NATIONAL CENTER FOR EDUCATION STATISTICS

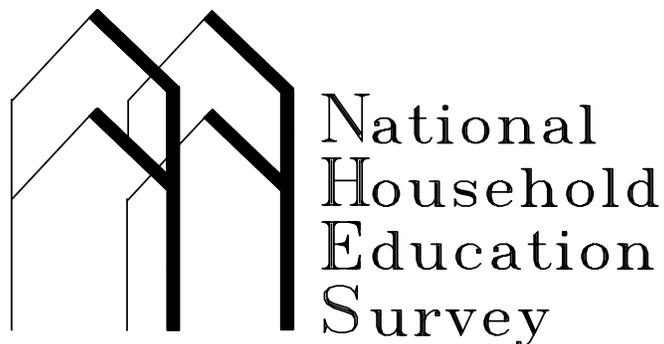**User's Manual**   **October 2000**

National Household Education Survey of 1999

Data File User's Manual

Volume I

National
Household
Education
Survey

NATIONAL CENTER FOR EDUCATION STATISTICS

User's Manual October 2000

National Household Education Survey of 1999

Data File User's Manual

Volume I

National Household Education Survey

Mary Jo Nolin, Project Director
Jill Montaquila, Senior Statistician
Jean Lennon
Brian Kleiner
Kwang Kim
Westat

Christopher Chapman
Kathryn Chandler
National Center for Education Statistics

Sean Creighton
Stacey Bielick
Education Statistics Services Institute

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
C. Kent McGuire
Assistant Secretary

National Center for Education Statistics
Gary W. Phillips
Acting 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
> Office of Educational Research and Improvement
> U.S. Department of Education
> 1990 K Street, NW
> Washington, DC 20006-5650

October 2000

The NHES World Wide Web Home Page is
http://nces.ed.gov/nhes

Suggested Citation

U.S. Department of Education. National Center for Education Statistics. *National Household Education Survey of 1999: Data File User's Manual, Volume I*, NCES 2000–076, by Mary Jo Nolin, Jill Montaquila, Jean Lennon, Brian Kleiner, Kwang Kim, Christopher Chapman, Kathryn Chandler, Sean Creighton, and Stacey Bielick. Washington, DC: 2000.

For ordering information on this report, write:	**Contact:**
U.S. Department of Education	Christopher Chapman
E.D. Pubs	(202) 502-7327
P.O. Box 1398	(e-mail) nhes@ed.gov
Jessup, MD 20794-1398	

or, http://www.ed.gov/pubs/edpubs.html
or call toll free 1-887-4ED-Pubs

CONTENTS

CONTENTS--Continued

CONTENTS--Continued

LIST OF TABLES

CONTENTS—Continued

LIST OF FIGURES

Figure

1. INTRODUCTION

The 1999 National Household Education Survey (NHES:1999) was a random digit dial (RDD) telephone survey of households in the United States developed by the National Center for Education Statistics (NCES) and conducted by Westat. The NHES:1999 was a compilation of key items from previous NHES survey years, and included three interviews:

■ The Parent Interview, in which data were collected on a variety of topics including early childhood program participation, types and frequency of family involvement in children's schooling, school practices to involve and support families, learning activities with children outside of school, and plans for their children's postsecondary education;

■ The Youth Interview, in which children in 6th through 12th grades, whose parents had completed a Parent Interview, were asked about school and family environments, civic involvement and community service, and plans for postsecondary education; and

■ The Adult Education Interview, in which data, such as type of program, employer support, and degree sought were collected on characteristics of participation in six types of adult educational activities, including English as a second language, adult basic education, credential programs, apprenticeships, work-related courses, and personal development courses.

There were three populations of interest for the NHES:1999:

■ Children birth through grade 12, whose parents completed the Parent Interview, which was structured according to the age or grade of their children;

■ Students in grades 6 through 12, who completed the Youth Interview; and

■ Adults, defined as persons 16 years old or older, not enrolled in grade 12 or below, and not on active duty in the military, who responded to the Adult Education Interview.

Data were collected from January 3 through April 3, 1999.

This manual provides documentation and guidance for users of the three public release data files of the NHES:1999: the Parent Interview file, the Youth Interview file, and the Adult Education Interview file. Information about the purpose of the study, the data collection instruments, the sample design, data collection and data processing procedures, and a brief guide to the data files is contained in Volume I of this manual. A reference guide to NHES:1999 variables in previous NHES data files is also provided.

Volumes II through IV of the NHES:1999 Data File User's Manual each contain a guide to a data file and codebook, a discussion of data considerations and anomalies and, in appendixes, the file layout, derived variable specifications, cross references of NHES:1999 variables with variables in previous NHES files, the codebook for one of the data files from the NHES:1999, and directions and sample code for linking the NHES:1999 files.

■ The Parent Interview file (Volume II) contains data from interviews completed with parents of 24,600 children—3,378 infants, 3,561 preschool children, 8,372 primary school children (grades K–5), 4,024 middle or junior high school children (grades 6–8), 4,980 high school children (grades 9–12), and 285 home school children;

■   The Youth Interview file (Volume III) contains data from interviews completed with 7,913 6th through 12th graders;

■   The Adult Education Interview file (Volume IV) contains data from interviews completed with 6,697 adults.

1.1   Background of Study

The National Household Education Survey is a household survey developed by the National Center for Education Statistics and conducted by Westat. This RDD, computer-assisted telephone interview (CATI) has been conducted in the spring of 1991, 1993, 1995, 1996, and 1999. The NHES complements the Center's other surveys, which primarily collect data through institution based surveys. By collecting data directly from households, the NHES collection allows the Center to gather data on issues that cannot easily be addressed through institutional surveys such as early childhood care and education, children's readiness for school, parent perceptions of school safety and discipline, participation in adult and continuing education, parent involvement in education, and civic involvement.

A large field test of the NHES was conducted in the fall of 1989. This first effort, which included the screening of about 15,000 households, included two topical components: school dropouts (interviews were conducted with adult household respondents and 14- to 21-year-old youth) and early childhood education (interviews were conducted with parents/guardians 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).

The first full-scale implementation of the NHES was conducted in the spring of 1991 (NHES:1991). The topical components in the survey were Early Childhood Education (ECE) for 3- to 8-year-olds and participation in adult education (AE). For the NHES:1991, more than 60,000 households were screened, nearly 14,000 Early Childhood Education interviews were conducted with the parents/guardians of eligible children, and about 12,500 interviews were conducted with adults regarding participation in adult education activities. More information about these data can be found in the *National Household Education Survey of 1991: Adult and Course Data Files User's Manual* (Brick et al. 1992a) and the *National Household Education Survey of 1991: Preprimary and Primary Data Files User's Manual* (Brick et al. 1992b).

The NHES was conducted again in 1993, addressing the topics of readiness for school and safety and discipline in school. The NHES:1993 School Readiness (SR) component focused on readiness for school in a broad sense and examined several relevant issues, such as experience in early childhood programs, the child's accomplishments and difficulties in several developmental domains, home activities, delayed kindergarten entry, and early school experience. Parents of 10,888 children aged 3 through 7 or in 2nd grade or below were interviewed. The other component in the NHES:1993, School Safety and Discipline (SS&D), focused on four areas: school environment, school safety, school discipline policy, and alcohol/other drug use and education. Parents of 12,680 children in 3rd through 12th grades were interviewed, as were 6,504 students in 6th through 12th grades whose parents had completed an interview. More information about these data can be found in the *National Household Education Survey of 1993: School Readiness Data File User's Manual* (Brick et al. 1994a) and the *National Household Education Survey of 1993: School Safety and Discipline Data File User's Manual* (Brick et al. 1994b).

In 1995, the topics from the NHES:1991 were repeated. The NHES:1995 Early Childhood Program Participation component (ECPP) focused on children's early experiences in various types of nonparental child care arrangements and early childhood programs. For the ECPP component of the NHES:1995, interviews were completed with parents of 14,604 children from birth through 3rd grade up to

age 10. The NHES:1995 Adult Education component (AE) focused on the participation of adults in a wide range of educational activities during the 12 months prior to the interview. The population for this survey was defined as persons age 16 and older who were not enrolled in elementary or secondary school and not on active duty in the U.S. Armed Forces. In all, 19,722 adult interviews were completed for the AE component of the NHES:1995: 11,713 with adult education participants and 8,009 with nonparticipants. More information about these data can be found in the *National Household Education Survey of 1995: Adult Education Data File User's Manual* (Collins et al. 1996a), the *National Household Education Survey of 1995: Early Childhood Program Participation Data File User's Manual* (Collins et al. 1996b), and the *National Household Education Survey of 1996: Data File User's Manual (*Collins et al. 1997*).*

In 1996, the NHES focused on the topics of Parent/Family Involvement in Education (PFI) and Civic Involvement (CI). The PFI component of the NHES:1996 focused on family involvement in children's education in four areas: children's schools, communication with teachers or other school personnel, children's homework and behavior, and learning activities with children outside of school. Interviews were completed for 20,792 children aged 3 through 20 and in 12th grade or below. The CI component of the NHES:1996 addressed sources of information about government and national issues, civic participation, and knowledge and attitudes about government. Items were administered to youth in grades 6 through 12 and their parents, as well as to a representative sample of U.S. adults. The CI component also addressed opportunities that youth have to develop the personal responsibility and skills that would facilitate their taking an active role in civic life. Interviews were completed with 9,393 parents of students in grades 6 through 12, 8,043 youth in grades 6 through 12, and 2,250 adults.

1.2   NHES:1999 Survey Topics

The NHES:1999 addressed a wide variety of educational topics that had been covered in previous NHES surveys as well as including a new focus on postsecondary education plans. The survey was designed to provide the Department of Education with end-of-decade measures of important education indicators. Topics were selected by identifying associated items most prevalently used in published estimates by the Department or other education researchers, evaluating the data needs for measuring the Department's Strategic Plan Indicators, consulting with NHES data users and education researchers about issues they considered important to measure at the end of the decade, and evaluating the content of other studies that could potentially overlap the content of the NHES:1999. In addition to the key issues identified through these sources, an Adult Special Study was included in this survey administration; it contained additional items to address a number of specific methodological issues, such as how different ways of asking race/ethnicity questions effect how respondents report these characteristics. (Except where noted, the Adult Special Study did not affect NHES:1999).

Parent Interview

The NHES:1999 Parent Interview included a variety of educational topics, each appropriate for certain age/grade groups of children. Some of the topics included were early childhood program participation, emerging literacy and numeracy, parent satisfaction with and involvement in the child's education, school practices to involve and support families, learning activities with children outside of school, family involvement in learning outside of school, and parent reports about child's postsecondary education plans. Other information collected in this component pertains to the child's personal and demographic characteristics, household characteristics, and the child's health and disability statuses. Interviews were completed for 24,600 children birth through 20 years old and in 12th grade or below.

Youth Interview

Asked of youth in grades 6 through 12 whose parents had completed a Parent Interview, the NHES:1999 Youth Interview focused on school, family environment, civic involvement and community service, and plans for postsecondary education. Interviews were completed with 7,913 youth in grades 6 through 12.

Adult Education Interview

The NHES:1999 Adult Education Interview focused on the participation of civilian adults (aged 16 years and older and not enrolled in 12th grade or below) in a wide range of educational activities. Respondents were asked about their participation in six broadly defined types of adult educational activities: English as a second language instruction, adult basic skills and GED preparation classes, courses taken towards college degrees or vocational diplomas or certificates, apprenticeship programs, career- or job-related courses, and personal interest/development courses. Interviews were completed with 6,697 adults, 3,999 with adult education participants and 2,698 with nonparticipants.

1.3   Overview of Design

The NHES:1999 was developed to provide reliable national estimates from the Parent, Youth and Adult Education Interviews. Because of the high costs associated with screening large numbers of households in order to meet the sample size requirements for precise estimates, more than one population and set of issues were addressed concurrently in the NHES:1999. This strategy is key to the NHES design. By including more than one topic within the framework of a single survey, the cost of screening households to find those eligible for the study could be partitioned over the three component interviews.

Another general feature of the NHES was developed in response to concerns about the burden placed upon those who respond to multiple survey components. With the introduction of multiple surveys within a single framework, the possibility of increasing response burden on the members of the sampled households arose. It is possible that the same household member could have been selected to respond to more than one interview and/or that more than one household member could be sampled. A Screener was used to collect information on household composition and interview eligibility. For the Parent Interview, if there were one or more children qualifying for one of the six possible interview paths (infant, preschool, elementary, middle school, high school, or home school), only one child was selected for that path. A maximum of three Parent Interviews were allowed per household but only if the household had not been designated for an Adult Education Interview. If a child was selected as the subject of a middle or high school path interview, following completion of that Parent Interview, he or she was asked to complete a Youth Interview. Only one 6th through 12th grader was sampled per household. To assure households did not have to participate in too many interviews, the maximum number of interviews per household was set to four. Three Parent Interviews were allowed in some households, and two Parents, one Adult, and one Youth in others. (See chapter 3 for a detailed discussion of sampling procedures for the NHES:1999.)

Even though sampling methods reduced the number of interviews per household, the length of the interview was considered to be a critical factor in obtaining good response rates and reliable estimates. Therefore, the number of items included in the NHES:1999 was limited in order to help improve response rates and reduce the demands made on survey respondents. The average administration time for the Screener was 1.9 minutes for Screeners with no extended interviews, 3.6 minutes for Screeners with Parent Interviews only, 4.0 minutes for Screeners with Parent and Youth Interviews, 3.0 minutes for Screeners with Adult Interviews only, 4.4 minutes for Screeners with Parent and Adult Interviews, and 4.9 minutes for Screeners with Parent, Youth, and Adult Interviews. The average administration time for the Parent

Interview was 14.0 minutes; for the Youth Interview it was 12.5 minutes. The Adult Education Interview took an average of 13.4 minutes to administer to adult education participants and 9.4 minutes for nonparticipants.

Because of the requirement to reduce respondent burden, the complex sampling techniques employed, and the need for quick and accurate administration, the NHES:1999 was conducted using computer-assisted telephone interviewing (CATI) technology. Some of the advantages of CATI included improved project administration, online sampling and eligibility checks, scheduling of interviews according to a priority scheme to improve response rates, managing data quality by controlling skip patterns and checking responses online for range and consistency, and an online "help" function to assist interviewers in answering respondents' questions. Items within each of the NHES:1999 instruments were programmed so that the appropriate items appeared on the interviewer's computer screen according to the respondent's answers to previous questions.

