2007 RDD and NHES:2007 Bias Study—Continued**

Characteristic	All RDD sample SR respondents			All Bias Study sample SR respondents			Bias[1]		Relative bias[2]
	Sample size	Estimate (percent)	Standard error	Sample size	Estimate (percent)	Standard error	Estimate (percent)	Standard error	Estimate (percent)
Parents report usually doing certain reading-related activity with child									
Ask child what is in a picture	1,390	55.5	1.26	159	52.9	4.00	2.6	4.20	4.9
Stop reading and point out letters	818	31.1	1.46	95	31.9	4.14	-0.8	4.39	-2.6
Ask child to read with parent	589	23.6	1.26	84	29.7	3.05	-6.1	3.30	-20.6
Talk about the story and what happened	1,419	56.1	1.44	165	60.3	3.18	-4.3	3.49	-7.0
Parents did home activities with child in the past week[3]	1,088	37.7	1.20	123	41.7	3.23	-4.0	3.45	-9.5
Parents took 3 or more outings with child in the past month[4]	1,175	39.0	1.10	138	46.0	3.71	-7.0	3.87	-15.3
Child watches 2 or more hours of TV in a typical weekday	1,538	61.6	1.19	179	71.1	2.10	-9.5	2.41	-13.3
Child has a disability	453	17.6	1.00	37	14.1	3.34	3.5	3.49	24.8
Child's age									
3 years	1,098	43.0	1.24	128	43.9	2.86	-0.9	3.12	-2.1
4 years	1,159	42.8	1.41	116	37.3	2.64	5.5	2.99	14.8
5 years and older	376	14.2	0.89	48	18.8	3.08	-4.6	3.20	-24.4

See notes at end of table.

Table 4-8. **Estimates of overall bias for various characteristics from the School Readiness Survey: NHES:2007 RDD and NHES:2007 Bias Study—Continued**

Characteristic	All RDD sample SR respondents			All Bias Study sample SR respondents			Bias[1]		Relative bias[2]
	Sample size	Estimate (percent)	Standard error	Sample size	Estimate (percent)	Standard error	Estimate (percent)	Standard error	Estimate (percent)
Child's sex									
Male	1,279	50.3	1.16	169	62.3	2.97	-12.0	3.19	-19.3
Female	1,354	49.7	1.16	123	37.7	2.97	12.0	3.19	31.9
Household urbanicity									
Urban	2,217	79.9	0.43	247	80.8	1.62	-0.9	1.67	-1.2
Rural	416	20.1	0.43	45	19.2	1.62	0.9	1.67	4.9
Home tenure									
Own	1,945	61.3	0.47	201	62.1	0.87	-0.8	0.99	-1.3
Rent/other	688	38.7	0.47	91	37.9	0.87	0.8	0.99	2.1
Parents' educational attainment									
High school diploma or below	603	27.8	1.26	91	34.5	3.80	-6.7	4.00	-19.5
Beyond high school diploma	2,030	72.2	1.26	201	65.5	3.80	6.7	4.00	10.3
Parents' language									
Both/only parent(s) speak(s) English	2,218	84.9	0.79	254	88.0	2.48	-3.1	2.60	-3.5
One of two parents speaks English	45	1.5	0.30	‡	‡	‡	0.5	0.78	50.0
No parent speaks English	370	13.6	0.77	29	11.0	2.60	2.6	2.71	23.9

See notes at end of table.

Table 4-8. Estimates of overall bias for various characteristics from the School Readiness Survey: NHES:2007 RDD and NHES:2007 Bias Study—Continued

Characteristic	All RDD sample SR respondents			All Bias Study sample SR respondents			Bias[1]		Relative bias[2]
	Sample size	Estimate (percent)	Standard error	Sample size	Estimate (percent)	Standard error	Estimate (percent)	Standard error	Estimate (percent)
Mothers' employment status									
35 hours or more per week	959	36.4	1.44	111	41.6	3.34	-5.2	3.64	-12.6
Less than 35 hours per week	597	20.9	0.94	69	20.4	3.21	0.5	3.35	2.3
Looking for work	108	5.5	0.70	23	10.0	2.49	-4.5	2.59	-44.7
Not in labor force	927	35.8	1.48	83	26.1	2.65	9.7	3.04	37.2
No mother in household	42	1.5	0.30	6!	1.9!	1.18!	-0.4	1.21	-20.2
Poverty status									
Poor	412	22.5	0.81	80	24.5	2.19	-2.0	2.34	-8.1
Nonpoor	2,221	77.5	0.81	212	75.5	2.19	2.0	2.34	2.6
Household income									
Less than $15,000	251	14.5	0.93	52	15.3	1.71	-0.8	1.94	-5.3
$15,001 to $30,000	356	17.4	1.09	62	17.6	2.39	-0.2	2.63	-0.9
$30,001 to $50,000	446	17.1	0.82	50	17.6	2.77	-0.5	2.89	-2.9
More than $50,000	1,580	51.0	0.82	128	49.5	2.92	1.5	3.03	3.0

See notes at end of table.

Table 4-8. **Estimates of overall bias for various characteristics from the School Readiness Survey: NHES:2007 RDD and NHES:2007 Bias Study —Continued**

Characteristic	All RDD sample SR respondents			All Bias Study sample SR respondents			Bias[1]		Relative bias[2]
	Sample size	Estimate (percent)	Standard error	Sample size	Estimate (percent)	Standard error	Estimate (percent)	Standard error	Estimate (percent)
Family structure									
Mother and father	2,192	78.7	0.96	219	73.4	4.36	5.3	4.46	7.2
Mother	346	17.2	0.92	60	22.4	3.09	-5.2	3.23	-23.1
Father	40	1.4	0.28	6!	1.9!	1.18!	-0.5	1.21	-25.5
Nonparent guardian(s)	55	2.6	0.54	7	2.3	0.96	0.3	1.10	11.6

\# Rounds to zero.

! Interpret with caution.

‡ Reporting standards not met.

[1] Bias is estimated as the difference between the "all RDD sample SR respondents" estimate and the "all Bias Study sample SR respondents" estimate.

[2] Relative bias is estimated as the bias estimate divided by the "all Bias Study sample SR respondents" estimate, and is expressed as a percentage.

[3] Told child a story; taught child letters, words, or numbers; taught child songs or music; did arts and crafts with child; played sports, active games or exercised together; and played board games or did puzzles with child.

[4] Any three or more of the following: Visited a library; visited a bookstore; went to a play, concert, or other live show; visited an art gallery, museum, or historical site; visited a zoo or aquarium; attended an event sponsored by a community, religious, or ethnic group; and attended an athletic or sporting event (outside of school) in which the child was not a player.

NOTE: Shading indicates a significant difference of 5 percentage points or more.

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Readiness (SR) Survey of the National Household Education Surveys Program, 2007.

Table 4-9. Estimates of overall bias for various characteristics from the Parent and Family Involvement Survey: NHES:2007 RDD and NHES:2007 Bias Study

| Characteristic | All RDD sample PFI respondents | | | All Bias Study sample PFI respondents | | | Bias[1] | | Relative bias[2] |
	Sample size	Estimate (percent)	Standard error	Sample size	Estimate (percent)	Standard error	Estimate (percent)	Standard error	Estimate (percent)
Parents participate in 5 or more activities in the student's school[3]	5,576	53.0	0.62	566	52.1	2.22	0.9	2.30	1.7
Parents report school provides information very well									
About how student is doing in school	6,385	60.9	0.76	668	58.2	2.20	2.7	2.33	4.6
About how to help student with his/her homework	4,740	46.6	0.65	525	47.9	2.07	-1.3	2.17	-2.7
About why student is placed in particular groups or classes	4,481	44.5	0.65	496	45.6	1.67	-1.1	1.79	-2.4
About how to help student plan for college or vocational school	2,064	34.0	0.86	237	38.9	1.65	-4.9	1.86	-12.6
About the family's expected role at student's school	5,024	48.4	0.72	539	46.9	1.93	1.5	2.06	3.2
Parent reports being very satisfied with 4 or more aspects of the student's school[4]	7,263	69.8	0.62	798	73.7	1.87	-3.9	1.97	-5.3
Parents participated in 5 or more home learning activities[5]	2,213	47.0	1.02	229	42.2	2.28	4.8	2.50	11.4

See notes at end of table.

Table 4-9. **Estimates of overall bias for various characteristics from the Parent and Family Involvement Survey: NHES:2007 RDD and NHES:2007 Bias Study—Continued**

Characteristic	All RDD sample PFI respondents			All Bias Study sample PFI respondents			Bias[1]		Relative bias[2]
	Sample size	Estimate (percent)	Standard error	Sample size	Estimate (percent)	Standard error	Estimate (percent)	Standard error	Estimate (percent)
Parents took 3 or more outings with student in the past month[6]	5,321	49.5	0.77	559	50.3	2.09	-0.8	2.23	-1.6
Parents check to see that student's homework gets done	8,190	85.4	0.46	885	87.9	0.99	-2.5	1.09	-2.8
Parents received information about free tutoring	4,552	43.9	0.65	465	42.6	2.15	1.3	2.25	3.1
Parent expects student to earn a college degree or higher	4,360	69.6	1.02	424	67.2	2.87	2.4	3.05	3.6
Family plans to help pay for student's education after high school	4,700	81.3	0.84	475	83.1	2.19	-1.8	2.35	-2.2
Student participated in school activities	5,965	56.0	0.76	624	55.8	1.89	0.2	2.04	0.4
Student has a disability	2,463	23.9	0.68	257	23.0	2.16	0.9	2.26	3.9
Student's sex									
Male	5,498	51.8	0.74	575	51.8	2.10	#	2.23	#
Female	5,183	48.2	0.74	548	48.2	2.10	#	2.23	#
Home tenure									
Own	8,438	70.0	0.24	837	69.7	0.80	0.3	0.84	0.4
Rent/other	2,243	30.0	0.24	286	30.3	0.80	-0.3	0.84	-1.0

See notes at end of table.

