NCES 2009-476

U. S. DEPARTMENT OF EDUCATION

# Technical Report and Data File User's Manual

## For the 2003 National Assessment of Adult Literacy

NATIONAL CENTER FOR EDUCATION RESEARCH
Institute of Education Sciences

# Technical Report and Data File User's Manual

## For the 2003 National Assessment of Adult Literacy

**U.S. Department of Education**
NCES 2009-476

**July 2009**

Stéphane Baldi
*Editor*
**American Institutes for Research**

Stéphane Baldi
Mark Kutner
Elizabeth Greenberg
Ying Jin
Justin Baer
Elizabeth Moore
Eric Dunleavy
**American Institutes for Research**

Martha Berlin
Leyla Mohadjer
Greg Binzer
Thomas Krenzke
Jacqueline Hogan
Michelle Amsbary
Barbara Forsyth
Lyn Clark
Terri Annis
**Westat**

Jared Bernstein
**Ordinate Corporation**

Sheida White
*Project Officer*
**National Center for Education Statistics**

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

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

**National Center for Education Statistics**
Stewart Kerachsky
*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. Unless specifically noted, all information contained herein is in the public domain.

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

National Center for Education Statistics
Institute of Education Sciences
U.S. Department of Education
1990 K Street NW
Washington, DC 20006-5651

July 2009

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

**Suggested Citation**
Baldi, S. (Ed.) et al. (2009). *Technical Report and Data File User's Manual for the 2003 National Assessment of Adult Literacy* (NCES 2009-476). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

**For ordering information on this report, write to:**
U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794-1398

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

**Content Contact**
Sheida White
(202) 502-7473
sheida.white@ed.gov

## CONTENTS

## CONTENTS—Continued

## CONTENTS—Continued

## CONTENTS—Continued

## CONTENTS—Continued

## CONTENTS—Continued

## CONTENTS—Continued

## CONTENTS—Continued

## CONTENTS—Continued

## CONTENTS—Continued

## CONTENTS—Continued

# LIST OF TABLES

**Table** **Page**

## LIST OF TABLES—Continued

## LIST OF TABLES—Continued

**Table** **Page**

## LIST OF TABLES—Continued

**Table** **Page**

## LIST OF TABLES—Continued

**Table** **Page**

## LIST OF FIGURES

**Figure** **Page**



## LIST OF EXHIBITS

**Exhibit** **Page**

*This page is intentionally left blank.*

# CHAPTER 1

# THE 2003 NATIONAL ASSESSMENT OF ADULT LITERACY: AN OVERVIEW

Mark Kutner and Elizabeth Moore, American Institutes for Research

## 1.1 INTRODUCTION

The 2003 National Assessment of Adult Literacy (NAAL) is the fourth national assessment of adult literacy supported by the federal government, and it is the third such assessment supported by the National Center for Education Statistics (NCES), which is part of the U.S. Department of Education's Institute of Education Sciences. The previous assessments were a 1985 household survey of the literacy skills of 21- to 25-year-old adults, a 1989–90 U.S. Department of Labor-funded survey of the literacy proficiencies of job seekers, and the 1992 National Adult Literacy Survey (NALS) of adults 16 years of age and older. The 2003 assessment, also supported by NCES, was designed to assess changes in adult literacy since 1992.

For the 2003 assessment, approximately 18,000 adults ages 16 and older were randomly selected to represent the adult household population in the United States. The sample included approximately 1,000 adults in each of six states that chose to participate in a concurrent State Assessment of Adult Literacy (SAAL). The state assessments were designed to produce state-level results comparable to the national data. Six states opted to participate: Kentucky, Maryland, Massachusetts, Missouri, Oklahoma, and New York. As in 1992, the 2003 assessment also included a survey of prison inmates. The prison sample consisted of approximately 1,200 prison-incarcerated adults from 107 state and federal prisons. Their participation helped provide accurate estimates of the literacy of the total U.S. population, as well as separate estimates of the literacy of the prison population.

Respondents selected for participation in the 2003 assessment were asked to provide background demographic information and information about activities that adults undertake that are thought to be related to literacy. Respondents were then asked to complete a booklet of literacy tasks that were constructed to measure respondents' ability to read and use a wide array of printed and written materials.