Table 1-1 summarizes the number of completed interviews and gives weighted completion and response rates for the Screener, the Parent, Youth, and Adult Education Interviews. Table 1-2 gives unweighted completion and response rates for the Screener and the Parent, Youth, and Adult Education Interviews. More details on the computation of these rates are given in chapter 4.

Table 1-1.—Summary of completed interviews and weighted completion and response rates

Interview type	Number of completed interviews	Completion rate[1]	Response rate[2]
Screener..	57,278	74.1	74.1
Parent Interview	24,600	88.3	65.4
Youth Interview	7,913	76.0	56.3
Adult Education Interview...	6,697	81.6	60.4

[1]The completion rate is the percentage of completed interviews for a specific stage of the survey (i.e., the Screener, the Parent Interview, the Youth Interview, or the Adult Education Interview). It is a ratio of the number of completed interviews to the number of units (e.g., households, household members) sampled for the interviews. For many telephone numbers sampled for the Screener interview, no contact was ever made. Based on previous research (Brick and Broene 1997), 40.5 percent of these numbers were assumed to be residential and were added to the denominator for the calculation of the Screener completion and response rates.

[2]The response rate indicates the percentage of possible interviews that have been completed, taking all sampling stages into account. It is the product of the Screener completion rate and the completion rate for the Parent Interview, the Youth Interview, or the Adult Education Interview. The response rate and completion rate are identical for the first stage of sampling and interviewing (i.e., the Screener).

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1999.

Table 1-2.—Summary of completed interviews and unweighted completion and response rates

Interview type	Number of completed interviews	Completion rate[1]	Response rate[2]
Screener...	57,278	73.5	73.5
Parent Interview	24,600	88.3	64.9
Youth Interview	7,913	75.5	55.5
Adult Education Interview...	6,697	83.5	61.4

[1]The completion rate is the percentage of completed interviews for a specific stage of the survey (i.e., the Screener, the Parent Interview, the Youth Interview, or the Adult Education Interview). It is a ratio of the number of completed interviews to the number of units (e.g., households, household members) sampled for the interviews. For many telephone numbers sampled for the Screener interview, no contact was ever made. Based on previous research (Brick and Broene 1997), 40.5 percent of these numbers were assumed to be residential and were added to the denominator for the calculation of the Screener completion and response rates.

[2]The response rate indicates the percentage of possible interviews that have been completed, taking all sampling stages into account. It is the product of the Screener completion rate and the completion rate for the Parent Interview, the Youth Interview, or the Adult Education Interview. The response rate and completion rate are identical for the first stage of sampling and interviewing (i.e., the Screener).

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1999.

1.4 Flow of the Interviews

Figure 1-1 shows the flow of the NHES:1999 interviews. Each household contact began with a Screener, which obtained information used to sample adults and children for extended interviews.

If the household contained any children from birth through 12th grade, up to three Parent Interviews were conducted with the parent or guardian most knowledgeable about each child's care and education. Any youth enrolled in grade 6 through 12 who was sampled was asked to respond to the Youth Interview after the Parent Interview was completed. Up to one adult was sampled in each household for an Adult Education interview. (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.

1.5 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, the Parent Interview, the Youth Interview, and the Adult Education Interview. Appendix B contains a summary of weighting and sample variance estimation variables. Volumes II through IV of the NHES:1999 Data File User's Manual provide information on the Parent Interview file, the Youth Interview File, and the Adult Education Interview file, respectively. Each contains a guide to the relevant data file and codebook, a discussion of data considerations and anomalies and, in appendixes, the file layout, derived variable specifications, a cross reference of NHES:1999 variables with variables from previous NHES administrations, the codebook for the relevant data file from the NHES:1999, and directions and sample code for linking the NHES:1999 data files.

Figure 1.1—Flow of the NHES:1999 interviews

```
┌─────────────────────────────────┐
│  Screener through household     │
│  composition and/or sampling    │
│  point.                         │
└─────────────────────────────────┘
              │
              ▼
        ╱╲
      ╱    ╲
    ╱  Any   ╲         no      ┌──────────────────────┐      ╭────────╮
   ╱ extended  ╲ ──────────▶   │ Remaining Screener   │ ──▶  │ End of │
    ╲interviews╱                │ items.               │      │ call.  │
      ╲to be  ╱                 └──────────────────────┘      ╰────────╯
        ╲done?╱
          ╲╱
           │
          yes
           │
           ▼
        ╱╲
      ╱    ╲
    ╱  Is    ╲        yes     ┌──────────────────────┐
   ╱ Screener ╲ ──────────▶   │ Conduct appropriate  │
   ╲respondent╱                │ interview(s) with    │
    ╲  for an ╱                │ Screener respondent. │
      ╲  ...  ╱                └──────────────────────┘
        ╲╱                              │
         │                              ▼
         no                          ╱╲
         │                         ╱    ╲
         ▼                       ╱  Any   ╲
┌──────────────────┐   yes      ╱remaining ╲
│ Conduct          │ ◀────────  ╲interviews╱
│ interview(s)     │             ╲to be    ╱
│ with other       │               ╲done? ╱
│ appropriate      │                 ╲╱
│ respondents.     │                  │
└──────────────────┘                  no
         │  ◀───────────────────────────
         ▼
      ╭────────╮
      │ End of │
      │ call.  │
      ╰────────╯
```
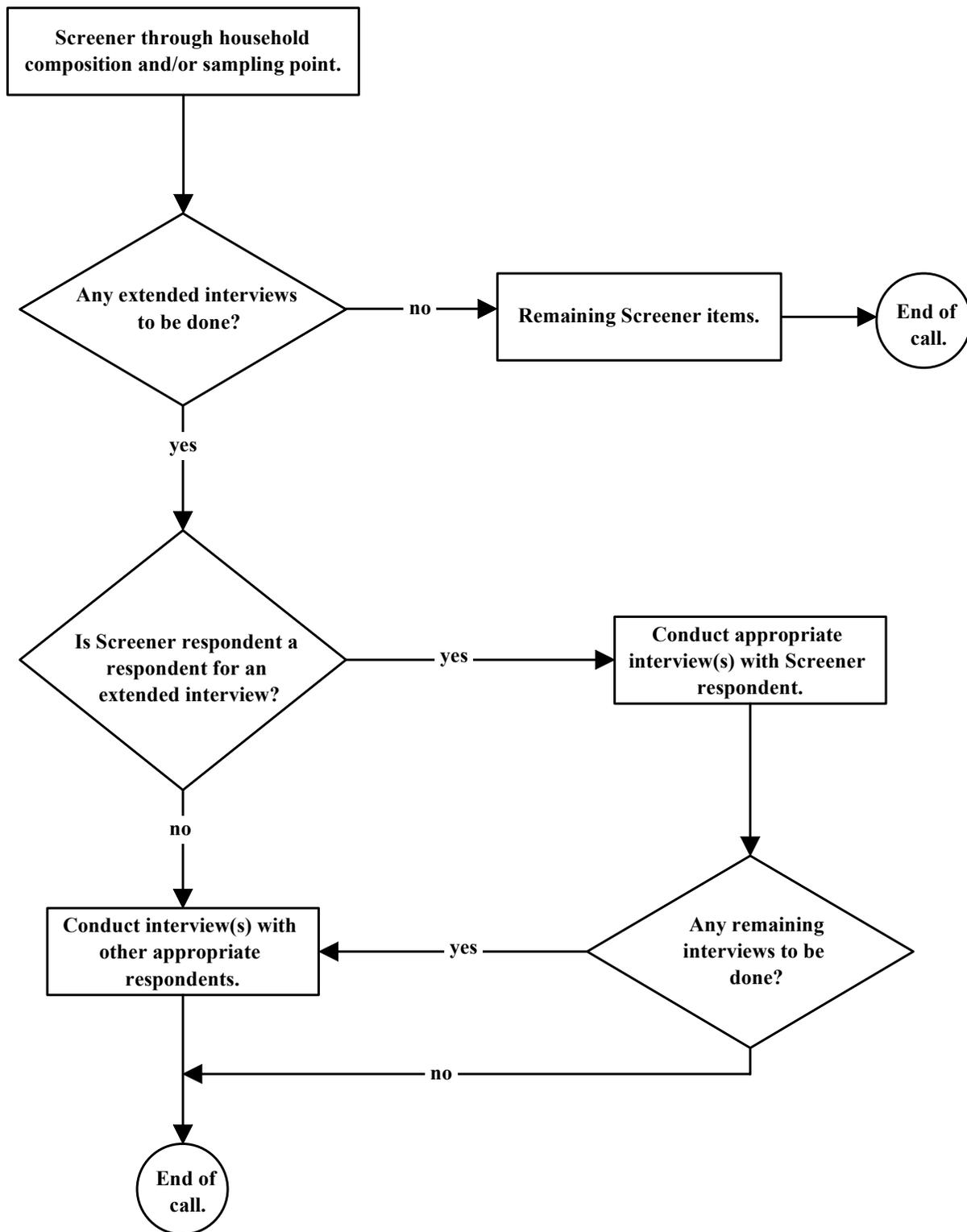
This page is intentionally blank.

2. DESCRIPTION OF DATA COLLECTION INSTRUMENTS

The sections that follow describe the instruments used to collect data contained in the Parent, Youth, and Adult Education data files in the NHES:1999. In addition to the Screener, through which eligibility was determined, three instruments were used: the Parent, Youth, and Adult Education Interviews. Appendix A contains a copy of each instrument.

2.1 The NHES:1999 Screener

As stated earlier, the screening interview in the NHES:1999 was used to determine whether sampled telephone numbers belonged to households, gather the information needed to sample household members to be interview subjects, and administer some household items in households in which no one was sampled for an extended interview. 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 and whether the household had been designated for an Adult Education Interview. In households not designated for an Adult Education Interview and in which there were children age 20 or younger, only those members age 20 and younger were enumerated. Screener questions directly following the enumeration determined whether they could be sampled for the study. The questions asked whether household members age 3 through 20 were attending school or being home schooled, and the grade or year of school in which they were enrolled. If a child was sampled for a Parent Interview, the parent/guardian most knowledgeable about the child's education and care was selected as the respondent.

In households designated for an Adult Education Interview, all household members were enumerated in the Screener. Participation in any educational activities during the past 12 months was determined for all household members age 16 or older and not currently enrolled in grade 12 or below. Following selection of an adult for the extended interview, eligibility was determined by asking whether the sampled adult was currently serving on active duty in the military. Active duty military personnel were not eligible for an Adult Education Interview. When appropriate, contact information was gathered for adults living in school-sponsored housing.

If no child was sampled for a Parent Interview and no adult was sampled for an Adult Education Interview, 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, information that was used for weighting and nonresponse adjustment.

2.2 **Parent Interview**

In the NHES:1999 Parent Interview, data were collected about children from birth through age 20 as of December 31, 1998. In order to be eligible, children aged 7 or older had to be enrolled in or home schooled in the 12th grade or below at the time of the interview. Youth who did not live with a parent/guardian or with an adult at least 12 years older than him or herself were ineligible.

Table 2-1 shows the overall structure of the interview and distribution of topics among the six groups of children that were in this survey. The respondent for the Parent Interview was the adult living 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, stepparent, adoptive parent, foster parent, grandparent, another relative, or nonrelative guardian. For simplicity, when referring to the most knowledgeable respondent in the manual, this person will be called the parent/guardian.

In the Parent Interview, subjects were routed to one of six questionnaire paths: infant, preschool, elementary school, middle/junior high, senior high, or home school. Irrespective of the questionnaire path for the child, parents were asked basic demographic questions about the child, questions about the child's health and disability status, questions about parent/guardian characteristics, and questions about household characteristics.

The **infant path** was for those children newborn through 2 years of age. Information was collected on early childhood care and programs, care and education by nonparental adults, training and support for parents of preschoolers, and family involvement in learning activities outside of school.

The **preschool path** was for those children who were age 3 or older and not yet attending kindergarten or primary school. These children were typically 3 to 5 years old. Information was collected about participation in early childhood care and programs, literacy-related skills and activities, training and support for families of preschoolers, and family involvement in programs and program practices to involve the family. Parents/guardians of preschoolers who attended Head Start or a center-based program received school-related questions, which parents of preschoolers who were not enrolled did not receive.

The **elementary school path** was administered to parents/guardians of children currently attending kindergarten through 5th grade, including transitional kindergarten and prefirst grade. As defined in the NHES, transitional kindergarten is a program before regular kindergarten for children who are old enough for kindergarten but not yet ready to start. Prefirst grade is an extra year between kindergarten and first grade. Elementary school children are typically 5 to 11 years old.

In the elementary school path, data were collected about enrollment in school, center-based program participation before school entry, school characteristics, student academics and behavior at school, parent satisfaction with school, family participation in school and school practices to involve families, before and after school care and activities, and family involvement in educational activities outside of school.

The **middle school/junior high school path** was administered to those parents/guardians of children who were attending grades 6 through 8. Data were collected about school enrollment, school characteristics, parent satisfaction with school, student academics and behavior at school, family participation in school and school practices to involve families, before and after school care and activities, and plans for postsecondary education.

The **senior high school path** was administered to parents/guardians of youth attending grades 9 through 12. This path collected data about school enrollment, school characteristics, parent satisfaction with

school, student academics and behavior at school, family participation in school and school practices to involve families, and plans for postsecondary education.

Table 2-1.—NHES:1999 Parent Interview content by major path

Interview section	Infant/ toddler (I)	Preschool (N)		Elementary school (E)		Middle school (M) 6–8	Senior high school (S) 9–12	Home school (H)
		Not enrolled	Center-based[1]	K–2	3–5			
Demographics	X	X	X	X	X	X	X	X
Current school/program status		X	X	X	X	X	X	X^2
Center-based participation before entering school				X				X
Home schooling								X
Preschool/school characteristics				X	X	X	X	X^2
Literacy-related skills		X	X					
Care/education by nonparental adults	X	X	X					
Support for families of preschoolers	X	X	X					
Parent satisfaction with school				X	X	X	X	X^2
Academics and behavior at school				X	X	X	X	X^2
Family/school involvement and school practices			X	X	X	X	X	X^2
Before-/after-school care				X	X	X		
Postsecondary plans						X	X	X
Family involvement outside of school	X	X	X	X	X			X
Health and disability	X	X	X	X	X	X	X	X
Parent/guardian characteristics	X	X	X	X	X	X	X	X
Household characteristics	X	X	X	X	X	X	X	X

[1]Center-based programs include day care centers, nursery schools, preschools, and prekindergartens.