Table 4-9. **Estimates of overall bias for various characteristics from the Parent and Family Involvement Survey: NHES:2007 RDD and NHES:2007 Bias Study—Continued**

Characteristic	All RDD sample PFI respondents			All Bias Study sample PFI respondents			Bias[1]		Relative bias[2]
	Sample size	Estimate (percent)	Standard error	Sample size	Estimate (percent)	Standard error	Estimate (percent)	Standard error	Estimate (percent)
Parents' educational attainment									
High school diploma or below	2,578	27.8	0.56	326	31.7	2.31	-3.9	2.38	-12.3
Beyond high school diploma	8,103	72.2	0.56	797	68.3	2.31	3.9	2.38	5.7
Parents' language									
Both/only parent(s) speak(s)									
English	9,437	88.5	0.34	1012	90.9	1.29	-2.4	1.33	-2.6
One of two parents speaks									
English	159	1.4	0.16	22	2.1	0.61	-0.7	0.63	-33.3
No parent speaks English	1,085	10.2	0.32	66	7.0	1.17	3.2	1.21	45.7
Mothers' employment status									
35 hours or more per week	4,993	44.2	0.66	521	47.0	2.12	-2.8	2.22	-6.0
Less than 35 hours per week	2,290	21.8	0.44	260	21.1	1.68	0.7	1.74	3.3
Looking for work	393	4.6	0.33	45	4.1	0.56	0.5	0.65	12.2
Not in labor force	2,611	26.1	0.59	241	23.4	1.24	2.7	1.37	11.5
No mother in household	394	3.3	0.21	56	4.4	0.55	-1.1	0.59	-25.0
Poverty status									
Poor	1,291	19.2	0.33	221	19.7	1.05	-0.5	1.10	-2.5
Nonpoor	9,390	80.8	0.33	902	80.3	1.05	0.5	1.10	0.6
Household income									
Less than $15,000	824	12.2	0.32	153	12.9	0.64	-0.7	0.72	-5.4
$15,001 to $30,000	1,321	15.0	0.40	179	15.2	1.22	-0.2	1.28	-1.3
$30,001 to $50,000	1,799	16.4	0.45	171	16.6	2.23	-0.2	2.27	-1.2
More than $50,000	6,737	56.4	0.46	620	55.3	2.22	1.1	2.27	2.0

See notes at end of table.

Table 4-9. Estimates of overall bias for various characteristics from the Parent and Family Involvement Survey: NHES:2007 RDD and NHES:2007 Bias Study—Continued

Characteristic	All RDD sample PFI respondents			All Bias Study sample PFI respondents			Bias[1]		Relative bias[2]
	Sample size	Estimate (percent)	Standard error	Sample size	Estimate (percent)	Standard error	Estimate (percent)	Standard error	Estimate (percent)
Family structure									
Mother and father	7,995	72.7	0.54	823	74.1	1.88	-1.4	1.96	-1.9
Mother	1,876	20.0	0.55	204	18.6	1.61	1.4	1.70	7.5
Father	356	3.0	0.21	54	4.0	0.42	-1.0	0.47	-25.0
Nonparent guardian(s)	454	4.3	0.39	42	3.3	0.55	1.0	0.67	30.3
School type									
Public	8,978	88.2	0.45	989	89.9	1.65	-1.7	1.71	-1.9
Private	1,392	11.8	0.45	114	10.1	1.65	1.7	1.71	16.8
School size									
Under 300	1,480	14.9	0.57	167	14.6	1.51	0.3	1.61	2.1
300-599	3,142	31.3	0.69	321	28.8	2.59	2.5	2.68	8.7
600-999	2,756	26.7	0.60	286	25.8	1.81	0.9	1.91	3.5
1,000 or more	2,910	27.0	0.52	323	30.8	1.65	-3.8	1.73	-12.3

\# Rounds to zero.

[1] Bias is estimated as the difference between the "all RDD sample PFI respondents" estimate and the "all Bias Study sample PFI respondents" estimate.

[2] Relative bias is estimated as the bias estimate divided by the "all Bias Study sample PFI respondents" estimate, and is expressed as a percentage.

[3] Any 5 or more of the following: Attended a general school meeting; attended a meeting of the parent-teacher organization or association; went to a regularly scheduled parent-teacher conference with the student's teacher; attended a school or class event because of the student; served as a volunteer in the student's classroom or elsewhere in the school; participated in fundraising for the school; served on a school committee; and met with a guidance counselor in person.

[4] Any 4 or more of the following: School student attends this year; teachers student has this year; academic standards of the school; order and discipline at the school; and way that school staff interact with parents.

[5] Any 5 or more of the following: Told student a story; did arts and crafts with student; played sports, active games or exercised together; worked on projects such as building, making, or fixing something with student not as a chore; talked with student about family history or ethnic heritage; and played board games or did puzzles with student..

[6] Any 3 or more of the following: Visited a library; visited a bookstore; went to a play, concert, or other live show; visited an art gallery, museum, or historical site; visited a zoo or aquarium; attended an event sponsored by a community, religious, or ethnic group; and attended an athletic or sporting event (outside of school) in which the child was not a player.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Parent and Family Involvement in Education (PFI) Survey of the National Household Education Surveys Program, 2007.

5. DATA PREPARATION

5.1 Disclosure Risk Analysis

Central to the mission of the National Center for Education Statistics (NCES) is a commitment to protecting the identity of respondents to its various data collections. Surveys that make up the National Household Education Surveys Program (NHES) are designed to protect respondent identity. This design includes an extensive respondent disclosure risk analysis. As in past NHES collections, results from this analysis led to modifications to some data included on the data files. The modifications included coarsening of response categories and swapping of certain data items between respondents. These confidentiality edits modify respondent data in order to prevent positive identification of individual respondents. Tests on the modified data were conducted to assure that the data remain accurate and useful.

Under law, data collected and distributed by NCES may be used only for statistical purposes.

Any effort to determine the identity of any reported case by data users is prohibited by law. Violations are subject to Class E felony charges of a fine up to $250,000, a prison term up to 5 years, or both. NCES does all it can to assure that the identity of data subjects cannot be disclosed. All direct identifiers, as well as any characteristics that might lead to identification, are omitted or modified in the public-use dataset to protect the true characteristics of individuals. Any intentional identification or disclosure of a person violates the assurances of confidentiality given to the providers of the information. Therefore, users must adhere to the following:

- Use the data in this dataset for statistical purposes only.

- Make no use of the identity of any person discovered inadvertently, and advise NCES of any such discovery.

- Not link this dataset with individually identifiable data from other NCES or non-NCES datasets.

- Signify their agreement to comply with the above-stated statutorily based requirements to proceed.

5.2 Coding and Editing Specifications

Most of the NHES:2007 interview data were coded by the interviewers during the interview using the computer-assisted telephone interviewing (CATI) system. As the interviewers entered the number of the response option given by the respondent, this number was written to the data file. Range and logic edits were developed for relevant items to maximize coding accuracy.

5.2.1 Range Specifications

The ranges of most of the items were determined by the codes available for responses (closed-ended responses). However, some items such as age did not have predefined response codes and required

an entry by the interviewer (open-ended responses). To help assure that reasonable entries were made for open-ended responses, reasonable ranges were defined prior to data collection and programmed into CATI.

Range edits included both hard- and soft-range edits. A "soft range" is one that represents the reasonable expected range of values but does not include all possible values. Responses outside the soft range were confirmed with the respondent and had to be entered a second time. For example, the number of hours each week a child spent doing homework had a soft range of 1 to 14. A value outside this range could be entered and confirmed as correct by the interviewer as long as it was within the hard range of values (1 to 36). "Hard ranges" are those that have a finite set of parameters for the values that can be entered into the CATI system. Out-of-hard-range values for either open- or closed-ended questions were not accepted. If the respondent insisted that a response outside the hard range was correct, the interviewer could enter the information in a comments data file. Out-of-hard-range values were accepted if the comments supported the response. Otherwise, the values were left as missing and later imputed.

After data collection was completed and imputation was performed, range edits for number-and-unit logic checks and hard-range-by-unit checks were rerun against the entire database to ensure that no outliers were inadvertently introduced during the post-data-collection updating process or during imputation. In addition, staff reviewed all continuous variable ranges. Therefore, any outliers that exist in the data files were reviewed during the data preparation process and originated from information entered into the comments data file.

5.2.2   Consistency Checks (Logic Edits)

Consistency or logic checks examine the relationships between responses to ensure that they do not conflict with one another or that the response to one item does not make the response to another unlikely. Logic specifications for the NHES:2007 interviews were contained within the CATI system. For example, the CATI system was programmed to control skip patterns so that inappropriate items were not asked. Other types of consistency (logic) checks for the NHES:2007 interviews also were included. For example, a parent/guardian may have reported that a child was attending a grade that was outside the normal range of grades for his age. If the logic check was violated, an error message appeared that explained that the response was inconsistent and allowed the interviewer to enter a correction. If the respondent confirmed an answer that appeared to be inconsistent, the interviewer entered it as a comment. The values and interviewer comments for cases violating logic edits were also checked against the entire database. Cases violating the edits were examined by data preparation and project staff and either the information violating the edit was kept or it was coded to "not ascertained" and later replaced with imputed data. Data were kept in circumstances where the data were judged to be plausible even though they violated the edit (e.g., an inconsistency between a child's age and his/her grade in school existed because the parent respondent indicated that the child had been accelerated in school).

5.2.3   Structural Edits

To facilitate imputation, person-level data collected in the Screener was structured vertically, one record per enumerated household member. SAS structural edits were run prior to and after imputation to ensure that appropriate person records existed for responses gathered during the interview. For example, if a birth father was indicated as living in the household, the structural edits checked for a

person record with sex equal to male and an age 12 or more years older than the child. Structural edits also checked interview completeness and parent relationship data.

5.2.4   Frequency and Cross-Tabulation Review

The frequencies of responses to all data items (both individually and in conjunction with related data items) were reviewed to ensure that appropriate skip patterns were followed and that inappropriate values were not introduced during data editing. Staff members checked each item to make sure the correct numbers of responses and legitimate skips were found. If a discrepancy was discovered, the problem case was identified and reviewed. If necessary, the audit trail for the interview, which provided a keystroke-by-keystroke record of an interview, was retrieved to determine the appropriate response. If the audit trail revealed no additional information, the item was coded as "not ascertained" and later imputed.

5.2.5   Review of "Other, specify" Items

Most "other, specify" text responses were reviewed to determine if they should be coded into one of the existing code categories. When a respondent gave a response other than those that were available, the interviewer entered the respondent's text response into a "specify" overlay that appeared on the screen. The "specify" responses were reviewed by the data preparation staff and, where appropriate, coded into one of the existing response categories. Review of the open-ended text responses revealed that with few exceptions, no particular text item occurred frequently enough to warrant the creation of a new response category. However, some additions were made to exiting categories in item (PN20) and (PN21) concerning television channels. Specifically, additional religious channels were added to the category including Christian/Cornerstone so that Trinity Broadcasting and others could be included there. Due to the high frequency of certain Spanish-language channels, they were added to the category including Telemundo and Galavision, so that MUN2 and others could be included there. All channels with the word "family" were included in the category for the ABC Family Channel and all sports-related channels were included in a sports channel category (previously ESPN). Additionally, a number of local channels were investigated and found to be Public Broadcasting Stations (PBS) and were included in the same category. Due to the high frequency of certain specified responses, two new variables were added to the list of television channels: The History Channel, and Food (or Cooking) Network. The response categories and open-ended items that were added appear in italics on the questionnaires. Verbatim strings of "other, specify" items appear only on the restricted-use data files. See chapter 6 of Volume II and Volume III for a discussion of the contents of both the public-use and restricted-use data files.

5.2.6   Coding Schools

During the PFI interview, parents were asked to provide the name and location of their child's school, and the interviewers used a lookup file to identify the school. When the school was identified, a school identification number was entered into the CATI record. If interviewers could not find the school in the lookup file during the interview, the parent was asked for the name and address of the school and other questions such as highest and lowest grade and the religious affiliation of a private school.

An attempt was made by study staff to identify the school using the information that was provided by the parent. Internet searches were conducted in an effort to identify schools; it was sometimes found that a school was located on a street named by the respondent but that the school name was different from the one provided by the respondent. In addition, using the online school lookup function for the Common Core of Data (CCD) at the U.S. Department of Education's National Center for Education Statistics, school identification numbers were identified for a number of schools (http://nces.ed.gov/ccd/schoolsearch/). For private schools, the lookup function for the Private School Universe Survey, also available through the U.S. Dept of Education's National Center for Education Statistics web site, offered a way to search for private schools online (http://nces.ed.gov/surveys/pss/).

If no NCES school identification number could be obtained through these methods, then the school was imputed using the procedures described in chapter 3.

REFERENCES

Blumberg, S.J., Luke, J.V., and Cynamon, M.L. (2006). Telephone Coverage and Health Survey Estimates: Evaluating the Need for Concern about Wireless Substitution. *American Journal of Public Health*, 96(5), 926-931.

Brick, J.M. (1996*). Undercoverage Bias in Estimates of Adults and 0- to 2-Year-Olds in the 1995 National Household Education Survey (NHES:95*).   (NCES Publication No. 92–101).   U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Brick, J.M., Burke, J., and West, J. (1992). *Telephone Undercoverage Bias of 14- to 21-Year-Olds and 3- to 5-Year-Olds.* (NCES Publication No. 92–101). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Brick, J.M., Celebuski, C.A., Collins, M.A., and West, J. (1992). *Overview of the NHES Field Test.* Technical Report No. 1. (NCES Publication No. 92–099). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Brick, J.M., Collins, M.A., and Chandler, K. (1997). *An Experiment in Random Digit Dial Screening.* (NCES Publication No. 98-255). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Brick, J.M.; Hagedorn, M.C.; Montaquila, J.; Roth, S.B., and Chapman, C. (2006). *Impact of Monetary Incentives and Mailing Procedures: An Experiment in a Federally Sponsored Telephone Survey* (NCES 2006-066). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Brick, J.M., Tubbs, E., Collins, M.A., Nolin, M.J., Cantor, D., Levin, K., and Carnes, Y. (1997). *Telephone Coverage Bias and Recorded Interviews in the 1993 National Household Education Survey. (NHES:93).* (NCES Publication No. 96–029). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Brick, J.M., Waksberg, J., Kulp, D., and Starer, A. (1995). "Bias in List-Assisted Telephone Samples." *Public Opinion Quarterly*, 59(2): 218–235.

Broene, P. and Rust, K. (2000). *Strengths and Limitations of Using SUDAN, Stata, and WesVar PC for Computing Variance from NCES Data Sets.* (NCES Publication No. 2000-2003). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Casady, R.J., and Lepkowski, J.M. (1993). "Stratified Telephone Survey Designs." *Survey Methodology*, 19(1), 103–113.

Deming, W.E., and Stephan, F.F. (1940). "On a Least Square Adjustment of a Sampled Frequency Table When the Expected Marginal Totals Are Known." *Annals of Mathematical Statistics*, 11, 427–444.

Kalton, G., and Kasprzyk, D. (1986). "The Treatment of Missing Survey Data." *Survey Methodology*, 12 (1), 1–16.

Montaquila, J.M., Brick, J.M., and Brock, S.P. (1997). *Undercoverage Bias in Estimates of Characteristics of Households and Adults in the 1996 National Household Education Survey*. (NCES 97–39). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Rao, J.N.K., and Shao, J. (1992). "Jackknife Variance Estimation with Survey Data Under Hot Deck Imputation." *Biometrika*, 79, 811–822.

Rubin, D.R. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons.

Shah, B.V., Barnwell, B.G., Hunt, P.N., and LaVange, L.M. (1995). *SUDAAN User's Manual*. Research Triangle Park, NC: Research Triangle Institute.

Shapiro, G., Battaglia, M., Camburn, D., Massey, J., and Tompkins, L. (1995). "Calling Local Telephone Company Business Offices to Determine the Residential Status of a Wide Class of Unresolved Telephone Numbers in a Random-Digit-Dialing Sample." *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 975–980.

Tucker, C., Lepkowski, J.M., and Piekarski L. (2002). "The Current Efficiency of List-Assisted Telephone Sampling Designs." *Public Opinion Quarterly*, 66, 321-338.

Wolter, K. (1985). *Introduction to Variance Estimation*. New York: Springer-Verlag, Chapter 4.

APPENDIX A

**NHES:2007 Screener, School Readiness and
Parent and Family Involvement
in Education Questionnaire**

This page is intentionally blank.

APPENDIX B

Summary of Weighting and
Sample Variance Estimation Variables

This page is intentionally blank.

Table B-1. Summary of weighting and sample variance estimation variables: 1991-2007

NHES data file	Full sample weight	Respondent ID	Replicate weights	Jackknife method	Sample design	Nesting variables	DEFT (Average Root Design Effect) for approximating sampling errors
			Computing sampling errors				
			Replication method (WesVar, SUDAAN, STATA, AM[1])		**Taylor series method (SUDAAN, Stata, SAS 8[2],AM, SPSS Complex Samples)**		
NHES:1991 *Early Childhood Education*		PERSID	EWREPL1 – EWREPL50				
● Primary file	EWGT		EWREPL1 –	JK1	WR	VSTRAT PSU	1.2
● Preprimary file	EWGT		EWREPL50	JK1	WR	VSTRAT PSU	1.2
NHES:1991 *Adult Education*							2.1 Full Sample
● Adult file	AEWT	PERSID	AEREPL1-AEREPL50	JK1	WR	VSTRAT PSU	1.5 Participants
● Course file[3]	AEWT	CLASID	AEREPL1-AEREPL50	JK1	WR	VSTRAT PSU	1.7 Nonparticipants
							2.0 Black (non-Hispanic)
							1.8 Hispanic
							1.7 White (non-Hispanic)
							1.6 Other races
NHES:1993 *School Readiness*	FWGT0	ENUMID	FWGT1 - FWGT60	JK2	WR	STRATUM PSU	1.3
NHES:1993 *School Safety & Discipline*							
● Parent interviews only	FWGT0	BASMID	FWGT1-FWGT60	JK2	WR	STRATUM PSU	1.4
● Parent & Emancipated Youth (EY) interviews	FWGT0 (for parents) & PFWGT0 (for EY)	BASMID	FWGT1-FWGT60, PFWGT1-PFWGT60	JK2	WR		1.4
● Youth interviews (including Emancipated Youth)		ENUMID	FWGT1-FWGT60	JK2	WR	STRATUM PSU	1.5
	FWGT0					STRATUM PSU	
NHES:1995 *Early Childhood Program Participation*	EWEIGHT	ENUMID	ERPL1 - ERPL50	JK1	WR	STRATUM PSU	1.2
NHES:1995 *Adult Education*[4]	AEWEIGHT	BASMID	ARPL1 - ARPL50	JK1	WR	STRATUM PSU	1.3
NHES:1996 *Screener/Household & Library*	FHWT	BASEID	FHWTR1-FHWTR80	JK1	WR	HSTRATUM HPSU	1.1
NHES:1996 *Parent PFI/CI*	FPWT	BASMID	FPWTR1-FPWTR80	JK1	WR	PSTRATUM PPSU	1.3

See notes at end of table.