[2]Asked of home schooled students who also attend regular school for 9 hours per week or more.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1999.

The **home school path** was administered to those parents/guardians of children who were not attending school but were being instructed at home and who had a grade equivalent of kindergarten through 12th grade. Those with a grade equivalent of nursery school were assigned to the preschool path. Parents of home schoolers were asked questions about the student's grade equivalent, center-based program participation before school entry (for those in grade equivalent K through 2), home schooling history, reasons for schooling their child at home, family involvement in educational activities outside of school, and, for those with a grade equivalent of 6 through 12, activities that may promote civic involvement among youth and plans for postsecondary education. For those students who were reported to be home schooled but also attended a school 9 or more hours per week, parents/guardians were administered the sections on school characteristics, parent satisfaction with school, student experiences and performance at school, and family participation in school and school practices to involve families. The number of children who followed the home school path was relatively small compared to the other paths.

To avoid redundancy and increased response burden, 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.

2.3   Youth Interview

In the NHES:1999 Youth Civic Involvement Interview, data were collected from students in grades 6 through 12, for whom there was a completed Parent Interview. Subjects were routed to one of two questionnaire paths: **middle school/junior high school** or **senior high school**. The **middle school/junior high school** path was administered to those students who were attending grades 6 through 8. Data were collected about school and family learning environments, activities that may promote personal responsibility and civic involvement, plans for postsecondary education, and community service. The **senior high school path** was administered to students attending grades 9 through 12. Youth in the senior high school path were asked the same questions as those in the middle school/junior high school path but were also given an additional set of questions regarding knowledge and attitudes about civic life. Home schoolers in the middle school/junior high school and senior high school paths who did not attend a school 9 or more hours per week skipped questions referring to schools.

2.4   Adult Education Interview

The Adult Education Interview was designed to provide national estimates of participation in adult educational activities for all adults. Respondents were asked about their participation in six types of educational activities: English as a second language, basic skills/GED preparation, credential courses, apprenticeships, career- or job-related activities, and personal interest/development activities. In addition, other items addressed general literacy activities, community service participation, and opinions about selected actions that might improve public education. The final sections gathered demographic, household, and employment information.

2.5   Authorship of the NHES:1999 Questionnaires

The questionnaires were designed by Mary Jo Nolin, Laura Loomis, and Kwang Kim of Westat, Kathryn Chandler and Christopher Chapman of NCES, and Sean Creighton of ESSI. They received advice and guidance from a Technical Review Panel. Panel members were Suzanne Bianchi, University of Maryland; Jennifer Day, U.S. Department of Commerce; Sandra Hofferth, University of Michigan; Scott Keeter, Virginia Commonwealth University; Oliver Moles, Office for Educational Research and Improvement (OERI); Cynthia Prince, National Education Goals Panel; and Seymour Sudman, University of Illinois.

3. SAMPLE DESIGN AND IMPLEMENTATION

This chapter describes the sample design for the NHES:1999, 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. (For a more detailed discussion of these topics, see *The 1999 National Household Education Survey Methodology Report* (Nolin et al. forthcoming).)

3.1   Precision Requirements for the NHES:1999

The number of telephone numbers required for the NHES:1999 was determined by the precision requirements for the Parent, Youth, and Adult estimates. The key precision requirement for the NHES:1999 Parent and Youth sample was the ability to detect a relative change of at least 10 to 15 percent at the 95 percent confidence level, if at least 30 percent of the children have the characteristic. For the NHES:1999 Adult Education sample, the key sample size determinants were the requirements to detect relative change of at least 10 to 15 percent in estimates of participation in adult education activities (overall) and participation in credential programs and work-related courses, as well as the requirement to estimate participation in adult education by race/ethnicity.

Target sample sizes of about 4,500 infants (age 0 to 2 years), 4,500 preschoolers (age 3 through 5 and not yet enrolled in kindergarten), 10,000 younger children (grades kindergarten through 5), and 10,800 older children (grades 6 through 12) were established for the NHES:1999. A target sample size of 8,150 was established for students responding to the Youth Interview. For adults, a target sample size of about 10,000 civilian adults ages 16 and older who were not in 12th grade or below was established. Adult education participants were sampled at a higher rate than non-participants in order to improve the precision of estimates of characteristics of participants. Taking into account all stages of sampling and expected response, an estimated total of 60,000 households needed to be screened.

3.2   Sampling Households

Different methods have been developed over the years for selecting random samples of telephone households. The Mitofsky-Waksberg method of random digit dialing as described in Waksberg (1978) is probably the best known of the methods. For the NHES:1991 and the NHES:1993, a modified Mitofsky-Waksberg method described by Brick and Waksberg (1991) was used.

Because the NHES is a telephone survey, coverage bias resulting from differences between telephone and nontelephone households is a concern. Various studies have been undertaken to examine the undercoverage bias for key subgroups in the NHES. Brick et al. (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 et al. (1997).

The NHES:1995, the NHES:1996, and the NHES:1999 used a different approach to RDD sampling, called a list-assisted method, described by Casady and Lepkowski (1993). This method reduces the number of unproductive calls to nonworking or nonresidential numbers (compared with simple random sampling of all numbers), produces a self-weighting sample, is a single stage and unclustered sample, and eliminates the sequential difficulties associated with the Mitofsky-Waksberg method. 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) were not sampled.

The disadvantage of the list-assisted method is that it incurs a coverage bias because 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 by Brick et al. (1995b) to address the question of coverage bias. These results show that the percentage of telephone numbers in the zero-listed stratum that are residential is very small (about 1.4 percent) and that about 3 to 4 percent of all telephone households are in the zero-listed stratum. The findings also show that the bias resulting from excluding the zero-listed stratum is generally small.

The sampling frame for the NHES:1999 was all telephone numbers in 100-banks with one or more listed telephone numbers as of December 1998. A stratified list-assisted sample was used in order to support design goals for national-level and subdomain statistics for the Parent, Youth, and Adult Education components of the NHES:1999.

In order to produce more reliable national estimates from the extended interviews for subdomains defined by race and ethnicity, telephone numbers in areas with high percentages of blacks and Hispanics were sampled at higher rates. The sampling frame used in the study contained the 1990 census 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 twice that of the low minority stratum.

A sample of 163,265 telephone numbers was selected for the NHES:1999. Assuming that 49 percent of the telephone numbers would belong to households and assuming a Screener response rate of 75 percent, it was expected that about 60,000 screening interviews would be completed. The actual number of households with completed screening interviews was 55,929.

3.3   Sampling Within Households

Because administering the Parent, Youth, and the Adult Education Interviews in the same household would have been too great a burden on respondents, a within-household sampling scheme was developed to control the number of persons sampled for extended interviews in each household. The sample of telephone numbers was randomly divided into three groups. The first group (21,279 telephone numbers or 13 percent of the sample) was designated for adult enumeration. The second group (20,871 telephone numbers or 13 percent of the sample) was designated for adult enumeration only if there were no eligible children in the household. The third group (121,115 telephone numbers or 74 percent of the sample) was designated for no adult enumeration.

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 Parent Interviews were conducted with parents/guardians of sampled children from birth through 20 years old who were in 12th grade or below. In households with one or more preschoolers, one child in this age/grade range in the household was sampled. The sampling of infants, younger children, older children, and adults was conducted using an algorithm designed to attain the sampling rates required to meet the target sample sizes while minimizing the number of interviews per household. The within-household sample size was limited to three eligible children (if no adults were to be selected), or two eligible children and one eligible adult. No more than one child from any given domain (i.e., infants, preschoolers, younger children, older children) was sampled in any given household. This sampling algorithm was designed to limit the amount of time

required to interview parents in households with a large number of eligible children. Table 3-1 gives the expected and actual response rates and numbers of completed interviews for each of the NHES:1999 interview components.

Table 3-1.—Expected and actual numbers of completed interviews and weighted response rates for the NHES:1999 Screener and interview components

	Expected		Actual	
	Number of completed interviews	Response rate (percent)	Number of completed interviews	Response rate (percent)
Screener	60,000	75.0	55,929	74.1
Parent Interview	27,342	67.5	24,600	65.4
Youth Interview	8,150	56.3	7,913	56.3
Adult Education Interview	8,182	60.0	6,697	60.4

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1999.

Estimates from the October 1996 Current Population Survey (CPS) indicated that 37 percent of all households have at least one child age 20 or below and enrolled in 12th grade or below. Using the within-household sampling algorithm developed for the NHES:1999, the 60,000 screened households should have yielded sample sizes of 4,732 infants, 4,585 preschoolers, 10,198 younger children (enrolled in kindergarten through 5th grade), and 10,867 older children (enrolled in 6th through 12th grade). Assuming a Parent Interview completion rate of 90 percent, the expected number of completed Parent Interviews was 27,342: 4,258 for infants, 4,126 for preschoolers, 9,178 for younger children, and 9,780 for older children. The actual number of completed Parent Interviews was 24,600, with 3,378 for infants, 3,561 for preschoolers, 8,513 for younger children, and 9,148 for older children. The difference between the expected and actual numbers of completed interviews was mainly due to the completion of fewer Screeners than expected, primarily because of a lower than expected residency rate.

All children in grades 6 through 12 whose parents completed a Parent Interview about the child were selected for a Youth Interview. The 10,867 older children expected to be sampled for the Parent Interview were expected to yield 8,150 completed Youth Interviews, assuming a completion rate of 75 percent. The actual number of completed Youth Interviews, 7,913, was lower than expected due to the lower than expected number of completed Screeners.

The Adult Education Interview may be used to generate estimates for all civilian, noninstitutionalized persons age 16 or older and not enrolled in grade 12 or below. Based on the sampling algorithm, the 60,000 screened households were expected to yield 10,227 sampled adults. Assuming an Adult Education Interview completion rate of 80 percent, the expected number of completed Adult Education Interviews was 8,182. The actual number of completed Adult Education Interviews was 6,697. The difference between the expected and observed numbers of interviews was due primarily to the lower than expected number of completed Screeners. Although the sample yield for adults was lower than expected, the lower yield did not affect the ability to detect differences between 1995 and 1999 in key statistics for which the sample was designed. (These key statistics were overall participation in adult education, participation in credential programs and work related courses, as well as participation by type of adult education and participation by blacks and Hispanics.) However, the reduction in sample size may affect the ability to detect differences in other statistics that were not used to design the sample.

3.4   Weighting Procedures

The objective of the NHES:1999 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 Current Population Survey (CPS) to achieve this goal. Beginning in 1994, the CPS weights were adjusted to totals of the number of persons that were adjusted to account for the undercoverage from the 1990 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. More complete details are presented in *The 1999 National Household Education Survey Methodology Report* (Nolin et al. forthcoming).

3.4.1   Household-Level Weights

The primary purpose of the Screener in the NHES:1999 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 collected during the extended interview. Thus, 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 a factor to reflect the subsampling for followup of no-answer telephone numbers that did not have available addresses.[1] This weight was adjusted for Screener nonresponse and then adjusted 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. (See chapter 4 for the results of the CHAID analysis.) These characteristics, which were primarily geographic characteristics, 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 3-2 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.

---

[1]An Adult Special Study was conducted in conjunction with the NHES:1999. (For details, refer to *The 1999 National Household Education Survey Methodology Report* (Nolin et al., forthcoming).) Telephone number level adjustments for Screener nonresponse and multiple telephone numbers in household were applied to the combined NHES:1999 Main Study and Adult Special Study samples. However, the household-level poststratification adjustment and all person-level adjustments were conducted for the two samples independently. Therefore, the presence of the Adult Special Study sample in the telephone number level adjustments is expected to have little or no effect on the final person-level weights.

Table 3-2.—Control totals for poststratifying the NHES:1999 household-level weights

Census region*	Children under 18 in household	Control total
Northeast	Yes	6,874,618
Northeast	No	12,946,397
South	Yes	13,725,789
South	No	22,870,844
Midwest	Yes	8,974,914
Midwest	No	15,284,737
West	Yes	8,587,220
West	No	13,319,699
TOTAL		102,584,218

*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 (CPS), March 1998.

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.

Person Weights for the Parent Interview

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 responded to the Parent Interview. Because sampling eligibility was defined in terms of the data collected in the Screener, the weighting procedures were developed with possible misclassification of children according to grade—resulting in a change in interview path—taken into account so that the estimates would not incur bias due to misclassification.

The first step in developing the person weights for the Parent Interview was to account for the probability of sampling the child's domain in the given household. For example, if there was one preschooler, one younger child (enrolled in kindergarten through 5th grade), and one older child (enrolled in 6th through 12th grade), then the preschooler was sampled with certainty and either the younger child or the older child was sampled, with each of these domains assigned a probability of 0.5; the domain sampling adjustment factor for the preschooler was one, and the factor for either the younger or older child (whichever was sampled) was 2. The second adjustment accounted for the probability of sampling the child from among all eligible children in the given domain. For example, if there were three preschoolers in the household, then one was sampled and the adjustment was 3, which is the reciprocal of the probability of selecting the child from among all children in that domain. The application of these two adjustments to the household weight created a person-level base weight for the Parent Interview.

The next step involved adjusting the person-level base weight for nonresponse to the Parent Interview. Nonresponse adjustment cells were created using age/grade combinations: age 0, age 1, age 2, preschool, kindergarten, and grade 1 through grade 12; enrolled children with no grade equivalent were included in the cell containing the modal grade for their age. For each cell, the ratio of the weighted number of eligible sampled children to the weighted number of responding children was then computed. This ratio was multiplied by the person-level base weight to create the nonresponse-adjusted person-level Parent Interview weight.