Table B-1. Summary of weighting and sample variance estimation variables: 1991-2007—Continued

| NHES data file | Full sample weight | Computing sampling errors | | | | | DEFT (Average Root Design Effect) for approximating sampling errors |
| | | Replication method (WesVar, SUDAAN, STATA, AM[1]) | | | Taylor series method (SUDAAN, Stata, SAS 8[2], AM, SPSS Complex Samples) | | |
		Respondent ID	Replicate weights	Jackknife method	Sample design	Nesting variables	
NHES:1996 Youth CI	FYWT	BASMID	FYWTR1-FYWTR80	JK1	WR	YSTRATUM YPSU	1.4
NHES:1996 Adult CI	FAWT	BASMID	FAWTR1-FAWTR80	JK1	WR	ASTRATUM APSU	1.2
NHES:1999 Parent Interview	FPWT	BASMID	FPWT1-FPWT80	JK1	WR	PSTRATUM PPSU	1.3
NHES:1999 Youth Interview	FYWT	BASMID	FYWT1-FYWT80	JK1	WR	YSTRATUM YPSU	1.3
NHES:1999 Adult Education Interview	FAWT	BASMID	FAWT1-FAWT80	JK1	WR	ASTRATUM APSU	1.3 Full sample 1.4 Participants 1.5 Black, non-Hispanic
NHES:2001 Early Childhood Program Participation	FEWT	BASMID	FEWT1-FEWT80	JK1	WR	ESTRATUM EPSU	1.2 Full sample 1.3 Black, non-Hispanic
NHES:2001 Before- and After-School Programs and Activities	FSWT	BASMID	FSWT1-FSWT80	JK1	WR	SSTRATUM SPSU	1.3 Full sample 1.4 Black, non-Hispanic
NHES:2001 Adult Education	FAWT	BASMID	FAWT1-FAWT80	JK1	WR	ASTRATUM APSU	1.3
NHES:2003 Parent and Family Involvement in Education	FPWT	BASMID	FPWT1-FPWT80	JK1	WR	PSTRATUM PPSU	1.3 Full sample 1.4 Race/ethnicity subgroups
NHES:2003 Adult Education for Work-Related Reasons	FAWT	BASMID	FAWT1-FAWT80	JK1	WR	ASTRATUM APSU	1.3 Full sample 1.4 Hispanics 1.4 Work-related adult education participants

See notes at end of table.

Table B-1.   Summary of weighting and sample variance estimation variables: 1991-2007—Continued

| NHES data file | Full sample weight | Computing sampling errors | | | | | DEFT (Average Root Design Effect) for approximating sampling errors |
| | | Replication method (WesVar, SUDAAN, STATA, AM[1]) | | | Taylor series method (SUDAAN, Stata, SAS 8[2], AM, SPSS Complex Samples) | | |
		Respondent ID	Replicate weights	Jackknife method	Sample design	Nesting variables	
NHES:2005 Early Childhood Program Participation	FEWT	BASMID	FEWT1-FEWT80	JK1	WR	ESTRATUM EPSU	1.4 Full sample 1.3 Preschoolers
NHES:2005 After-School Programs and Activities	FSWT	BASMID	FSWT1-FSWT80	JK1	WR	SSTRATUM SPSU	1.4 Full sample 1.3 Home schoolers 1.3 White, non-Hispanic 1.5 Black, non-Hispanic
NHES:2005 Adult Education	FAWT	BASMID	FAWT1-FAWT80	JK1	WR	ASTRATUM APSU	1.6 Full sample 1.5 White, non-Hispanic 1.5 Black, non-Hispanic 1.5 Nonparticipants 1.7 Less than high school 1.4 High school diploma/ equiv. 1.4 Bachelors or higher 1.5 Associates degree
NHES:2007 School Readiness	FRWT	BASMID	FRWT1-FRWT80	JK1	WR	RSTRATUM RPSU	1.4 Full sample 1.5 Preschoolers 1.6 Black, non-Hispanic

See notes at end of table.

Table B-1. Summary of weighting and sample variance estimation variables: 1991-2007—Continued

NHES data file	Full sample weight	Computing sampling errors					DEFT (Average Root Design Effect) for approximating sampling errors
		Replication method (WesVar, SUDAAN, STATA, AM[1])			Taylor series method (SUDAAN, Stata, SAS 8[2], AM, SPSS Complex Samples)		
		Respondent ID	Replicate weights	Jackknife method	Sample design	Nesting variables	
NHES:2007 Parent and Family Involvement in Education	FPWT	BASMID	FPWT1-FPWT80	JK1	WR	PSTRATUM PPSU	1.4 Full sample 1.5 Elementary schoolers 1.5 Middle schoolers 1.5 High schoolers 1.5 Black, non-Hispanic

[1] WesVar Complex Samples software, version 5, is available from Westat (www.westat.com). Information on SUDAAN can be obtained at www.rti.org. SUDAAN performs replication using the JK1 procedure but not the JK2 procedure. Information on Stata can be obtained at www.stata.com. Information on AM can be obtained at www.am.air.org.

[2] Information on SUDAAN can be obtained at www.rti.org. Information on Stata can be obtained at www.stata.com. Additionally, SAS version 9 includes survey procedures that use the Taylor series method for variance estimation. (See www.sas.com.) Information on AM can be obtained at www.am.air.org. Information on SPSS Complex Samples can be obtained at www.spss.com/complex_samples.

[3] Unlike the NHES:1995 Adult Education data file, no course weights are provided in the NHES:1991 course file. The full sample weight and variables for computing sampling errors are provided in the course file for making adult-level estimates. Information as to the total number of courses that adults took is also available, and procedures similar to those described in the NHES:1995 *Adult Education Data File User's Manual* (Collins et al. 1996) could be used to create weights for making course-related estimates. However, it is important to note that the course information collected in the NHES:1991 pertains to the four most recent courses taken, rather than a random sample of courses as was the case in the NHES:1995.

[4] This data file contains weights for making "person-course" estimates pertaining to work-related and other formal structured courses. A simple way of doing this is to create a new variable that is the product of the course weight and the variable of interest. The standard weight and variance estimation methods are then applied to the new variable. The weight variables are called WRWGT, for adjusting for the courses adults took in work-related classes, and SAWGT, for adjusting for personal development courses. Weights are required for these types of courses because course-related data were collected only for a random subsample of courses. See the NHES:1995 *Adult Education Data File User's Manual* (Collins et al. 1996) for more details.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Surveys Program (NHES), 1991-2007.

APPENDIX C

Reasonableness Tables

This page is intentionally blank.

Table C-1. Percentage distribution for age of subjects of interviews: SR-NHES:2007, PFI-NHES:2007, and CPS:2005

Age category	SR-NHES:2007 and PFI-NHES:2007[1]		CPS:2005	
	Percent	s.e.	Percent	s.e.
3 – 5 years	4	#	4	0.1
6 – 9 years	6	0.1	6	0.1
10 – 15 years	9	#	9	0.1
16 – 19 years	6	0.3	6	0.1

Rounds to zero.

[1] Estimates of children age 3 through 6 and not yet enrolled in kindergarten were obtained from the School Readiness (SR) Survey. Estimates of children/youth age 3 through 19 and enrolled in kindergarten through grade 12 were obtained from the Parent and Family Involvement in Education (PFI) Survey. Estimates for the 16–19 years age category include adults who were not in grades 12 or below who were interviewed for the Adult Education for Work-Related Reasons (AEWR) Survey.

NOTE: s.e. is standard error.

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Readiness Survey of the National Household Education Surveys Program (NHES), 2007 and Parent and Family Involvement in Education Survey of the NHES, 2007. U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 2005.

Table C-2A. Percentage distribution of children ages 3 through 20 not enrolled in school or enrolled in kindergarten through grade 12: SR-NHES:2007 and PFI-NHES:2007

Child's age	Number of children (thousands)	Child's current grade														
		U	N	K	1	2	3	4	5	6	7	8	9	10	11	12
3	3,993	59	39	3												
4	3,854	27	66	7												
5	3,674	4	9	81	5											
6	3,829			14	80	5										
7	4,103				20	73	6	1								
8	3,871				1	18	78	3								
9	3,791					1	15	77	7							
10	4,043						1	19	75	5						
11	3,947							1	17	74	6	1				
12	4,013								1	19	73	7				
13	4,201									4	18	71	6			
14	4,245										4	21	68	7		
15	4,323											1	22	70	6	
16	4,530												3	21	71	6
17	3,811													3	23	73
18	802														5	94
19	93														4	96
20	26															100

NOTE: For the National Household Education Surveys Program (NHES) kindergarten (K) includes grades classified as kindergarten, transitional kindergarten, and prefirst grade. Age in NHES:2007 was recalculated to match the Current Population Survey definition of the child's age as of September 30. Homeschoolers are excluded from the NHES estimates. Because of rounding, percentages may not add to 100. Blank cells in the table represent estimates that round to zero.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Readiness (SR) Survey of the National Household Education Surveys Program (NHES), 2007; Parent and Family Involvement in Education Survey of the NHES, 2007.

Table C-2B. **Standard errors of the percentage distribution of children ages 3 through 20 not enrolled in school or enrolled in kindergarten through grade 12: SR-NHES:2007 and PFI-NHES:2007**

Child's age	Number of children (thousands)	Child's current grade														
		U	N	K	1	2	3	4	5	6	7	8	9	10	11	12
3	138	1.6	1.6	1.6												
4	132	1.8	1.8	1.1												
5	117	0.9	1.1	1.6	1.4											
6	174			1.9	2.1	1.0										
7	179				3.0	2.8	1.1	0.6								
8	125				0.4	1.7	1.6	0.7								
9	130					0.3	2.2	2.2	1.3							
10	109						0.5	1.6	1.6	0.7						
11	116							0.5	1.5	1.6	1.0	0.4				
12	124								0.4	1.8	1.9	1.4				
13	137									1.9	1.7	2.5	1.1			
14	118										0.9	1.4	1.8	1.0		
15	131											0.5	1.7	1.8	1.0	
16	138												0.7	1.7	1.9	1.1
17	140													1.5	2.0	2.0
18	110														1.9	2.0
19	27														4.1	4.1
20	13															

NOTE: Standard errors increase for children who are 18, 19, and 20 years old. This is because there are small numbers of those children in the grade categories shown above. Blank cells in the table represent estimates that round to zero.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Readiness (SR) Survey of the National Household Education Surveys Program (NHES), 2007; Parent and Family Involvement in Education Survey of the NHES, 2007.

Table C-2C.   Percentage distribution of children ages 3 through 20 not enrolled in school or enrolled in kindergarten through grade 12: CPS:2005

Child's age	Number of children (thousands)	U	N	K	1	2	3	4	5	6	7	8	9	10	11	12
								Child's current grade								
3	4,151	59	40	2												
4	4,028	34	59	7												
5	3,955	7	13	74	6	1										
6	3,889	3	2	16	73	6	1									
7	3,875			1	21	72	5	1								
8	3,904				3	19	71	6	1							
9	3,849				4	2	20	67	5	1						
10	4,005					1	3	20	70	5	1					
11	3,979						1	3	22	67	6	1				
12	3,993							1	2	24	65	6	2			
13	4,331								1	4	24	65	6			
14	4,175									1	3	24	67	5		
15	4,184										1	3	22	66	7	1
16	4,443												5	27	62	6
17	3,864												1	5	29	65
18	1,137												1	2	14	82
19	246										2		3	8	18	70
20	77												9		32	59

NOTE: Homeschoolers are included in the Current Population Survey estimates. Because of rounding, percentages may not add to 100. Blank cells in the table represent estimates that round to zero.
SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 2005.

Table C-2D. Standard errors of the percentage distribution of children ages 3 through 20 not enrolled in school or enrolled in kindergarten through grade 12: CPS:2005

Child's age	Number of children (thousands)	U	N	K	1	2	3	4	5	6	7	8	9	10	11	12
3	4,151	1.3	1.3	0.3												
4	4,028	1.3	1.3	0.7												
5	3,955	0.7	0.9	1.2	0.7	0.3										
6	3,889	0.4	0.4	1.0	1.2	0.7	0.3									
7	3,875			0.3	1.1	1.2	0.6	0.3								
8	3,904				0.5	1.1	1.3	0.6	0.3							
9	3,849				0.5	0.4	1.1	1.3	0.6	0.3						
10	4,005					0.2	0.5	1.1	1.3	0.6	0.3					
11	3,979						0.2	0.5	1.1	1.3	0.7	0.3				
12	3,993							0.3	0.4	1.2	1.3	0.6	0.3			
13	4,331								0.3	0.5	1.1	1.2	0.6			
14	4,175									0.3	0.5	1.1	1.3	0.6		
15	4,184										0.2	0.5	1.1	1.3	0.7	0.2
16	4,443												0.5	1.1	1.3	0.6
17	3,864												0.3	0.6	1.3	1.3
18	1,137												0.5	0.8	1.8	2.0
19	246										1.5		1.8	3.0	4.2	5.1
20	77												5.7		9.2	9.7

NOTE: Blank cells in the table represent estimates that round to zero.
SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 2005.

Table C-3. Number of children enrolled in kindergarten through grade 12, by school type and by student grade level: PFI-NHES:2007 and CPS:2005

School type and grade	PFI-NHES:2007 Number (thousands)	s.e. (thousands)	CPS:2005 Number (thousands)	s.e. (thousands)
Total number of children in kindergarten through 12th grade	53,186	0	53,328	330
School type[1]				
Public	45,517	269	48,018	320
Private	6,082	232	5,309	124
Homeschooled	1,586	118	—	—
Student grade level				
K	3,902	0	3,912	107
1	4,135	0	4,146	110
2	3,918	0	3,928	107
3	3,915	0	3,925	107
4	3,850	0	3,860	106
5	4,047	0	4,058	109
6	4,053	0	4,064	109
7	4,143	0	4,154	110
8	4,240	0	4,251	111
9	4,272	0	4,283	112
10	4,365	0	4,376	113
11	4,415	0	4,427	113
12	3,932	0	3,942	107

— Not available.

[1] The Current Population Survey did not identify homeschoolers.

NOTE: s.e. is standard error. Because the standard error of an estimate is a measure of sampling error variance, a standard error of zero indicates the absence of sampling error variance. When the NHES estimates of totals are adjusted to exactly match CPS totals, all sampling error in those estimated totals is eliminated, under the assumption that the CPS total is the true population value. Because of rounding, details may not add to totals.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Parent and Family Involvement in Education Survey of the National Household Education Surveys Program, 2007; and U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 2005.

Table C-4. Number and percentage of children in kindergarten through grade 12 enrolled in public and private schools: PFI-NHES:2007 and CPS:2005

Child's current grade	School type					
	Public			Private		
	Number (thousands)	Percent	Percent s.e.	Number (thousands)	Percent	Percent s.e.
PFI-NHES:2007						
K	3,078	81	2.2	699	19	2.2
1	3,394	86	1.7	563	14	1.7
2	3,329	88	1.7	453	12	1.7
3	3,360	88	1.6	458	12	1.6
4	3,386	90	1.3	359	10	1.3
5	3,502	89	1.7	455	11	1.7
6	3,445	88	2.3	448	12	2.3
7	3,587	89	1.3	435	11	1.3
8	3,657	88	1.5	486	12	1.5
9	3,708	90	1.2	416	10	1.2
10	3,813	89	1.2	450	11	1.2
11	3,805	89	1.8	488	11	1.8
12	3,454	90	1.1	372	10	1.1
CPS:2005						
K	3,349	86	1.0	563	14	1.0
1	3,663	88	0.9	483	12	0.9
2	3,490	89	0.9	438	11	0.9
3	3,555	91	0.8	370	9	0.8
4	3,475	90	0.8	385	10	0.8
5	3,619	89	0.8	439	11	0.8
6	3,651	90	0.8	413	10	0.8
7	3,738	90	0.8	416	10	0.8
8	3,836	90	0.8	415	10	0.8
9	3,906	91	0.7	377	9	0.7
10	4,061	93	0.7	315	7	0.7
11	4,016	91	0.8	411	9	0.8
12	3,659	93	0.7	284	7	0.7

NOTE: s.e. is standard error. For the National Household Education Surveys Program: 2007, kindergarten (K) includes grades reported as kindergarten, transitional kindergarten, and prefirst grade. NHES:2007 estimates exclude children who are homeschooled. The Current Population Survey did not identify homeschool.ers
SOURCE: U.S. Department of Education, National Center for Education Statistics, Parent and Family Involvement in Education Survey of the National Household Education Surveys Program, 2003; and U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 2005.

Table C-5.   Percentage of children in ages 3 through 6 and not enrolled in school, by household income: SR-NHES:2007 and CPS:2005

Household income	SR-NHES:2007		CPS:2005	
	Percent	s.e.	Percent	s.e.
$5,000 or less	4	0.7	4	0.4
$5,001 to $10,000	4	0.6	6	0.5
$10,001 to $15,000	7	0.8	5	0.4
$15,001 to $20,000	4	0.6	7	0.5
$20,001 to $25,000	8	0.9	6	0.5
$25,001 to $30,000	5	0.7	6	0.5
$30,001 to $35,000	4	0.6	5	0.5
$35,001 to $40,000	6	0.7	9	0.6
$40,001 to $50,000	7	0.6	5	0.5
$50,001 to $60,000	10	0.8	9	0.6
$60,001 to $75,000	12	0.7	11	0.6
Over $75,000	30	0.8	26	0.9

NOTE: s.e. is standard error. Current Population Survey estimates exclude cases with missing income data. Because of rounding, percentages may not add to 100.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Readiness Survey of the National Household Education Surveys Program, 2007; and U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 2005.