The final stage of weighting for the Parent Interview involved raking the nonresponse-adjusted person-level weights to national totals obtained using the percentage distributions from the October 1997 CPS and the total number of children from the March 1998 CPS. The October 1997 CPS contains variables not available on the March 1998 CPS, but the totals in the latter are more current. In the procedure used in the NHES:1999, the control total for a raking cell is the proportion in that cell from the October 1997 CPS multiplied by the estimate of the total number of children from the March 1998 CPS. The three raking dimensions 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 age or grade of child (with those enrolled in school but having no grade equivalent assigned to the modal grade for their age; that is, they were assigned to the grade that most children their age are enrolled in). 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). Table 3-3 shows the control totals used for raking the Parent Interview weights.

Person Weights for the Youth Interview

The person-level weight for the Youth Interview used the final person-level weight for the Parent Interview as the base weight. Two adjustments were made to produce the person-level weight for the Youth Interview. The first adjustment was for children for whom the Parent Interview was completed but the Youth Interview was not completed. Adjustment cells were created for each grade of 6 through 12; enrolled children with no grade equivalent were included in the cell containing the modal grade for their age. For each cell, the ratio of the weighted number of sampled children with completed Parent Interviews to the weighted number of older children responding to the Youth Interview was then computed. This adjustment was applied to the Youth Interview base weight to create the nonresponse-adjusted person-level Youth weight.

The second step was to adjust the nonresponse-adjusted Youth weight to national control totals using a raking procedure. The control totals were computed using the percentage distributions of older children from the October 1997 CPS and the total number of children from the March 1998 CPS. The three raking dimensions 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 grade of child (for children who did not have a grade equivalent, the modal grade for their age was assigned). 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). Table 3-4 contains the control totals used for raking the Youth Interview weights.

Table 3-3.—Control totals for raking the NHES:1999 person-level Parent Interview weights

Race/ethnicity of child	Household income	Control total
Black, non-Hispanic	$10,000 or less	3,262,444
Black, non-Hispanic	$10,001-$25,000	3,665,320
Black, non-Hispanic	$25,001 or more	4,285,309
Hispanic	$10,000 or less	2,366,399
Hispanic	$10,001-$25,000	4,390,088
Hispanic	$25,001 or more	4,068,648
Other	$10,000 or less	3,529,724
Other	$10,001-$25,000	7,684,402
Other	$25,001 or more	38,899,073

Census region*	Urbanicity	Control total
Northeast	Urban	11,631,543
Northeast	Rural	3,108,352
South	Urban	17,016,507
South	Rural	7,771,589
Midwest	Urban	12,407,797
Midwest	Rural	4,902,234
West	Urban	13,208,312
West	Rural	2,105,073

Home tenure	Age/grade of child	Control total
Rent	Age 0	1,612,771
Rent	Age 1	1,643,573
Rent	Age 2	1,603,392
Rent	Age 3–6, not enrolled	1,896,297
Rent	Nursery/Preschool/Head Start	1,526,931
Rent	Transitional kindergarten/ Kindergarten/Pre-1st grade	1,471,803
Rent	1st grade	1,703,358
Rent	2nd grade	1,390,018
Rent	3rd grade	1,437,235
Rent	4th grade	1,195,640
Rent	5th grade	1,231,200
Rent	6th grade	1,263,620
Rent	7th grade	1,116,492
Rent	8th grade	1,071,239
Rent	9th grade	1,174,732
Rent	10th grade	1,093,938
Rent	11th grade	919,073
Rent	12th grade	897,639

Table 3-3.—Control totals for raking the NHES:1999 person-level Parent Interview weights—
Continued

Home tenure	Age/grade of child	Control total
Own or other	Age 0	2,226,665
Own or other	Age 1	2,283,331
Own or other	Age 2	2,329,237
Own or other	Age 3–6, not enrolled	2,143,463
Own or other	Nursery/Preschool/Head Start	2,990,042
Own or other	Transitional kindergarten/ Kindergarten/Pre-1st grade	2,476,210
Own or other	1st grade	2,836,843
Own or other	2nd grade	2,692,744
Own or other	3rd grade	2,638,868
Own or other	4th grade	2,752,372
Own or other	5th grade	2,767,977
Own or other	6th grade	2,766,607
Own or other	7th grade	2,876,585
Own or other	8th grade	2,698,162
Own or other	9th grade	2,880,556
Own or other	10th grade	2,932,074
Own or other	11th grade	2,771,081
Own or other	12th grade	2,839,639
TOTAL		72,151,408

*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 (CPS), March 1998. U.S. Department of Commerce, Bureau of the Census, Current Population Survey (CPS), October 1997.

Table 3-4.—Control totals for raking the NHES:1999 person-level Youth Interview weights

Race/ethnicity of child	Household income	Control total
Black, non-Hispanic	$10,000 or less	979,615
Black, non-Hispanic	$10,001-$25,000	1,397,009
Black, non-Hispanic	$25,001 or more	1,791,803
Hispanic	$10,000 or less	739,271
Hispanic	$10,001-$25,000	1,342,728
Hispanic	$25,001 or more	1,460,141
Other	$10,000 or less	1,103,620
Other	$10,001-$25,000	2,667,179
Other	$25,001 or more	15,820,073

Census region*	Urbanicity	Control total
Northeast	Urban	4,401,270
Northeast	Rural	1,176,172
South	Urban	6,438,892
South	Rural	2,940,699
Midwest	Urban	4,694,998
Midwest	Rural	1,854,961
West	Urban	4,997,906
West	Rural	796,541

Home tenure	Grade of child	Control total
Rent	6th	1,263,620
Rent	7th	1,116,492
Rent	8th	1,071,239
Rent	9th	1,174,732
Rent	10th	1,093,938
Rent	11th	919,073
Rent	12th	897,639
Own or other	6th	2,766,607
Own or other	7th	2,876,585
Own or other	8th	2,698,162
Own or other	9th	2,880,556
Own or other	10th	2,932,074
Own or other	11th	2,771,081
Own or other	12th	2,839,639
TOTAL		27,301,439

*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 (CPS), March 1998. U.S. Department of Commerce, Bureau of the Census, Current Population Survey (CPS), October 1997.

Person Weights for the Adult Education Interview

Four adjustments were made to the household-level weight to compute the person-level weight for the Adult Education Interview. The first adjustment accounted for the probability of sampling adults in the household. 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. For example, if there were no eligible children in the household and there were two eligible adults—one adult education participant and one non-participant—then an adult was sampled with probability 0.258. In such an example, if an adult was sampled, then the domain sampling adjustment factor for that adult was 3.876, which is the reciprocal of the probability of sampling the adult domain. The second adjustment was used to account for the probability of selecting the adult from among all adults in the household. Persons identified by the Screener respondent as adult education participants were given twice the probability of selection of persons identified as non-participants. In the above example, the adult education participant had a probability of selection equal to two-thirds, and the adult education non-participant had a probability of selection equal to one-third. If the adult education participant was selected, then the adjustment factor was 1.5; if the adult education non-participant was sampled, then the adjustment factor was 3. The application of these two adjustments to the household weight created a person-level base weight for the Adult Education Interview.

The third adjustment for the Adult Education Interview person-level weights was the nonresponse adjustment. Three variables were used to create the nonresponse adjustment cells. The first was the sex of the adult, the second was the adult education participation status of the adult (as reported by the Screener respondent), and the third was an indicator of whether the sampled adult was the Screener respondent. These variables were used because they are available for all sampled adults (both respondents and nonrespondents) and were associated with Adult Education Interview response propensity. Within each cell, the ratio of the weighted number of sampled adults to the weighted number of responding adults was computed and used to create the nonresponse-adjusted person-level weight.

The nonresponse-adjusted weight was adjusted in the final step to national totals using a raking procedure. The control totals were obtained from the March 1998 CPS. The four dimensions for the raking cells were a cross of the adult's race/ethnicity (black, non-Hispanic/Hispanic/other) and household income ($10,000 or less/$10,001-$25,000/$25,001 or more), a cross of age (16–29 years/30–49 years/50 years or more) and sex, a cross of Census region (Northeast/South/Midwest/West) and urbanicity (urban/rural), and a cross of home tenure (rent/own or other) and highest educational attainment (less than high school diploma/high school diploma or equivalent/some college). These raking dimensions were used because they include important analysis variables (e.g., educational attainment) and characteristics that have been shown to be associated with telephone coverage (e.g., race/ethnicity). The control totals used for raking the Adult Education Interview person-level weights are given in table 3-5.

3.5   Computing Sampling Errors

In surveys with complex sample designs, such as the NHES:1999, direct estimates of the sampling errors assuming a simple random sample will typically underestimate the variability in the estimates. The NHES:1999 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

Table 3-5.—Control totals for raking the NHES:1999 person-level Adult Education Interview weights

Race/ethnicity	Household income	Control total
Black, non-Hispanic	$10,000 or less	3,261,602
Black, non-Hispanic	$10,001-25,000	5,575,328
Black, non-Hispanic	$25,001 or more	13,291,836
Hispanic	$10,000 or less	2,133,951
Hispanic	$10,001-25,000	5,267,484
Hispanic	$25,001 or more	12,090,037
Other	$10,000 or less	8,939,081
Other	$10,001-25,000	26,706,488
Other	$25,001 or more	117,359,401

Age	Sex	Control total
16–29 years	Male	20,621,334
16–29 years	Female	21,256,430
30–49 years	Male	40,852,187
30–49 years	Female	42,157,257
50 years or more	Male	31,663,625
50 years or more	Female	38,074,374

Census region*	Urbanicity	Control total
Northeast	Urban	31,375,568
Northeast	Rural	8,384,642
South	Urban	45,901,269
South	Rural	20,963,514
Midwest	Urban	33,469,480
Midwest	Rural	13,223,559
West	Urban	35,628,833
West	Rural	5,678,341

Home tenure	Highest educational attainment	Control total
Rent	Less than high school diploma	12,800,580
Rent	High school diploma or equivalent	17,837,673
Rent	Some college	24,341,621
Own or other	Less than high school diploma	20,542,819
Own or other	High school diploma or equivalent	47,963,995
Own or other	Some college	71,138,520

TOTAL		194,625,207

*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 (CPS), March 1998.

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 the NHES:1991–1999.

A total of 80 replicates were defined for the NHES:1999 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 the NHES:1999 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 Parent Interview file as FPWT1 through FPWT80. In the Youth Interview file, they are FYWT1 through FYWT80, and in the Adult Education Interview file, they are FAWT1 through FAWT80. The computation of the sampling errors using these replicate weights can be done easily using the SAS procedure WESVAR or the Windows-based software package WesVar Complex Sample Software; with either of these packages, the replication method should be specified as JK1. The current version of WesVar Complex Samples is available from SPSS. Additional information can be obtained at http://www.spss.com. A previous version of WesVarPC (version 2.12) is available free of charge at http://www.westat.com or by sending an e-mail message to wesvar@westat.com. Please note that version 2.12 of WesVarPC is no longer being updated or revised.

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 (Shah et al. 1995), 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. On the Parent Interview file, the PSU and stratum variables are called PPSU and PSTRATUM; on the Youth Interview file, they are YPSU and YSTRATUM; and on the Adult Education Interview file, they are APSU and ASTRATUM. 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 Youth Interview file, use YSTRATUM YPSU in the NEST statement. (Information on obtaining SUDAAN can be found at http://www.rti/patents/sudaan/sudaan.html.)

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 Youth Interview file, use the following commands to specify these design parameters:

svyset strata ystratum
svyset psu ypsu
svyset pweight fywt

The full sample weight to be used for analysis of the Parent Interview file is FPWT. For the Youth Interview file, the full sample weight is FYWT. For the Adult Education Interview file, the full sample weight is FAWT.

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.

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 below.

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 the NHES:1999.

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 the NHES:1999, the DEFT is typically greater than one due to the clustering of the sample and the differential weights attached to the observations. In the NHES:1999 both of these factors contributed to making the average DEFT greater than one. A fuller description of these factors for the NHES:1999 is given *The 1999 National Household Education Survey Methodology Report* (Nolin et al. forthcoming). (See also appendix B for the DEFT for each data file of the NHES:1991–1999.)

The average DEFT computed for estimates in the three interviews in the NHES:1999 ranged from 1.2 to 1.4. For the Parent file estimates, the average DEFT was 1.3 overall**.** It did not vary appreciably for estimates by path of child (grouped as infant, preschooler, younger child, older child, or home schooler) or by race/ethnicity. Therefore, a DEFT of **1.3** is recommended to approximate the standard error of the estimates in the Parent Interview file.

The average DEFT for estimates in the Youth file is also 1.3, and this does not vary for estimates by path of student (grouped in grades 6 though 8 and 9 through 12) or by race/ethnicity. Therefore, a DEFT of **1.3** is recommended to approximate the standard error of the estimates from the Youth Interview file.

For estimates from the Adult Education file, the average DEFT is 1.3. For estimates by race/ethnicity, the average DEFT ranges from 1.2 to 1.3. For estimates of characteristics of Adult Education participants the average DEFT is 1.4. Therefore, for estimates of the characteristics of the adult population as a whole a DEFT of **1.3** could be used to approximate the standard error; for estimates of characteristics of adult education participants, a DEFT of **1.4** is recommended; and for estimates of characteristics of white, non-Hispanic adults, a DEFT of **1.2** is recommended.

As stated above, the average DEFT can be used to approximate the standard error for an estimate. An example of how to do this on a **percent** estimate is as follows. If a weighted estimate of 46 percent is obtained for some characteristic in the Adult Education file (suppose that 46 percent of adults participated in Adult education activities, excluding full-time credential programs), 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 46 percent statistic would be 0.61 percent (the square root of

(46*54)/6,697, where the weighted estimate (p) is 46 percent, 54 is 100 minus the estimated percent (1-p), and the unweighted sample size (n) is 6,697.) The approximate standard error of the estimate from the NHES:1999 is this quantity (the simple random sample standard error) multiplied by the DEFT for the Adult Education file estimates of 1.3. In this example, the estimated standard error would be 0.79 percent (1.3 x 0.61 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 Youth 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 standard error from the unweighted analysis is multiplied by the mean DEFT for the Youth file estimates of 1.3 to approximate the standard error of the estimate under the NHES:1999 design. For example, suppose that the estimated (weighted) mean number of hours per week worked by students in grades 6 through 12 (among those who worked) was 12 and the simple random sampling standard error (unweighted) was 10 hours. Then, the approximate standard error for the estimate would be 13 hours (10 hours x 1.3).