Table C-6. **Number and percentage of children ages 3 through 6 and not enrolled in school, by household income and race/ethnicity: SR-NHES:2007 and CPS:2005**

Race/ethnicity	Number of children (thousands)	Household income							
		Less than $15,000		$15,001 to $30,000		$30,001 to $50,000		More than $50,000	
		Percent	s.e.	Percent	s.e.	Percent	s.e.	Percent	s.e.
SR-NHES:2007									
White, non-Hispanic	4,680	6	0.9	11	1.3	16	1.3	66	1.6
Black, non-Hispanic	1,319	30	4.1	23	4.4	17	2.7	30	3.0
Hispanic	1,919	25	1.8	28	2.7	19	1.8	28	2.1
Other	817	12	3.0	18	4.8	17	3.3	53	3.9
CPS:2005									
White, non-Hispanic	4,882	8	0.7	11	0.8	21	1.1	60	1.3
Black, non-Hispanic	1,323	34	2.5	22	2.2	20	2.1	23	2.2
Hispanic	1,924	20	1.7	32	2.0	22	1.8	26	1.9
Other	629	12	2.5	17	2.8	18	2.9	53	3.8

NOTE: s.e. is standard error. Current Population Survey percentage estimates exclude cases with missing income data. Because of rounding, percentages may not add to 100.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Readiness Survey of the National Household Education Surveys Program, 2007; and U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 2005.

Table C-7. Percentage of children in kindergarten through grade 12, by household income: PFI-NHES:2007 and CPS:2005

Household income	PFI-NHES:2007		CPS:2005	
	Percent	s.e.	Percent	s.e.
$5,000 or less	3	0.3	3	0.1
$5,001 to $10,000	4	0.3	5	0.2
$10,001 to $15,000	5	0.3	4	0.2
$15,001 to $20,000	5	0.3	6	0.2
$20,001 to $25,000	6	0.3	6	0.2
$25,001 to $30,000	4	0.3	6	0.2
$30,001 to $35,000	4	0.3	5	0.2
$35,001 to $40,000	5	0.2	9	0.2
$40,001 to $50,000	8	0.3	4	0.2
$50,001 to $60,000	8	0.3	9	0.2
$60,001 to $75,000	12	0.4	11	0.3
Over $75,000	36	0.4	30	0.4

NOTE: s.e. is standard error. Current Population Survey estimates exclude cases with missing income data. Because of rounding, percentages may not add to 100.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Parent and Family Involvement in Education Survey of the National Household Education Surveys Program, 2007; and U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 2005.

Table C-8.   Number and percentage of children in kindergarten through grade 12, by household income and race/ethnicity: PFI-NHES:2007 and CPS:2005

Race/ethnicity	Number of children (thousands)	Household income							
		Less than $15,000		$15,001 to $ 30,000		$30,001 to $50,000		More than $50,000	
		Percent	s.e.	Percent	s.e.	Percent	s.e.	Percent	s.e.
PFI-NHES:2007									
White, non-Hispanic	31,035	6	0.3	9	0.4	14	0.6	70	0.6
Black, non-Hispanic	7,898	29	1.3	23	1.4	18	1.0	30	1.3
Hispanic	9,929	18	0.9	27	1.3	22	1.2	33	1.4
Other	4,324	12	2.2	14	1.5	15	1.8	58	2.5
CPS:2005									
White, non-Hispanic	31,689	6	0.3	10	0.3	19	0.4	64	0.5
Black, non-Hispanic	7,919	31	1.0	23	0.9	20	0.9	26	0.9
Hispanic	9,955	20	0.7	27	0.8	24	0.8	29	0.8
Other	3,765	11	1.0	15	1.1	20	1.3	53	1.6

NOTE: s.e. is standard error. Current Population Survey percentage estimates exclude cases with missing income data. Because of rounding, percentages may not add to 100.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Parent and Family Involvement in Education Survey of the National Household Education Surveys Program, 2003; and U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 2005.

Table C-9. Number and percentage of children enrolled in kindergarten through grade 12 in public and private schools, by race/ethnicity: PFI-NHES:2007 and CPS:2005

| Race/ethnicity | PFI-NHES:2007 | | | | | CPS:2005 | | | | |
	Number of children (thousands)	Public Percent	s.e.	Private Percent	s.e.	Number of children (thousands)	Public Percent	s.e.	Private Percent	s.e.
White, non-Hispanic	31,035	85	0.6	15	0.6	31,689	87	0.3	13	0.3
Black, non-Hispanic	7,898	92	1.3	8	1.3	7,919	94	0.5	6	0.5
Hispanic	9,929	93	0.6	7	0.6	9,955	95	0.4	5	0.4
Other	4,324	90	1.1	10	1.1	3,765	91	0.8	9	0.8

NOTE: s.e. is standard error. Percentages include only those students for whom public/private enrollment was reported, that is, children whose parents indicated they were enrolled in school.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Parent and Family Involvement in Education Survey of the National Household Education Surveys Program, 2007; and U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 2005.

Table C-10. Percentage of children ages 3 through 5 not yet in kindergarten participating in center-based arrangements, by race/ethnicity: SR-NHES:2007, ECPP-NHES:2005, ECPP-NHES:2001, Parent-NHES:1999, and ECPP-NHES:1995

Child's race/ethnicity	Number of children (thousands)	Percent	s.e.
SR-NHES:2007			
White, non-Hispanic	4,664	65	1.4
Black, non-Hispanic	1,311	67	5.5
Hispanic	1,899	41	2.4
Other	812	63	5.0
ECPP-NHES:2005			
White, non-Hispanic	5,177	63	1.1
Black, non-Hispanic	1,233	69	3.1
Hispanic	1,822	47	2.1
Other	834	64	3.4
ECPP-NHES:2001			
White, non-Hispanic	5,313	62	0.9
Black, non-Hispanic	1,251	67	2.8
Hispanic	1,506	42	1.9
Other	482	64	4.2
Parent-NHES:1999			
White, non-Hispanic	5,389	61	0.9
Black, non-Hispanic	1,214	71	2.4
Hispanic	1,376	44	2.2
Other	547	65	4.1
ECPP-NHES:1995			
White, non-Hispanic	6,334	58	1.5
Black, non-Hispanic	1,396	61	3.5
Hispanic	1,042	39	2.3
Other	457	54	5.4

NOTE: s.e. is standard error. Center-based programs include nursery schools, preschools, center-based Head Start programs, and prekindergartens.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Readiness (SR) Survey of the National Household Education Surveys Program (NHES), 2007; ECPP-NHES:2005; ECPP-NHES:2001; Parent-NHES:1999; and ECPP-NHES:1995.

Table C-11. **Percentage of children ages 3 through 5 not yet in kindergarten participating in center-based programs, by high and low income: SR-NHES:2007, ECPP-NHES:2005, ECPP-NHES:2001, Parent-NHES:1999, PFI/CI-NHES:1996, ECPP-NHES:1995, and SR-NHES:1993**

Income level	SR-NHES:2007		ECPP-NHES:2005		ECPP-NHES:2001		Parent-NHES:1999		PFI/CI-NHES:1996		ECPP-NHES:1995		SR-NHES:1993	
	Per-cent	s.e.	Per-cent	s.e.	Per-cent	s.e.	Per-cent	s.e.	Per-cent	s.e.	Per-cent	s.e.	Per-cent	s.e.
High income	70	1.6	67	1.2	69	1.3	71	1.4	72	1.6	76	1.8	75	1.4
Low income	42	6.2	53	4.5	46	3.8	56	3.2	43	2.9	49	3.2	47	2.0

NOTE: s.e. is standard error. Center-based programs include nursery schools, preschools, center-based Head Start programs, and prekindergartens. High income was defined as household income of over $50,000. Low income was defined as household income of $10,000 or less.

SOURCE: U.S. Department of Education, National Center for Education Statistics, School Readiness (SR) Survey of the National Household Education Surveys Program (NHES), 2007; ECPP-NHES:2005; ECPP-NHES:2001; Parent Survey of NHES, 1999; Parent and Family Involvement in Education/Civic Involvement Survey (PFI/CI) of NHES, 1996; ECPP-NHES:1995; and School Readiness Survey (SR) of NHES, 1993.

Table C-12.   Percentage of children ages 3 through 5 not yet in kindergarten, by family structure, parents' highest level of education, and urbanicity of ZIP Code area: SR-NHES:2007; ECPP-NHES:2005, ECPP-NHES:2001, Parent-NHES:1999, and ECPP-NHES:1995

Family and community characteristics	SR-NHES:2007		ECPP-NHES:2005		ECPP-NHES:2001		Parent-NHES:1999		ECPP-NHES:1995		CPS 2005-2006	
	Percent	s.e.	Percent	s.e.	Percent	s.e.	Percent	s.e.	Percent	s.e.	Percent	s.e.
Family structure												
Mother and father	79	1.0	77	0.8	75	0.9	71	0.9	73	0.7	71	0.7
Mother	17	0.9	19	0.8	21	0.8	24	0.9	24	0.7	23	0.7
Father	1	0.3	2	0.3	2	0.3	3	0.3	2	0.2	4	0.3
Nonparent guardian(s)	3	0.5	2	0.3	2	0.3	2	0.3	2	0.3	3	0.3
Parents' highest education												
Less than high school	7	0.8	7	0.6	9	0.6	8	0.5	8	0.5	9	0.5
High school graduate	21	1.2	27	1.1	28	1.0	27	0.9	31	0.8	25	0.8
Some college	29	1.2	27	1.0	29	0.9	30	0.9	29	0.8	29	0.8
College graduate	22	1.0	21	0.9	19	0.8	19	0.7	17	0.7	23	0.8
Graduate school	21	1.0	18	0.8	15	0.7	15	0.7	14	0.7	14	0.6
Household urbanicity												
Urban	80	0.4	77	0.6	74	0.8	75	0.8	74	0.7	—	—
Rural	20	0.4	23	0.6	26	0.8	25	0.8	26	0.7	—	—

—Not available.
NOTE: s.e. is standard error. Mother and father refer to birth, adoptive, step, or foster parents. Detail may not sum to totals because of rounding. Parents' highest level of education for SR-NHES:2007 was derived by taking into account the education level of second mothers/female guardians and second fathers/male guardians whereas parents' highest level of education for prior years was derived by taking into account only the education level of primary mothers/female guardians and primary fathers/male guardians. Current Population Survey percentage estimates by family structure are for children ages 3 through 5, excluding emancipated minors, from CPS March 2006. Current Population Survey percentage estimates by parents' highest education are approximated by highest education attainment within households, from CPS October 2005.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Readiness (SR) Survey of the National Household Education Surveys Program (NHES), 2007; ECPP-NHES:2005; ECPP-NHES:2001; Parent Survey of NHES, 1999; and ECPP-NHES:1995.