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 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:1999 design. For example, if a given parameter in a model involving items from the Parent Interview file has a weighted estimate of 2.33 and a standard error of 0.45, then the adjusted standard error would be 1.3 x 0.45 = 0.59.

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 the Adult Education file, sum the final weights for all 6,697 cases on the file. Next, divide this sum by the number of cases to generate an average final weight. (In the above example, the number of cases is 6,697). 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 1.3, or 1.69.) 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 1.69 and the average final weight.) 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 would be affected by small differences in the estimated standard errors.

3.7   Standardization of Weights for the Split Half Samples and Other Subsamples

In the NHES:1999, two versions of a five-item knowledge of government test were developed and administered to split half samples of youth. The first set (Set A), comprising the items YE8a-e (CYVP, CYLAW, CYHOUSE, CYVETO, and CYCONSRV) on the Youth file was administered to a random subsample of about half the respondents, and the second set (Set B), comprising the items YE9a-e (CYVP, CYJUDGE, CYSENATE, CYCONST, and CYDFENS) on the Youth file was administered to the remaining respondents. The half samples were randomly determined by the telephone number of the household. One set of questions was administered in households with telephone numbers ending in an even digit, and the other set was administered in households with telephone numbers ending in an odd digit.

In the NHES:1999 Youth Interview, a set of detailed follow-up questions were asked of a subsample of youth who participate in service activities. The follow-up questions are the following items, which appear as items YG2 through YG8 in the Youth Interview questionnaire: SAPYMT1-3, SAARNG1-3, SAHELP1-3, SAHCHIL1-3, SAHADLT1-3, SAHELDR1-3, SARELA1-3, SAPOOR1-3, and SADISB1-3. Additionally, the service activity descriptions were coded for this subsample of youth (BCODE1-3, SPCODEA1-3, SPCODEB1-3, and SPCODEC1-3).

As discussed in Section 3.4, the Parent Interview, Youth Interview, and Adult Education Interview weights have been adjusted to national totals. However, weights for the random subsamples described above were not adjusted separately to national totals. As a result, totals for demographic characteristics for each of these subsamples do not agree with each other and do not match the national totals. If it is desirable for the subsamples to each be nationally representative, then a simple weighting adjustment may be performed using WesVar Complex Samples Software. (For more detail, see WesVar Complex Samples 3.0 User's Guide, 1998.) Hereafter, such an adjustment will be referred to as "standardization."

For users of WesVarPC, the following instructions list the steps required to standardize the weights of the split half sample or subsample to the full sample weighted totals, which may be computed from the weights on the NHES:1999 files Using the terminology of standardization, the demographic variables to which the subsamples are standardized are called the dimensions and the categories of these dimensions are called the levels. An example of standardizing using the two sets of civic knowledge questions in the Youth Interview is given below.

1. For each dimension to be used in the standardization, compute the control totals (*i.e.*, the weighted estimates of totals for the levels of the dimension based on the full sample). For example, suppose the weights are to be standardized to two dimensions: sex (SEX) and grade (ALLGRADE, recoded so that children with no grade equivalent are included in the modal grade for their age). To do so, submit Table requests in WesVarPC using the full-sample data file. Submit one Table request for each dimension (SEX and the recode of ALLGRADE) and generate weighted frequencies by specifying the full sample weight.

2. For each dimension, create an ASCII file containing a field for the level of the dimension variable and a field for the control total, with the two fields separated by a space. Continuing the example, then two ASCII files should be created. The first, say SEXTOTL.DAT, would look like this (generating the control totals from the full sample weight in WesVarPC):

    1 13831839
    2 13469599

    The second, say GRADTOTL.DAT, would look like this (generating the control totals from the full sample weight in WesVar):

    6 4030226
    7 3993078
    8 3769401
    9 4055288
    10 4026012
    11 3690155
    12 3737278

    Note that the sum of the two levels of SEX equals the sum of the seven levels of the control totals of the recoded ALLGRADE. This is a requirement of standardization. The sum of the control totals for the dimensions must be equal.

3.   Subset the file to the particular subsample of interest. For example, on the Youth Interview file, to standardize the weights for the subsample that was administered the civic knowledge questions in Set A, subset the file by extracting only those cases with CYLAW ≠ -1.

4.   Using WesVarPC, choose Import Data File from the Prep menu. Provide the required parameters (see Brick et al. 1995a for details on the specifications for each WesVar screen), and press the Create button. This will create a WesVarPC file containing only the cases administered the Set A questions. (This file will have a .VAR extension.)

5.   Next, from the Prep menu, select Poststratification. Specify the WesVarPC file that was created in the previous step. Specify the first dimension variable as the Cell Identifier, and specify the corresponding file of control totals as the File with Control Totals. Keeping with the last example, one might specify SEX as the Cell Identifier and C:\SEXTOTL.DAT as the File with Control Totals. A new WesVarPC file will be created. For the sake of illustration, suppose this new file is called YUTH_A.VAR.

6.   Repeat the poststratification process, using the second variable as the Cell Identifier and its corresponding control totals file as the File with Control Totals. The input file for this step should be the output file from the previous poststratification step (YUTH_A.VAR). In the example, the recode of ALLGRADE would be the Cell Identifier, and C:\GRADTOTL.DAT would then be the File with Control Totals and the output file might be called YUTH_B.VAR. At this point, the sum of the weights of the file by the recode of ALLGRADE equal the control totals in GRADTOTL.DAT.

7.   Continue this poststratification process until all dimensions have been exhausted. When this occurs, you have completed one iteration of this process. In the example, one iteration will have been completed after one poststratification step has been completed using the recode of ALLGRADE.

8.   Proceed with further iterations of this poststratification process until the new weighted totals converge to the control totals. To check on convergence, submit a set of Table requests after each iteration, with each table producing weighted frequencies for a cell identifier variable (i.e., dimension). In the example, there will be two Table requests: SEX and the recode of ALLGRADE. If the new weighted totals <u>for each dimension</u> match the control totals (to within some tolerable amount of error), then terminate the poststratification process. If the new weighted total <u>for any dimension</u> do not match the control totals, then continue with further iterations. In the example, after each iteration, the new weighted totals for the recode of ALLGRADE will match the control totals, since the recode of ALLGRADE was the last dimension to which the weights were poststratified. However, if the totals for SEX do not match the control totals, then you must proceed with another iteration of poststratification; i.e., poststratify to SEX and then poststratify to the recode of ALLGRADE.

This process will bring the new weighted totals for the particular subsample up to national levels. However, caution should be taken in combining samples. For example, one might use this standardization process on the weights for youth in the subsample receiving the Set A questions, and then also use this

process on the weights for youth in the subsample receiving the Set B questions. In that case, each of the two subsamples is standardized to national totals. However, the standardized weights should be **divided by two** for any analyses where the two subsamples are combined and standardization to national totals is desirable.

3.8   Imputation

In the NHES:1999, 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 for the item 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 the NHES:1999, the item response rate was very high. The median item response rate for all imputed items from the Parent Interview was 98.96 percent; for the Youth Interview data, 98.41 percent; and for the Adult Education Interview, 99.25 percent. (Item response rates are discussed in more detail in chapter 4.) Despite the high item response rate, data items with missing data on the file were imputed with the exception of the nine knowledge about government items. (For those items, answers of "don't know" or "refused" are regarded as incorrect responses. As such, they 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 complete responses 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.

The standard set of sort order variables for the household-level items collected in the Parent, Youth, and Adult Education Interviews consisted of:

- ■   CENREG -- the Census region in which the household was located;

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

- ■   KIDINHH -- a variable derived for imputation from the age (AGE) of household members indicating whether or not children under age 18 resided in the household; and

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

The standard sort order variables for the person-level items on the Parent Interview file were:

- ■   MAINRSLT -- the final completion code for the interview;

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

- ■   SEX -- sex of the sampled child;

■ PARGRADS -- a variable that indicates the highest education level attained by either parent in the household as less than high school diploma, high school diploma but no bachelor's degree, or college graduate derived from MOMGRADE, MOMDIPL, DADGRADE, and DADDIPL; and

■ HHPARNS -- a variable derived for imputation from HHMOM and HHDAD indicating whether there were two parents in the household or not.

The standard sort order variables for the person-level items from the Youth Interview were:

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

■ SPUBLIC -- whether the sampled child attends a public or private school;

■ SEX -- sex of the sampled child; and

■ PARGRADR -- a variable derived for imputation that indicates the highest education level attained by either parent in the household as less than high school diploma, high school diploma but no bachelor's degree, or college graduate; derived from MOMGRADE, MOMDIPL, DADGRADE, and DADDIPL.

The standard sort order variables for the person-level items from the Adult Education Interview file were:

■ PARTIC -- a variable derived for imputation that indicates whether the adult participated in any Adult Education activities (including full-time credential) in the last year;

■ EDUC -- a variable derived for imputation that indicates whether or not the adult has at least a high school diploma or the equivalent;

■ AGECAT -- a variable derived for imputation from AGE for the respondent with the categories 18 through 29 years, 30 through 49 years, and 50 or older;

■ ARACETH -- a variable derived for imputation that classifies the respondent as black, non-Hispanic; Hispanic; or other; and

■ HINCMRNG -- the household income range.

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 (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.

For some items, the missing values were imputed manually rather than using the hot-deck procedure. In the NHES:1999, hand imputation was done (1) to impute certain person-level demographic characteristics; (2) to impute whether a child is home schooled, attends regular school for some classes, and the number of hours attending regular school;[2] (3) to impute variables that involved complex relationships

---

[2] Home schoolers in the Youth CI interview received most of the questions the other youth received and no complex skip patterns were involved; therefore, no special procedures were used for imputing missing values for the home schoolers on the Youth Interview file.

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 hand imputation of the person-level demographic items and of the home schooling items, three sort variables were used: state, the 3-digit ZIP code, and the person identification number, to ensure that all household members were grouped together.

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. There were 51 items in the Parent file with response rates of less than 90 percent, 23 items in the Youth file, and 9 items in the Adult Education file. 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 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 either 1, 2, 3, or 4. The value of the imputation flag indicates the specific procedure used to impute the missing value. The imputation flag was typically set to 1 if the missing value was imputed using the standard hot-deck approach.

The procedure for hot-deck imputation only recognizes missing value codes as those that need to be replaced by imputed values. For the NHES:1999, these missing codes were -7 = refused, -8 = don't know, and -9 = not ascertained. Therefore, in some cases, variables that originally equaled -1 (inapplicable) had to be recoded to a missing value code (i.e., -9 = not ascertained) 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, if the value of SCHOICE (PD2) equaled -8 for a child, then SDISRCT (PD3) was never asked and thus equaled -1 (inapplicable). During the imputation process for this child, if SCHOICE was imputed to equal 2 (chosen), the 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.

For some items with complex skip patterns and only a few missing values, the item was imputed manually and 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. Code 4 was utilized whenever a response of "don't know" was obtained except for the political knowledge questions.

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 NHES:1999, including the use of computer-assisted telephone interviewing (CATI), staff training, interviewer assignments and contact procedures, and quality control. More detailed descriptions of these topics can be found in *The 1999 National Household Education Survey Methodology Report* (Nolin et al. forthcoming).

4.1.1 Special Pre-Collection Procedures

Before the beginning of data collection, special procedures were implemented to remove some nonresidential and nonworking telephone numbers from the sample, and in addition, an advance mailing was conducted.

Business purge. The 163,265 telephone numbers in the NHES:1999 sample were drawn by Genesys, a commercial firm that draws samples to specific requirements. In addition to the NHES: 1999 numbers, an additional 4,082 numbers were drawn by Genesys for the Adult Special Study. These 167,347 telephone numbers were matched by Genesys against residential White Pages listings and Yellow Pages listings. Numbers appearing in both listings were classified as residential listings. Numbers that appeared only in the Yellow Pages business listings were classified as nonresidential. Of the total sample, 7,527 telephone numbers were found to be nonresidential through this business purge and were given a final status of nonresidential. This total includes 7,324 numbers from the main NHES:1999 sample and 203 numbers from the Adult Special Study sample.

Tritone test. Genesys also conducted tritone tests on the remaining 159,821 telephone numbers that were not identified as business numbers or residential numbers through a check of White and Yellow page listings. The numbers were computer dialed; those that received tritone signals (the three-bell sound heard when dialing a nonworking number) prior to a ring on two separate calls on different days were classified as nonworking and assigned that final disposition code. In all, 17,682 telephone numbers were classified as nonworking through the tritone check. This total includes 17,253 numbers from the main NHES:1999 sample and 429 numbers from the Adult Special Study sample.

Advance mailing. In an effort to increase Screener-level response, a mailing was sent to the 51,004 households for which an address was obtained by the commercial firm Telematch.[3] (Because of time constraints imposed by requiring Genesys to draw the sample from the fourth quarter 1998 Marketing Systems Group frame and the time needed for preparing for the advance mailing, all of the telephone numbers in the sample were sent to Telematch for address matching.) A brief letter explaining the purpose of the NHES:1999 and encouraging participation in the study was mailed to these households; however, 7,615 letters were returned by the postmaster, the vast majority containing no forwarding address. (Because additional mailings were planned, an address was updated if a forwarding address was supplied. See section 4.1.5 below for a description of additional mailings.) The field test of the NHES:1996 showed that households receiving an advance letter were more likely to respond to the survey (Brick and Collins 1997). The NHES:1999 result was consistent with those findings. The weighted response rate for households to which letters were mailed in advance of calling was higher than that for households for which addresses were sought but could not be obtained (80.8 percent compared to 66.5 percent).

---

[3]After the Tritone tests, the NHES:1999 and Adult Special Study samples were treated as distinct samples. The remaining discussion of data collection and response rate calculations refers only to the NHES:1999 sample.

4.1.2 CATI System Applications

The use of a CATI system for the NHES:1999 included a number of applications that facilitated the implementation of the survey. Briefly, the most salient features of the CATI system for the NHES:1999 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.

- **Copying Responses:** The CATI system was used to copy responses from one interview to another to prevent unnecessary repetition of questions. For example, when two children with the same parents were sampled in a household, the parent characteristics series and household information items were asked only once. This helped to reduce response burden.

- **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 each item in the screener and extended interviews. These screens, which could be accessed with a keystroke by the interviewer, clarified terminology, explained the intent of questions, and helped the interviewer obtain correct information.

4.1.3 Interviewer Training

Training for interviewers was conducted over a 7-week period, 3 weeks in December 1998 and throughout the month of January 1999. Interviewers were trained in groups of about 30. Inexperienced interviewers received 18 hours of training and experienced interviewers, 14 hours. Prior to the NHES:1999 project training, all interviewers had participated in a basic training in general interviewing techniques and the use of the CATI system that typically lasted 6 hours. Each training group had an assigned time for the first interviewing shift so that their work could be closely monitored. In total, 421 interviewers were trained for the study.

Interviewer training was conducted using the CATI system. The trainees entered information in the CATI system during training presentations, providing them with hands-on experience prior to beginning data collection. The topics covered in the training session included an introduction to the study, interactive lectures based on each of the survey questionnaires, details about survey procedures, and techniques for refusal avoidance. Prior to live interviewing, trainees practiced interviews in pairs using several role-play scripts. The majority of training time (about 12 hours) was spent on interactive lectures and practice interviews using role-play scripts. Most of the remaining time was spent on procedures for contacting households and respondents and techniques for gaining cooperation.

The survey staff included 21 interviewers bilingual in English and Spanish. These interviewers received the same English training as all other interviewers and 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.

4.1.4   Interviewing Procedures

The CATI system scheduled cases automatically, based on an algorithm that was customized for the NHES:1999 survey. 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;

- Initial cases, until they had received one day and one evening call attempt;

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

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

At least seven attempts were made by NHES interviewers to screen households in order to determine the presence of household members eligible for extended interviews, that is, an eligible child or adult. These calls were staggered on different days of the week and at different times of the day over a period of at least 2 weeks. This included at least two daytime calls, three evening calls, and two weekend calls. In addition, nearly all cases for which this seven-call limit was reached were released at several points during data collection to receive additional calling attempts. Some cases received more than 20 calls in this effort to complete screening interviews. Cases that were coded as problems were referred to a telephone supervisor to discuss appropriate methods of completing an interview (e.g., holding a case for some time and releasing it for additional attempts later in the data collection period). Below is an overview of the specific calling strategies used during the NHES:1999 data collection and their results. Because most nonresponse in a random-digit-dialing (RDD) survey occurs at the screening level, these procedures emphasized increasing the Screener response rate. Please see *The 1999 National Household Education Survey Methodology Report* (Nolin et al. forthcoming) for a more detailed account of these procedures and their results.

Procedures for Other than English. The NHES:1999 was conducted primarily in English, but provisions were made to interview persons who spoke only Spanish. As mentioned above, the questionnaires were translated into Spanish, Spanish versions of the CATI instruments were 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." Bilingual interviewers were the only ones who could access these "language problem" cases for follow up. 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.

There were 38 Screeners that were classified by at least one interviewer as a hearing or speech problem. None of these cases were completed.

A total of 2,515 Screeners were classified by the first interviewer making contact as Spanish-speaking. About 72 percent of all these cases were finalized as completes. About 93 percent of these completed cases were completed in Spanish.

Of the 899 Screeners with respondents identified by the initial interviewer as speaking some language other than English or Spanish, about 22 percent were completed. Most were completed in English; less than 0.5 percent were completed in Spanish. Of those that were not completed, 84 percent were finalized as language problems and the rest were refusals (6 percent) or other nonresponse (10 percent).

Answering Machines., Interviewers made at least seven attempts to reach households in which call attempts resulted in an answering machine status in order to complete the screening and determine whether any household members were sampled for interviews. The first time an answering machine was reached, the interviewer left a brief message explaining the nature of the call, providing the 800 telephone number for the prospective respondent, and explaining that an interviewer would call again at a later time. A different message was left upon reaching an answering machine only if the case changed strategy, for instance, became a refusal case. Up to four such messages could have been left in a household if the case changed strategies; however, that would have been a rare occurrence. At the end of the data collection period, a fifth message may have been left in an attempt to gain the cooperation of the household.

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 the NHES:1999. Unless an interviewer indicated that the initial refusal was "hostile" (e.g., profane or abusive), one 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 14-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 close-out while maximizing response rates.

A strategy used in the NHES:1999 to increase the likelihood of successful refusal conversion attempts was to send a letter to households for which a good address had been obtained from Telematch or through a forwarding address from a postmaster. The letters were sent by Federal Express at a special corporate rate in order to draw attention to their importance. A total of 5,765 letters were mailed, and 4,980 of these cases or 90 percent were completed, versus 69 percent of the 7,878 cases to which a letter was not mailed. Taken together, 19 percent of the completed screeners were completed in those households that received refusal conversion letters.

An additional refusal conversion attempt was made in a subset of cases which had twice refused to participate in the Screener interview. The cases included in this effort were those for which neither the first nor second refusal received a code of "hostile." This effort resulted in the completion of 2,794 additional Screeners (or 5 percent of the total completed). All refusals were considered to be final if a third contact with the household resulted in a code of refusal.

Refusal conversion efforts were successful at the extended interview level as well. Nine hundred ninety-two Parent Interviews, 198 Youth Interviews, and 437 Adult Education Interviews were completed after initially refusing the interview. In the NHES:1999, an additional refusal conversion attempt was also made on selected extended interview cases for which two refusals, both coded "mild," had been received.

These efforts also yielded completed interviews; 131 additional Parent Interviews were completed out of 304 cases refielded, 222 additional Youth Interviews were completed out of 471 cases refielded, and 83 additional Adult Education Interviews out of 252 were completed.

In summary, the additional refusal conversion activities for the NHES:1999 were productive. Ninety percent of the Screener refusal cases that were mailed a Federal Express letter and called after the first refusal were completed versus 69 percent of the cases called after an initial refusal but not mailed a Federal Express letter. Twenty-two percent of the cases released for another call attempt after two refusals were completed. Of the extended interviews released for an additional refusal conversion attempt, from 26 to 47 percent of the refielded cases were completed.

"Maximum call" cases. Other efforts to increase the Screener response rate focused on "maximum call" cases, in which a person had answered on at least one of the seven initial attempts. If a good mailing address had been obtained for the household, a first-class letter describing the survey and the importance of cooperating was mailed to it in an 8 1/2 x 11 inch envelope. The cases were held for a period of time, mailed to if possible (1,974 cases were mailed letters), and released for additional attempts periodically during the data collection period. Initially, only maximum call cases that had never been released for additional attempts were released. However, toward the end of the data collection period, all maximum call cases were released for additional call attempts, regardless of how many additional attempts they had already received. This effort resulted in the completion of 2,308 additional Screeners, 42 percent of cases refielded and 4 percent of all completed screeners.

Extended interview "maximum call" cases were also released for additional call attempts. These efforts resulted in the completion of 1,230 Parent Interviews, 337 Youth Interviews, and 404 Adult Education Interviews.

"No answer" calls. Numbers that had been answered only by answering machines and never by a person ("no answer, answering machine" cases) and numbers at which no answer was ever received during the seven attempts, ("no answer" cases) also received special treatment during the NHES:1999 data collection. All "no answer, answering machine" cases for which a good address had been obtained were mailed a first class letter in an 8 1/2 by 11 inch envelope; all "no answer, answering machine" cases were released for additional call attempts. Twenty-six percent of the 5,237 Screener "no answer, answering machine" cases were completed after re-releasing. There was little difference between the 1,872 cases that were mailed a letter (27 percent completed) and the 3,365 cases not mailed to (24 percent completed). The "no answer, answering machine" cases represent 2 percent of all completed screeners.

"No answer" cases with mailable addresses were mailed a first class letter in an 8 1/2 by 11 inch envelope and re-released for additional call attempts as well. However, previous experience has shown that "no answer" cases are the most unproductive cases, and those without a mailable address are likely to be the least productive of that group. Therefore, a subsampling strategy was employed before re-releasing them, so that efforts could be concentrated on the cases more likely to yield completes. Half of the nonmailable "no answer" cases were randomly sampled and released for additional attempts. To reflect this subsampling, the nonmailable cases that were re-released were weighted up to represent the cases that were not re-released. Twelve percent of the 1,286 mailable "no answer" cases were completed versus 3 percent of the 5,606 subsampled, nonmailable "no answer" cases. Together these complete interviews represented less than .05 percent of all completed screeners. However, refielding efforts for all "no answer" and "no answer – answering machine" cases resulted in 1,619 telephone numbers identified as nonworking or nonresidential.

Telephone company business office checks done in 1996 indicated that approximately 40.5 percent of cases finalized with unknown residency status were residential. (See Brick and Broene 1997 for details regarding the use of telephone company business office checks to estimate the residential status of

telephone numbers.) Based on this information, the final no-answer cases were proportionally allocated to residential and nonresidential status in the calculation of final response rates.

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. A live field test was conducted in households between September 24 and September 27, 1998. Four hundred twenty-seven Screeners, 206 Parent Interviews, 67 Youth Interviews, and 61 Adult Education Interviews were completed. The purpose of the field test was to ensure that the CATI system was working properly. Minor modifications to the instruments were made following the field test.

Quality control activities continued during training and data collection. During interviewer training, interviewers were paired with one another, and they conducted 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 occasionally. 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 recurring problems were identified. Additional training was provided as necessary.

Throughout data collection, supervisors and telephone monitors (experienced telephone interviewers who were trained for monitoring) monitored the interviews by listening 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 monitor completed a special monitoring form that covered five major areas of telephone interviewing:

- Reading and general skills;

- Listening skills and probing;

- Recording;

- Handling refusals and questions; and

- Telephone manner and relationship with respondent.

The monitors recorded their impressions of the interviewer's skills and abilities along with suggestions for improvement. Interviewers were 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:1999 interviewing pool. Over 9,000 monitoring sheets were completed for NHES:1999 interviewers. Only four interviewers were released because of inadequate performance.

In addition, at least once a week, the CATI management system produced computer-generated reports that displayed response rates, refusal rates, and refusal conversion rates for each NHES:1999 interviewer. These reports assisted telephone center supervisors in identifying differences in interviewer performance that might not be detected through monitoring. For example, some interviewers may have

been more effective in refusal conversion or gaining cooperation from the elderly. Supervisors relied on both monitoring sheets and standard reports to make staff assignments.

4.2   Response Rates in the NHES:1999

A response rate is the ratio of the number of units with completed interviews (for example, 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 numerator or denominator of the ratio must be estimated.

For reporting the results from the NHES:1999, the response rate indicates the percentage of possible interviews completed taking all survey stages into account, while the completion rate measures the percentage of interviews completed for a specific stage of the survey. For example, household members were identified for interviews in a two-stage process. Screener interviews were conducted to enumerate and sample household members, and then questionnaires were administered for the sampled members. If the responding household member failed to complete the first-stage Screener, no members could be sampled for other interviews. Under this design, the completion rate for the second stage (Parent, Youth, or Adult Education Interview) is the percentage of sampled persons who completed these interviews. The response rate is the product of the first- and second-stage completion rates.[4]

Response and completion 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. Both rates are usually similar unless the probabilities of selection and the response rates in the categories with different selection probabilities vary considerably. All of the response rates discussed in the report are weighted unless noted specifically in the text.

Response rates and completion rates are identical for the first stage of sampling and interviewing (i.e., the Screener). The next section discusses the response rate for the Screener and provides a profile of the characteristics of the respondents. The discussion of response and completion rates for the Parent, Youth, and the Adult Education Interviews are given in the following sections. Additional information on the NHES:1999 response rates, including the findings of additional nonresponse bias analyses, is included in *The 1999 National Household Education Survey Methodology Report* (Nolin et al. forthcoming).

---

[4]Because Youth Interviews could only be conducted after the Parent Interview for that sampled child had been completed, the Youth Interview could be treated as a third survey stage. The Youth Interview response rate would have been calculated as the Screener completion rate times the completion rate times the completion rate of Parent Interviews for 6th through 12th grade students sampled for the Youth Interview times the Youth Interview completion rate. Because the completion rate of Parent Interviews for 6th through 12th grade students sampled for the Youth Interview was not developed, this approach was not taken.

4.2.1 Screener Response Rate

The first panel of table 4-1 shows the disposition of the 163,265 telephone numbers that were sampled for the NHES:1999. 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.

As shown in the lower part of the table, the first weighted response rate of 74.1 percent for the screener was calculated using the business office method. For this method, the numerator is the weighted number of telephone numbers in households that participated in the survey (55,929). The denominator is the weighted total number of residential telephone numbers plus the 40.5 percent of numbers with unknown residential status that are assumed to be residential. Both the numerator and the denominator have been weighted by the probabilities of selecting the telephone numbers and weighted for the subsampling for extensive followup nonmailable no-answer telephone numbers that were not refielded. The 40.5 percent estimate was based on a special study conducted at the conclusion of the NHES:1995 survey in which telephone business offices were contacted to provide the status of a sample of telephone numbers that had unresolved residency status (Brick and Broene 1997).

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

Screener response category	Number	Percentage of all numbers	Percentage of residential numbers
Total	163,265	100.0	
Identified as residential	70,621	43.3	100.0
Responded	55,929	34.3	79.2
Did not respond	14,692	9.0	20.8
Identified as nonresidential	79,054	48.4	
Unknown residential status	13,590	8.3	

Estimated screener response rates*	Weighted rate (percent)*	Unweighted rate (percent)*
Business office method response rate	74.1	73.5
Survival analysis response rate	76.1	75.7
CASRO response rate	73.2	72.6
Conservative response rate	67.4	66.4
Liberal response rate	79.4	79.2

*All of the 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 estimated response rate using the business office method, the proportion of unknown residential status numbers included in the denominator was based upon the proportion identified in checks with telephone business offices. For the survival analysis method response rate, the proportion of unknown residential status numbers included in the denominator was estimated using survival analysis methods that incorporate information about the cases (including listed status, interviewers' coding of answering machine call results, and the number of call attempts the telephone number received). For the CASRO (Council of American Survey Research Organizations) 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 response rate, all of the unknown residential status numbers were included. For the liberal response rate, none of the unknown residential status numbers were included.