Table C-13. **Number and percentage of children ages 3 through 5 not yet in kindergarten, by parents' highest level of education and race/ethnicity: SR-NHES:2007, ECPP-NHES:2005, ECPP-NHES:2001, Parent-NHES:1999, and ECPP-NHES:1995**

Race/ethnicity	Number of children (thousands)	Parents' highest level of education									
		Less than high school		High school		Some college		College graduate		Graduate school	
		Percent	s.e.	Percent	s.e.	Percent	s.e.	Percent	s.e.	Percent	s.e.
SR-NHES:2007											
White, non-Hispanic	4,664	2	0.5	15	1.4	28	1.6	30	1.5	26	1.3
Black, non-Hispanic	1,311	9	3.4	29	4.9	38	4.4	11	1.7	14	2.9
Hispanic	1,899	17	2.1	34	2.0	29	2.5	11	1.5	9	1.5
Other	812	6	2.8	15	3.5	19	3.3	22	4.3	38	4.8
ECPP-NHES:2005											
White, non-Hispanic	5,177	2	0.5	21	1.3	27	1.3	27	1.3	23	1.2
Black, non-Hispanic	1,233	7	1.6	42	3.6	33	3.9	12	2.4	5	1.1
Hispanic	1,822	21	1.9	35	2.4	25	1.9	11	1.4	8	1.0
Other	834	4	2.8	19	3.0	28	3.4	20	2.6	28	3.6
ECPP-NHES:2001											
White, non-Hispanic	5,313	4	0.6	26	1.3	30	1.2	23	1.2	18	1.0
Black, non-Hispanic	1,251	17	2.1	32	3.0	32	2.2	12	1.7	7	1.1
Hispanic	1,506	24	1.8	36	2.2	23	1.7	10	1.1	6	0.9
Other	482	6	2.1	24	4.0	19	2.8	18	3.2	32	4.3
Parent-NHES:1999											
White, non-Hispanic	5,389	2	0.5	24	1.1	31	1.3	24	1.0	19	0.9
Black, non-Hispanic	1,214	12	1.6	37	2.6	32	2.2	11	1.5	8	1.4
Hispanic	1,376	27	1.9	31	1.8	28	1.7	9	1.1	5	0.8
Other	547	6	2.0	24	3.5	30	3.2	20	3.0	20	3.0

See notes at end of table.

Table C-13. Number and percentage of children ages 3 through 5 not yet in kindergarten, by parents' highest level of education and race/ethnicity: SR-NHES:2007, ECPP-NHES:2005, ECPP-NHES:2001, Parent-NHES:1999, and ECPP-NHES:1995—Continued

Race/ethnicity	Number of children (thousands)	Parents' highest level of education									
		Less than high school		High school		Some college		College graduate		Graduate school	
		Percent	s.e.	Percent	s.e.	Percent	s.e.	Percent	s.e.	Percent	s.e.
ECPP-NHES:1995											
White, non-Hispanic	6,334	4	0.5	29	1.1	29	1.0	21	0.9	18	1.0
Black, non-Hispanic	1,396	16	2.0	39	2.6	32	2.5	8	1.5	5	1.3
Hispanic	1,042	27	2.0	37	1.6	24	1.7	6	0.9	6	1.1
Other	457	4!	1.9!	26	4.2	32	4.4	19	3.5	19	3.1

NOTE: s.e. is standard error. Detail may not sum to totals because of rounding. Parents' highest level of education for SR-NHES:2007 was derived by taking into account the education level of second mothers/female guardians and second fathers/male guardians whereas parents' highest level of education for prior years was derived by taking into account only the education level of primary mothers/female guardians and primary fathers/male guardians.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Readiness (SR) Survey of the National Household Education Surveys Program (NHES), 2007; ECPP-NHES:2005; ECPP-NHES:2001; Parent Survey of NHES, 1999; and ECPP-NHES:1995.

Table C-14. **Percentage of children ages 3 through 5 not yet in kindergarten whose parents reported reading to them three times a week or more: SR-NHES:2007, ECPP-NHES:2005, ECPP-NHES:2001, Parent-NHES:1999, PFI/CI-NHES:1996, ECPP-NHES:1995, and SR-NHES:1993**

Survey	Percent	s.e.
SR-NHES:2007	83	1.1
ECPP-NHES:2005	86	0.7
ECPP-NHES:2001	84	0.8
Parent-NHES:1999	82	0.7
PFI/CI-NHES:1996	83	0.9
ECPP-NHES:1995	84	0.6
SR-NHES:1993	78	0.6

NOTE: s.e. is standard error.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Readiness (SR) Survey of the National Household Education Surveys Program (NHES), 2007; ECPP-NHES:2005; ECPP-NHES:2001; Parent Survey of NHES, 1999; Parent and Family Involvement in Education/Civic Involvement (PFI/CI) Survey of NHES, 1996; ECPP-NHES:1995; and School Readiness (SR) Survey of NHES, 1993.

Table C-15. Percentage of children ages 3 through 5 not yet in kindergarten with specific disabilities: SR-NHES:2007, ECPP-NHES:2005, ECPP-NHES:2001, Parent-NHES:1999, and PFI/CI-NHES:1996

Disability	SR-NHES:2007 Percent	s.e.	ECPP-NHES:2005 Percent	s.e.	ECPP-NHES:2001 Percent	s.e.	Parent-NHES:1999 Percent	s.e.	PFI/CI-NHES:1996 Percent	s.e
Learning disability	4	0.6	2	0.3	1	0.2	2	0.3	2	0.4
Mental retardation	#	†	#	†	#	†	#	†	#	†
Speech impairment	12	0.8	10	0.6	6	0.5	7	0.5	7	0.6
Serious emotional disturbance	1	0.2	1	0.2	1	0.2	1	0.2	1	0.2
Deafness or another hearing impairment	2	0.3	1	0.2	1	0.3	1	0.2	1	0.2
Blindness or another visual impairment	1	0.4	1	0.3	2	0.3	2	0.3	1	0.2
An orthopedic impairment	2	0.4	1	0.3	1	0.2	1	0.2	2	0.3
Another health impairment lasting 6 months or more	6	0.6	3	0.4	5	0.5	5	0.4	6	0.5
Percent with any disability	18	1.0	15	0.8	13	0.8	14	0.8	15	0.6

Rounds to zero.
† Standard errors are not provided for estimates of less than 1 percent.
NOTE: s.e. is standard error.
SOURCE: U.S. Department of Education, National Center for Education Statistics, School Readiness (SR) Survey of National Household Education Surveys Program (NHES), 2007; ECPP-NHES:2005; ECPP-NHES:2001; Parent Survey of NHES, 1999; and Parent and Family Involvement/Civic Involvement (PFI/CI) Survey of NHES, 1996.

Table C-16. Percentage of children in kindergarten through grade 2 whose parents reported reading to them three or more times per week: PFI-NHES:2007, PFI-NHES:2003, Parent-NHES:1999, PFI/CI-NHES:1996, ECPP-NHES:1995, and SR-NHES:1993

Survey	Percent	s.e.
PFI-NHES:2007	69	1.6
PFI-NHES:2003	73	1.0
Parent-NHES:1999	78	0.9
PFI/CI-NHES:1996	70	0.9
ECPP-NHES:1995	78	0.7
SR-NHES:1993	66	0.7

NOTE: s.e. is standard error.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Parent and Family Involvement in Education Survey of the National Household Education Surveys Program (NHES), 2007; Parent and Family Involvement in Education Survey of the National Household Education Surveys Program (NHES), 2003; Parent Survey of NHES, 1999; Parent and Family Involvement in Education/Civic Involvement Survey of NHES, 1996; Early Childhood Program Participation Survey of NHES, 1995; and School Readiness Survey of NHES, 1993.

Table C-17. Percentage of children in kindergarten through grade 12, by school size: PFI-NHES:2007, PFI-NHES:2003, Parent-NHES:1999, and PFI/CI-NHES:1996

School size	PFI-NHES:2007		PFI-NHES:2003		Parent-NHES:1999		PFI/CI-NHES:1996	
	Percent	s.e.	Percent	s.e.	Percent	s.e.	Percent	s.e.
Under 300	15	0.6	18	0.5	17	0.4	18	0.3
300–599	31	0.7	37	0.6	38	0.5	39	0.5
600–999	27	0.6	22	0.5	22	0.4	22	0.4
1,000 or more	27	0.5	23	0.5	23	0.4	22	0.4

NOTE: s.e. is standard error. Students who are homeschooled are not represented. Because of rounding, percentages may not add to 100. The estimates of PFI-NHES:2007 were based on the school size information on the CCD/PSS data files and excluded cases with missing school size. The estimates of previous NHES surveys were based on the school size reported by parents.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Parent and Family Involvement in Education Survey of the National Household Education Surveys Program (NHES), 2007; Parent and Family Involvement in Education Survey of the National Household Education Surveys Program (NHES), 2003; Parent Survey of the NHES, 1999; and Parent and Family Involvement/Civic Involvement Survey of the NHES, 1996.