NOTE: Because of rounding, details may not add to totals.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1999.

Other estimates of the response rates were computed by allocating different proportions of the numbers with unknown residency status into the residential category. The footnote to table 4-1 explains five different schemes for estimating the response rate. It is reasonable to say that the Screener response rate is

between 67 and 79 percent. The variability in the estimates arises because it is not possible to identify precisely the residential status for each telephone number. The response rate calculated by the business office method has traditionally been reported as the NHES response rate. Using this approach, the best estimate for the NHES:1999 response rate is 74 percent.

The lower right part of table 4-1 also shows unweighted screener response rates calculated using each of the approaches described above. If the raw count of telephone numbers had not been weighted, the Screener response rate using the business office method would have been 73.5 percent.

Table 4-2 presents the Screener 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 response rate for listed residential telephone numbers was higher than for telephone numbers that were not listed. The response rate was also higher for telephone numbers with mailable addresses than for those without mailable addresses. The Screener response rate also varied somewhat by region of the country. Response rates for the Northeast and West were lower than for the Midwest and South. Areas with lower proportions of renters and lower proportions of college graduates had higher response rates than those with higher proportions. Response rates for households not located in an MSA were higher than for households in other areas.

4.2.2   Extended Interview Response Rates

The number of children enumerated and sampled, and those with completed interviews for each component of the NHES:1999, are given in table 4-3. Of the enumerated 38,993 children eligible for sampling in the Parent Interview, a sample of 28,011 children was selected. About 0.5 percent of the sampled children were not actually in the age and grade range for the survey as determined by the Parent Interview respondent. These children were classified as ineligible. Complete interviews were obtained for 24,600 of the sampled children for an estimated 88 percent completion rate and a response rate of 65 percent. The bulk of the unit nonresponse for the Parent Interview was due to refusal of the parent/guardian to respond (64.2 percent of nonresponse). Other reasons for Parent Interview nonresponse were inability to make contact with the parent/guardian (27.5 percent of nonresponse), language problems (2.5 percent of nonresponse), and other miscellaneous reasons such the parent/guardian being unavailable for an interview during the field period (5.9 percent of nonresponse).

The number of older children enumerated, sampled, and the final status of each sampled child for the Youth Interview are also given in table 4-3. About 68 percent of the 15,563 enumerated older children were sampled for the Youth Interview. Less than 2 percent of the sampled older children were classified as ineligible because the Parent Interview respondent reported that they were not actually enrolled in grades 6 through 12. In all, 7,913 interviews were completed with the sampled youth. The estimated completion rate for the Youth Interview is 76 percent and the response rate is 56 percent. The main reason for Youth Interview nonresponse was failure to complete the Parent Interview (54.2 percent of Youth Interview nonresponse). Other reasons for nonresponse to the Youth Interview were the refusal of the parent to permit the youth to respond to the Youth Interview (17.7 percent of Youth Interview nonresponse), refusal of the youth to respond (16.1 percent of Youth Interview nonresponse), inability to make contact with the sampled youth (7.2 percent of Youth Interview nonresponse), language problems (0.5 percent of Youth Interview nonresponse), and other miscellaneous reasons for nonresponse such as the youth not being capable of responding to the interview for health reasons (4.4 percent).

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

Characteristic	Total	Residential, responded	Residential, did not respond	Non- residential	Unknown residential status	Estimated response rate* (percent)
Total	163,265	55,929	14,692	79,054	13,590	74.1
Census region						
Northeast	29,537	10,003	3,212	13,252	3,070	70.1
South	61,684	21,646	5,311	30,024	4,703	75.5
Midwest	33,920	11,420	2,634	17,446	2,420	76.7
West	38,124	12,860	3,535	18,332	3,397	72.8
Listed status						
Not listed	108,794	25,730	8,070	63,751	11,243	67.5
Listed residential number	47,147	30,199	6,622	7,979	2,347	80.4
Listed business number	7,324	0	0	7,324	0	NA
Advance mailing status						
Mailable address, mailing sent	43,389	28,455	5,711	7,360	1,863	81.9
No mailable address	112,261	24,633	8,125	68,472	11,031	66.5
Mailable address, but returned	7,615	2,841	856	3,222	696	70.8
Minority concentration						
High	81,454	26,855	7,523	40,431	6,645	72.7
Not high	81,811	29,074	7,169	38,623	6,945	74.7
Percent college graduates						
Less than 20 percent	63,993	22,752	5,226	32,056	3,959	77.8
20 to 29 percent	50,939	17,966	4,964	23,628	4,381	73.5
30 to 39 percent	25,171	8,442	2,420	11,801	2,508	71.7
40 to 59 percent	20,100	6,031	1,833	9,922	2,314	69.1
60 percent or more	3,062	738	249	1,647	428	65.1
Percent renters						
Less than 20 percent	4,243	1,502	390	1,990	361	74.0
20 to 39 percent	76,617	28,339	6,797	35,711	5,770	75.8
40 to 49 percent	35,219	12,665	3,278	16,318	2,958	74.3
50 to 59 percent	20,370	6,413	1,792	10,376	1,789	72.2
60 to 69 percent	10,820	3,001	940	5,880	999	69.2
70 to 79 percent	8,863	2,358	876	4,782	847	66.3
80 percent or more	7,133	1,651	619	3,997	866	62.7
Metropolitan status						
In county in central city	67,562	21,970	6,305	33,320	5,967	72.3
In county not in central city	29,074	10,424	2,960	13,198	2,492	73.0
Subcounty of MSA	29,299	10,542	2,748	13,464	2,545	73.8
MSA in its own county	7,842	2,554	889	3,546	853	66.4
Not MSA	29,488	10,439	1,790	15,526	1,733	80.9

*The estimated response rate is the number of completed interviews divided by the sum of the number of completed interviews, nonresponses, and 40.5 percent of the unresolved telephone numbers, weighted by the probability of selection.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1999.

The bottom section of table 4-3 gives the numbers of adults enumerated and sampled, and the final status of the Adult Education Interview for sampled adults. Adults were enumerated in only a subsample of households. Of the 20,266 enumerated adults, 8,114 were sampled for Adult Education Interviews. A total of 6,697 adults completed the Adult Education Interview. The estimated completion rate for the Adult Education Interview is 82 percent and the response rate is 60 percent. Almost all of those sampled were eligible for the interview; those classified as ineligible were either in the military or currently

enrolled in high school. For the Adult Education Interview, the bulk of the nonresponse was due to refusal of the sampled adult to respond (66.7 percent of nonresponse). Other reasons for Adult Education Interview nonresponse were inability to make contact with the sampled adult (19.5 percent of nonresponse), language problems with the sampled adult (4.2 percent of nonresponse), and other miscellaneous reasons such as the sampled adult being unable to respond due to illness (9.5 percent of nonresponse).

Table 4-3.—Number of enumerated children and adults, completed interviews, and weighted completion and response rates, by type of extended interview

Type of interview	Number	Estimated completion rate (percent)	Estimated response rate (percent)*
Parent Interview			
Enumerated	38,993		
Sampled	28,011		
Ineligible	151		
Did not respond	3,260		
Complete	24,600	88.3	65.4
Youth Interview			
Enumerated	15,563		
Sampled	10,651		
Ineligible	170		
Parent Interview not completed	1,376		
Parent refused for youth	460		
Youth did not respond	732		
Complete	7,913	76.0	56.3
Adult Education Interview			
Enumerated	20,266		
Sampled	8,114		
Ineligible	96		
Did not respond	1,321		
Complete	6,697	81.6	60.4

*The estimated response rate is computed by multiplying the Screener response rate of 74.1% by the appropriate completion rate.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1999.

The completion rates for the Parent, Youth, and the Adult Education Interviews can be examined by variables available for both respondents and nonrespondents. The variables shown for the Parent Interview are census region (based on the telephone number) and grade of the sampled child. Grade was collected during the Screener. Table 4-4 shows the number of sampled children by response status and completion rate for each of these variables. The completion rates are quite consistent across all Census regions and grades.

For the Youth Interview, three variables about each sampled youth are used for examining the response profile: Census region, grade of the youth, and type of school (i.e., public *vs.* private *vs.* home school). Census region was obtained based on the telephone number, grade was obtained from the Screener, and type of school was obtained from the Parent Interview. The distribution of cases for these variables and the completion rates are shown in table 4-5. There is little variation in the completion rates for region or for students whose grade is known. The completion rates by type of school are more variable, with the lowest rate for home schoolers.

For the Adult Education Interview, four variables were considered in examining the response profile: Census region (based on the telephone number), sex (from the Screener), adult education participation status as reported by the Screener respondent, and an indicator of whether the sampled adult was the Screener respondent (see table 4.6). There is little variation in completion rates across regions. The completion rate for females is higher than that for males, and the completion rate for adults reported by the Screener respondent to be adult education participants is higher than the completion rate for those reported to be non-participants. Sampled adults who were the Screener respondents completed the Adult Education interview at higher rates than those who were not the Screener respondents.

Table 4-4.—Number of sampled Parent Interviews, by response status and weighted completion rates

Parent Interviews	Total	Responded	Did not respond	Ineligible	Estimated completion rate (percent)
Total	28,011	24,600	3,260	151	88.3
Census region					
Northeast	4,912	4,270	625	17	87.3
South	10,761	9,473	1,224	64	88.4
Midwest	5,511	4,899	578	34	89.8
West	6,827	5,958	833	36	87.4
Grade of child (Screener)					
Not enrolled	5,486	4,952	506	28	91.0
Nursery/Preschool	2,225	1,983	241	1	89.7
Kindergarten	1,642	1,480	157	5	91.1
1st grade	1,648	1,458	189	1	88.9
2nd grade	1,596	1,392	204	0	87.3
3rd grade	1,588	1,403	182	3	88.9
4th grade	1,565	1,377	185	3	89.0
5th grade	1,622	1,402	219	1	88.5
6th grade	1,500	1,306	191	3	87.1
7th grade	1,607	1,407	197	3	87.7
8th grade	1,566	1,373	193	0	87.1
9th grade	1,563	1,376	177	10	87.3
10th grade	1,472	1,258	200	14	85.6
11th grade	1,420	1,229	171	20	86.4
12th grade	1,484	1,191	242	51	81.4
Unknown	12	0	5	7	0.0
Other*	15	13	1	1	94.9

* "Other" includes special education and ungraded.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1999.

Table 4-5.—Number of sampled Youth Interviews, by response status and weighted completion rates

Youth Interviews	Total	Responded	Did not respond	Ineligible	Parent Interview not completed	Estimated percent completed among those with a completed Parent Interview[3]
Total	10,651	7,913	1,192	170	1,376	86.8
Census region						
Northeast	1,818	1,292	225	24	277	84.7
Midwest	2,110	1,622	233	35	220	87.7
South	4,192	3,137	455	76	524	87.4
West	2,531	1,862	279	35	355	86.8
Grade of child (Screener)						
6th grade	1,502	1,106	186	21	189	84.9
7th grade	1,599	1,221	171	10	197	87.4
8th grade	1,567	1,191	178	5	193	87.0
9th grade	1,557	1,192	173	15	177	86.8
10th grade	1,464	1,089	152	23	200	88.7
11th grade	1,411	1,039	176	25	171	85.6
12th grade	1,506	1,059	148	58	241	87.9
Other[1]	31	15	7	6	3	61.5
Unknown	14	1	1	7	5	77.6
Type of school (Parent Interview)						
Public	8,103	7,036	1,016	51	0	87.2
Private	924	779	142	3	0	84.2
Home schoolers	134	98	34	2	0	76.4
Unknown[2]	1,490	0	0	114	1,376	0.0

[1]"Other" includes special education, ungraded, and grades other than 6 through 12.

[2]Characteristics obtained during the Parent Interview are unknown for some ineligible youths and for youths for whom no Parent Interview was completed.

[3]This may be viewed as a conditional Youth Interview completion rate, where the denominator is the weighted number of youth with completed Parent Interviews rather than the weighted number of youth sampled for a Youth Interview. Because the rates reported here are conditional, they differ from the rates reported in table 4-3 and throughout this manual, which are not conditional on the completion of the Parent Interview.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1999.

Table 4-6.—Number of sampled Adult Education Interviews, by response status and weighted completion rates

Adult Education Interviews	Total	Responded	Did not respond	Ineligible	Estimated completion rate (percent)
Total	8,114	6,697	1,321	96	81.6
Census region					
Northeast	1,405	1,162	233	10	81.0
South	3,126	2,566	506	54	81.0
Midwest	1,710	1,451	253	6	83.7
West	1,873	1,518	329	26	80.6
Sex (Screener)					
Female	4,508	3,829	650	29	84.1
Male	3,606	2,868	671	67	78.3
Adult education participation status (Screener)					
Adult education participant	4,542	3,953	519	70	86.4
Adult education non-participant	3,572	2,744	802	26	77.8
Screener respondent					
Sampled adult	5,145	4,620	469	56	89.6
Person other than sampled adult	2,969	2,077	852	40	69.6

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1999.

4.3   Item Response Rates

For most of the data items collected in the NHES:1999, the item response rate was very high. The tables in this section show the item response rates for a representative group of items for each interview. The items included were selected to represent key items 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.

Table 4-7 shows the item response rates for a representative group of items from the Parent Interview. Youth Interview item response rates for selected items are represented in table 4-8; table 4-9 presents the selected Adult Interview item response rates. For the Parent, Youth, and the Adult Education questionnaires the median item response rates for imputed items were 98.96 percent, 98.41 percent, and 99.25 percent, respectively. Items that are rarely asked, e.g. the items pertaining to basic skills courses in the Adult Education Interview, may involve unusual situations where the circumstances are less amenable to being collected in the standard set of questionnaire items. For such rarely asked items, a small number of missing values could result in a low item response rate.

For more details, including a complete listing of all item response rates, see *The 1999 National Household Education Survey Methodology Report* (Nolin et al. forthcoming).