Table C-18. Percentage of children in kindergarten through grade 12, by family structure, parents' highest level of education, and urbanicity: PFI-NHES:2007, PFI-NHES:2003, Parent-NHES:1999, PFI/CI-NHES:1996, and CPS: 2005-2006

Family and community characteristics	PFI-NHES:2007		PFI-NHES:2003		Parent-NHES:1999		PFI/CI-NHES:1996		CPS:2005-2006	
	Percent	s.e.	Percent	s.e.	Percent	s.e.	Percent	s.e.	Percent	s.e.
Family structure										
Mother and father	73	0.5	71	0.6	66	0.4	69	0.4	68	0.4
Mother	20	0.5	22	0.6	27	0.4	24	0.4	24	0.3
Father	3	0.2	4	0.3	4	0.2	3	0.2	5	0.2
Nonparent guardian(s)	4	0.4	3	0.2	3	0.2	3	0.2	4	0.1
Parents' highest education										
Less than high school	7	0.4	7	0.4	9	0.3	10	0.3	9	0.2
High school graduate	21	0.6	25	0.6	28	0.4	31	0.4	24	0.3
Some college	29	0.6	31	0.6	30	0.4	30	0.5	33	0.4
College graduate	22	0.5	19	0.5	16	0.3	15	0.4	21	0.3
Graduate school	21	0.5	17	0.5	17	0.4	14	0.4	13	0.2
Household urbanicity										
Urban	79	0.0	79	0.0	74	0.2	—	—	—	—
Rural	21	0.0	21	0.0	26	0.2	—	—	—	—

—Not available.
NOTE: s.e. is standard error. Because the standard error of an estimate is a measure of sampling error variance, a standard error of zero indicates the absence of sampling error variance. When the NHES estimates of totals are adjusted to exactly match CPS totals, all sampling error in those estimated totals is eliminated, under the assumption that the CPS total is the true population value. Mother and father refer to birth, adoptive, step, or foster parents. Because of rounding, percentages may not add to 100. Parents' highest level of education for PFI-NHES:2007 was derived by taking into account the education level of second mothers/female guardians and second fathers/male guardians whereas parents' highest level of education for prior years was derived by taking into account only the education level of primary mothers/female guardians and primary fathers/male guardians. Current Population Survey percentage estimates by family structure are for children ages 5 through 17, excluding emancipated minors, from CPS March 2006. Current Population Survey percentage estimates by parents' highest education are approximated by highest education attainment within households, from CPS October 2005.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Parent and Family Involvement in Education Survey of the National Household Education Surveys Program (NHES), 2007; Parent and Family Involvement in Education Survey of the National Household Education Surveys Program (NHES), 2003; Parent Survey of the NHES, 1999; and Parent and Family Involvement/Civic Involvement Survey of the NHES, 1996. U.S. Department of Education, National Center for Education Statistics, Parent and Family Involvement in Education Survey of the National Household Education Surveys Program, 2003; and U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 2005 and March 2006.

Table C-19. **Number and percentage of students in kindergarten through grade 12, by parents' highest level of education and race/ethnicity: PFI-NHES:2007, PFI-NHES:2003, Parent-NHES:1999, and PFI/CI-NHES:1996**

Race/ethnicity	Number of children (thousands)	Less than high school Percent	s.e.	High school Percent	s.e.	Some college Percent	s.e.	College graduate Percent	s.e.	Graduate school Percent	s.e.
PFI-NHES:2007											
White, non-Hispanic	31,035	2	0.3	17	0.7	28	0.8	27	0.7	26	0.6
Black, non-Hispanic	7,898	11	1.6	30	2.0	33	2.2	15	1.6	11	1.0
Hispanic	9,929	19	1.4	29	1.6	29	1.5	13	1.2	10	0.7
Other	4,324	4	1.1	15	2.3	30	2.2	22	1.7	29	1.9
PFI-NHES:2003											
White, non-Hispanic	32,844	3	0.4	21	0.8	32	0.8	23	0.7	21	0.7
Black, non-Hispanic	8,274	11	1.4	33	1.6	35	1.6	12	1.0	10	1.1
Hispanic	8,322	22	1.1	32	1.3	28	1.3	10	0.9	8	0.8
Other	3,143	2	0.8	21	2.7	28	2.7	22	2.0	27	2.6
Parent-NHES:1999											
White, non-Hispanic	33,512	3	0.2	25	0.6	32	0.6	19	0.5	20	0.5
Black, non-Hispanic	8,343	13	1.1	40	1.3	29	1.2	10	0.7	9	0.6
Hispanic	7,322	31	1.3	28	1.0	25	0.9	9	0.6	7	0.5
Other	2,719	7	1.2	24	1.8	26	2.2	20	1.7	23	2.1
PFI/CI-NHES:1996											
White, non-Hispanic	33,730	5	0.3	28	0.6	32	0.5	18	0.5	17	0.4
Black, non-Hispanic	7,865	15	0.9	41	1.5	30	1.2	9	0.6	5	0.5
Hispanic	6,424	32	1.2	32	1.2	22	1.2	7	0.8	7	0.7
Other	2,108	6	1.0	25	1.8	31	2.1	20	1.8	19	1.6

NOTE: s.e. is standard error. Because of rounding, percentages may not add to 100. Parents' highest level of education for PFI-NHES:2007 was derived by taking into account education levels of second mothers/female guardians and second fathers/male guardians whereas parents' highest level of education for prior years was derived by taking into account only the education level of primary mothers/female guardians and primary fathers/male guardians.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Parent and Family Involvement in Education Survey of the National Household Education Surveys Program (NHES), 2007; Parent and Family Involvement in Education Survey of the National Household Education Surveys Program (NHES), 2003; Parent Survey of NHES, 1999; and Parent and Family Involvement/Civic Involvement Survey of NHES, 1996.

Table C-20.   Percentage of students enrolled in kindergarten through grade 12 whose parents reported selected school contacts with family: PFI-NHES:2007, PFI-NHES:2003, Parent-NHES:1999, and PFI/CI-NHES:1996

School effort to contact family	PFI-NHES:2007		PFI-NHES:2003		Parent-NHES:1999		PFI/CI-NHES:1996	
	Percent	s.e.	Percent	s.e.	Percent	s.e.	Percent	s.e.
School contacted parents about student's academic performance	23	0.6	26	0.5	19	0.3	27	0.4
School contacted parents about student's behavior	23	0.6	19	0.4	23	0.4	22	0.4

NOTE: s.e. is standard error. Students who are homeschooled are not represented.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Parent and Family Involvement in Education Survey of the National Household Education Surveys Program (NHES), 2007; Parent and Family Involvement in Education Survey of the National Household Education Surveys Program (NHES), 2003; Parent Survey of NHES, 1999; and Parent and Family Involvement/Civic Involvement Survey of NHES, 1996.

Table C-21. Percentage of students enrolled in kindergarten through grade 12 whose parents reported attendance at selected school meetings and events, volunteering, and fundraising: PFI-NHES:2007, PFI-NHES:2003, Parent-NHES:1999, and PFI/CI-NHES:1996

School effort to contact family	PFI-NHES:2007 Percent	s.e.	PFI-NHES:2003 Percent	s.e.	Parent-NHES:1999 Percent	s.e.	PFI/CI-NHES:1996 Percent	s.e.
Attended a general school meeting (open house), back-to-school night, meeting of parent-teacher organization	89	0.5	88	0.4	78	0.5	77	0.4
Went to a regularly scheduled parent-teacher conference with child's teacher	78	0.5	77	0.4	73	0.5	72	0.4
Attended a school or class event (e.g., play, sports event, science fair) because of child	74	0.6	70	0.4	65	0.4	67	0.4
Acted as a volunteer at the school or served on a committed	44	0.6	42	0.6	37	0.4	39	0.4

NOTE: s.e. is standard error. Students who are homeschooled are not represented.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Parent and Family Involvement in Education Survey of the National Household Education Surveys Program (NHES), 2007; Parent and Family Involvement in Education Survey of the National Household Education Surveys Program (NHES), 2003; Parent Survey of NHES, 1999; and Parent and Family Involvement/Civic Involvement Survey of NHES, 1996.

Table C-22.   Percentage of children in kindergarten through grade 12 with specific disabilities: PFI-NHES:2007, PFI-NHES:2003 and Parent-NHES:1999

Disability	PFI-NHES:2007		PFI-NHES:2003		Parent-NHES:1999	
	Percent	s.e.	Percent	s.e.	Percent	s.e.
Learning disability	10	0.5	9	0.3	9	0.4
Mental retardation	1	0.1	1	0.1	1	0.1
Speech impairment	9	0.4	6	0.3	4	0.2
Serious emotional disturbance	3	0.3	3	0.2	3	0.2
Deafness or another hearing impairment	2	0.2	2	0.1	2	0.1
Blindness or another visual impairment	2	0.1	8	0.3	5	0.2
An orthopedic impairment	2	0.2	3	0.2	2	0.1
Another health impairment lasting 6 months or more	8	0.4	8	0.3	6	0.2
Percent with any disability	24	0.7	26	0.5	21	0.4

NOTE: s.e. is standard error.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Parent and Family Involvement in Education Survey of the National Household Education Surveys Program (NHES), 2007; Parent and Family Involvement in Education Survey of the National Household Education Surveys Program (NHES), 2003; Parent Survey of NHES, 1999; and Parent and Family Involvement/Civic Involvement Survey of NHES, 1996.