Table 4-7.—Item response rates for selected items in the Parent Interview

Item	Number attempted	Percent response
Current school status and developmental characteristics		
Child being schooled at home	17,999	99.99
Child attended kindergarten	2,857	99.72
Child can identify colors	3,561	99.80
Child's care arrangements		
Attends program before/after school or both	2,351	99.70
Child attends center based program	6,939	99.96
Location of Head Start program	478	99.37
Head Start located at workplace	463	98.70
Amount household pays for nonrelative care	1,707	90.80
Child receives care from a nonrelative	19,335	99.88
Relative care before/after school	2,495	99.76
Number of days per week child attends Head Start	467	99.57
School characteristics		
Child attends public/private school	17,400	99.80
Child attends church related school	2,110	99.57
Highest grade at child's school	17,400	98.28
Number of students in child's grade	524	82.63
Family involvement and school practices		
Satisfied with child's school	17,400	99.52
Family attended general school meeting	19,581	99.84
Family attended meeting with teacher	19,581	99.78
School tells family how child is doing in school	19,581	99.31
Plans for and cost of postsecondary education		
Child will attend school after high school	9,147	95.38
Saving money to pay for child's education	8,557	95.07
Cost of tuition at a specific 4-year college	1,755	71.00
Talked about financial aid with someone	8,557	95.28
Family involvement outside of school		
Attended parenting class	6,939	99.87
Took child on errands in past week	5,041	99.94
Visited library with child in past month	12,075	99.78
Number of times parent read to child in past week	12,672	99.83
Parents arrange after school activities	12,396	99.85
Parent and household items		
Mother worked for pay last week	23,650	99.67
Highest grade mother completed	23,650	99.22
Father worked for pay last week	18,314	99.69
Highest grade father completed	18,314	98.28
Family received WIC in past 12 months	24,600	99.63
Family received TANF/AFDC in past 12 months	24,600	99.54
Household income	24,600	89.39

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1999.

Table 4-8.—Item response rates for selected items in the Youth Interview

Item	Number attempted	Percent response
Family involvement in education and attitudes about school		
Family visited library with child in past month	7,913	99.90
Family has rules about TV viewing time	7,913	99.72
Family discusses decisions with child	7,913	99.89
Teachers maintain discipline	7,806	99.77
Child enjoys school	7,806	99.78
Activities that promote responsibility		
School has student government	7,806	92.49
Served/worked in student government	6,487	92.51
Works for pay	7,913	99.96
Service activities		
Does community service activities	7,913	99.81
School arranges service activities	4,080	99.31
First service activity helps elderly	1,022	93.74
Participating in first service activity now	4,141	99.86
First service activity schedule	4,141	99.61
Talk in class/group about service activities	4,080	99.56
Activities that promote civic involvement		
Watch/listen to national news with family in past week	6,812	99.66
Frequency for talking about national new with family	7,913	99.81
Can't understand politics/government	4,381	99.20
Family has no say in what government does	4,381	96.37
Frequency at which youth reads national news	7,913	99.95
Class increased interest in government issues	5,540	99.30
Took part in debate in class	7,913	99.68
Could write letter to government official	4,381	99.66
Plans for and cost of postsecondary education		
Will attend school after high school	7,913	98.33
Discuss colleges/schools with parents	7,619	98.29
Will start college education at 2 or 4 year school	7,132	95.26
Talks with teacher about college academic requirements	7,528	98.41
Likely to attend public/private 4 year college	3,382	93.70
Can estimate tuition at in-state 4 year college	4,283	94.19
Parent and household items		
Household income	7,913	91.01
Exact household income/nearest $1,000	1,284	64.95
Family received TANF/AFDC in past 12 months	7,913	99.82

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1999.

Table 4-9.—Item response rates for selected items in the Adult Education Interview

Item	Number attempted	Percent response
Civic involvement		
Teacher evaluations would improve education	6,697	97.30
Heard of Lifetime Learning Tax Credit	6,697	99.42
Used public library in past year	6,697	99.28
How often read newspaper	6,697	99.28
Participate in community service	6,697	99.34
Employment		
Looking for work in past 4 weeks	1,443	98.13
Earnings	5,183	84.20
Participation in adult education activities		
Apprenticeship program	6,697	99.91
Main reason for ABE/GED classes	107	97.20
Employer required credential program	1,319	99.62
Vocational/technical program	6,697	99.90
ESL classes	488	99.80
Career or job related courses	6,697	99.88
Employer required work-related courses	1,901	99.63
Work-related course through distance education—video conference	173	99.42
Household items		
Total household income	6,697	80.86
Total household income range	6,697	87.71

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1999.

This page is intentionally blank.

5. DATA PREPARATION

5.1   Coding and Editing Specifications

Most of the NHES:1999 interview data were coded by the interviewers during the interview using the 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.

5.1.1   Range Specifications

The ranges of most of the items were determined by the codes available for the responses, since most were closed ended. For open-ended items that required an entry by the interviewer (for example, ages and dates), there were no specific sets of responses; therefore, reasonable ranges were defined.

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 attended center-based care had a soft range of 1 to 50. 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 70). "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. These comments were reviewed by data preparation and project staff. 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, range edits were rerun against the entire database to ensure that no outliers were inadvertently introduced during the post-data-collection updating process or during imputation. Therefore, any outliers that exist in the data files were reviewed during the data preparation process and most often originated from information entered into the comments data file.

5.1.2   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:1999 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. Additional consistency (logic) checks for the NHES:1999 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 out of the soft range and allowed the interviewer to enter a correction. If the interviewer passed through the error screen once and information was still outside the soft range, but within the hard range, the interviewer was asked to reverify the information. After the second attempt, the inconsistent information was accepted. However, if an initial response was outside the hard range, the error message appeared and continued to reappear unless a response within the hard range was entered. If the respondent confirmed an answer outside of a hard range, the interviewer entered it as a comment. These verified responses were allowed in the data file. At several points during data collection, logic edits were also checked against the entire data base. Cases violating the edits were examined 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., a 15-year-old enrolled in preschool). In such circumstances, there was supporting information available in the comments data file.

5.1.3   Structural Edits

Because of the survey's complexity, the CATI database was a highly complex, hierarchical file. The relationships of database records were often dependent on values of variables contained in other database records; therefore, structural edit specifications were developed to check the structural integrity of the database. This ensured that all variables that should exist did exist and those that should not exist did not exist in the database. For example, if there was a completed Parent Interview for a child, the data record that contained the parent items must have existed in the database. Structural edits were run against the entire database during the data preparation.

5.1.4   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. Members of the data preparation team checked each item to make sure the correct number of responses was represented for all items. 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, either a call back was made to the household to obtain the information or the item was coded as "not ascertained," and later imputed.

5.1.5   Review of "Other, specify" Items

The "other, specify" open-ended text responses were reviewed to determine if they should be coded into one of the existing code categories. When a respondent selected an "other" response, the interviewer entered text 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 one exception, no one particular text item occurred frequently enough to warrant the creation of a new response category for any of the questionnaire items. The exception was the addition of new response categories for the item asking the main reason child will not attend school after high school (PSNOTREA in the Parent Interview file and YSNOTREA in the Youth Interview file); these appear in italics on the questionnaire Verbatim strings of all "other, specify" items do not appear on the public-use data files. However, verbatim strings do appear on the restricted data file. See sections 6.3 of Volume II through Volume V for a discussion of the restricted files.

5.1.6   Coding of Open-Ended Items

Some open-ended items in the Youth Interview and Adult Education Interview were coded using coding schemes developed specifically for those items. In the Youth Interview, a subsample of those youth who indicated they had participated in community service in the past year were asked to describe their service activities. Codes were assigned to verbatim strings given in response to item YG1 ("Earlier you said that you participated in (SERVICE ACTIVITY) as community service or volunteer work. Please describe in

one sentence the work that you did."). Each service activity was assigned a broad category code and a specific activity code. These codes are included in the public data file (BCODE1 through BCODE3, SPCODEA1-3, SPCODEB1-3, and SPCODEC1-3). The coding manual for the youth service activities is found in appendix H of Volume III of this manual. Verbatim strings used in coding youth service activities are included in the proprietary file of the NHES:1999 Youth Interview.

In the Adult Education Interview, the open-ended items that were coded were related to the industry and occupation of jobs reported by respondents and the major field of study for participants in credential programs. Codes for industry and occupation are included in the public data file (FSIC for industry; FSOC for occupation). The coding manual for industry and occupation, an aggregation of "Standard Industry and Standard Occupation Codes (SIC and SOC), is found in appendix H of Volume IV of this manual. Codes for major field of study are also included in the public data file (CIPF) and the Major Field of Study coding manual is found in appendix I of Volume IV of this manual. Verbatim strings used in coding industry and occupation and major field of study are included in the proprietary file of the NHES:1999 Adult Education Interview.

This page is intentionally blank.

REFERENCES

Brick, J.M., and Broene, P. (1997). Unit and Item Response Rates, Weighting, and Imputation Procedures *in the 1995 National Household Education Survey*. NCES Working Paper. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, 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). Washington, D.C.: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

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). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, 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). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, 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). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, 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). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Brick, J.M., Collins, M.A., Celebuski, C.A., Nolin, M.J., Squadere, T.A., Ha, P.C., Wernimont, J., West, J., Chandler, K., Hausken, E.G. (1992a). *National Household Education Survey: Adult and Course Data Files User's Manual*. (NCES Publication No. 92–019). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Brick, J.M., Collins, M.A., Celebuski, C.A., Nolin, M.J., Squadere, T.A., Ha, P.C., Wernimont, J., West, J., Chandler, K., Hausken, E.G. (1992b). *National Household Education Survey of 1991: Preprimary and Primary Data Files User's Manual*. (NCES Publication No. 92–057). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Brick, J.M., Collins, M.A., Nolin, M.J., Ha, P.C., Levinsohn, M., and Chandler, K. (1994a). *National Household Education Survey of 1993: School Readiness Data File User's Manual*. (NCES Publication No. 94–193). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Brick, J.M., Collins, M.A., Nolin, M.J., Ha, P.C., Levinsohn, M., and Chandler, K. (1994b). *National Household Education Survey of 1993: School Safety and Discipline Data File User's Manual*. (NCES Publication No. 94–218). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Brick, J.M., Judkins, D., Morganstein, D.R., and Rust, K. (1995a). *A User's Guide to WesVarPC*, Rockville, MD: Westat.

Brick, J.M., and Waksberg, J. (1991). "Avoiding Sequential Sampling With Random Digit Dialing," *Survey Methodology*, 17 (1), 27–42.

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

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

Collins, M.A., Brick, J.M., Kim, K., Gilmore, S., and Stowe, P. (1996a). *National Household Education Survey of 1995: Adult Education Data File User's Manual*. (NCES Publication No. 96–826). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Collins, M.A., Brick, J.M., Loomis, L.S., Gilmore, S., and Chandler, K. (1996b). *National Household Education Survey of 1995: Early Childhood Program Participation Data File User's Manual*. (NCES Publication No. 96–825). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Collins, M.A., Brick, J.M., Nolin, M.J., Vaden-Kiernan, N., Gilmore, S., and Chandler, K. (1997). *National Household Education Survey of 1996: Data File User's Manual, Volumes I through V*. (NCES Publication No. 97–425). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

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 Publication No. 97–39). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Nolin, M.J., Montaquila, J., Nicchitta, P., Kim, K., Kleiner, B., Lennon, J., and Chapman, C. (forthcoming). *The 1999 National Household Education Survey Methodology Report*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Rao, J.N.K., and Shao, J. (1992). "Jacknife 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.

Waksberg, J. (1978). "Sampling Methods for Random Digit Dialing," *Journal of the American Statistical Association*, 73(361), 40–46.

Westat. (1998). *WesVar Complex Samples 3.0 User's Guide*. (1998). Rockville, MD: Westat.

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

This page is intentionally blank.

APPENDIX B

**SUMMARY OF WEIGHTING AND
SAMPLE VARIANCE ESTIMATION VARIABLES**

Summary of weighting and sample variance estimation variables

NHES Data File	Full Sample Weight	Computing Sampling Errors					Approximating Sampling Errors
		Replication Method (WesVarPC[1])			Taylor Series Method (SUDAAN and STATA[2])		DEFT (Average Root Design Effect)
		Respondent ID	Replicate Weights	Jackknife Method	Sample Design	Nesting Variables	
NHES:1991 *Early Childhood Education* ■ Primary file ■ Preprimary file	 EWGT EWGT	PERSID	 EWREPL1 - EWREPL50 EWREPL1 – EWREPL50	 JK1 JK1	 WR WR	 VSTRAT PSU VSTRAT PSU	 1.3 1.3
NHES:1991 *Adult Education* ■ Adult file ■ Course file[3]	 AEWT AEWT	 PERSID CLASID	 AEREPL1-AEREPL50 AEREPL1-AEREPL50	 JK1 JK1	 WR WR	 VSTRAT PSU VSTRAT PSU	4.5 Full Sample 2.3 Participants 2.8 Nonparticipants 3.8 Black (non-Hispanic) 3.2 Hispanic 2.8 White (non-Hispanic) 2.4 Other races
NHES:1993 *School Readiness*	FWGT0	ENUMID	FWGT1 - FWGT60	JK2	WR	STRATUM PSU	1.3
NHES:1993 *School Safety & Discipline* ■ Parent interviews only ■ Parent & Emancipated Youth (EY) interviews ■ Youth interviews (including Emancipated Youth)	 FWGT0 FWGT0 (for parents) & PFWGT0 (for EY) FWGT0	 BASMID BASMID ENUMID	 FWGT1-FWGT60 FWGT1-FWGT60, PFWGT1-PFWGT60 FWGT1-FWGT60	 JK2 JK2 JK2	 WR WR WR	 STRATUM PSU STRATUM PSU STRATUM PSU	 1.4 1.4 1.5
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

Summary of weighting and sample variance estimation variables—Continued

NHES Data File	Full Sample Weight	Computing Sampling Errors					Approximating Sampling Errors
		Replication Method (WesVarPC[1])			Taylor Series Method (SUDAAN and STATA[2])		DEFT (Average Root Design Effect)
		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 1.2 White, non-Hispanic

[1]WesVar Complex Samples software is available from SPSS; information can be obtained at www.spss.com. WesVarPC software, version 2.1, is available at no charge at www.westat.com.

[2]Information on SUDAAN can be obtained at www.rti.org. Information on STATA can be obtained at www.stata.com.

[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 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 for more details.