A central objective of the 2003 assessment was to provide data that could measure changes in adult literacy between 1992 and 2003. The National Assessment of Adult Literacy also included three new features that were designed to enhance the information produced, while preserving trend:

- The Fluency Addition to NAAL (FAN) measured the oral fluency and basic reading skills of adults and produced a basic reading skill score.

- The Adult Literacy Supplemental Assessment (ALSA) was a performance-based assessment that used actual products and materials, rather than representations, to assess the basic literacy skill level of the lowest performing adults.

- The health literacy component measured the ability of adults to navigate and understand health materials.

## 1.2 DEFINING LITERACY

The 2003 adult literacy assessment covered the same content as the 1992 assessment, and both assessments used the same definition of literacy:

*Using printed and written information to function in society, to achieve one's goals, and to develop one's knowledge and potential.*

This definition implies that literacy goes beyond simply being able to sound out or recognize words and understand text. A central feature of the definition is that literacy is related to achieving an objective and that adults often read for a purpose.

### 1.2.1 Prose, Document, and Quantitative Literacy

As in 1992, three literacy scales—prose literacy, document literacy, and quantitative literacy—were used in the 2003 assessment:

- Prose literacy. The knowledge and skills needed to perform prose tasks (i.e., to search, comprehend, and use information from continuous texts). Prose examples include editorials, news stories, brochures, and instructional materials. Prose texts can be further broken down as expository, narrative, procedural, or persuasive.

- Document literacy. The knowledge and skills needed to perform document tasks (i.e., to search, comprehend, and use information from noncontinuous texts in various formats). Document examples include job applications, payroll forms, transportation schedules, maps, tables, and drug and food labels.

- Quantitative literacy. The knowledge and skills required to perform quantitative tasks (i.e., to identify and perform computations, either alone or sequentially, using numbers embedded in printed materials). Examples include balancing a checkbook, figuring out a tip, completing an order form, and determining the amount of interest on a loan from an advertisement.

The literacy tasks in the assessment were drawn from actual texts and documents, which were either used in their original format or reproduced in the assessment booklets. Each question appeared before the materials needed to answer it, thus encouraging respondents to read with purpose. Respondents could correctly answer many assessment questions by skimming the text or document for the information necessary to perform a given literacy task. All tasks were open-ended.

### 1.2.2 Establishing Literacy Levels

In response to a request from the National Center for Education Statistics (NCES), the National Research Council (NRC) convened a Committee on Performance Levels for Adult Literacy to set standards for the prose, document, and quantitative scales. The committee's goal was to do the following in an open and public way: evaluate the literacy levels used by NAAL's 1992 predecessor survey and recommend a set of new performance levels that could be used in reporting the 2003 results and also be applied to the 1992 results to make comparisons across years.

After reviewing information about the 1992 and 2003 assessments as well as feedback from stakeholders (e.g., practitioners), the committee specified a new set of performance levels intended to correspond to four policy-relevant categories of adults, including adults in need of basic adult literacy services. These four levels were *Below Basic*, *Basic*, *Intermediate*, and *Proficient*. The next step was to determine the score ranges to be included in each level for each of the three NAAL literacy scales: prose, document, and quantitative literacy.

To determine the score ranges for each level, the committee decided to use the Bookmark method. The initial implementation of the method involved describing the literacy skills of adults in the four policy-relevant levels and holding two sessions with separate panels of judges consisting of adult literacy practitioners, officials with state offices of adult education, and others. One group of judges focused on the 1992 assessment tasks, and the other group focused on the 2003 assessment tasks.

For each literacy area (prose, document, and quantitative), the judges were given, in addition to descriptions of the performance levels, a booklet of assessment tasks arranged from easiest to hardest. The judges' job was to place "bookmarks" in the set of tasks that adults at each level were "likely" to get right. The term *likely* was defined as "67 percent of the time," or two out of three times, and statistical procedures were used to determine the score associated with a 67 percent probability of performing the task correctly. The bookmarks designated by the judges at the two sessions were combined to produce a single bookmark-based cut score for each performance level on each of the three literacy scales.

To refine the bookmark-based cut scores, which indicated the lowest score to be included in each performance level, the committee used a procedure it called the "quasi-contrasting groups approach." Committee members compared the bookmark-based cut scores with the 1992 scores associated with various background variables, such as educational attainment. The criterion for selecting the background variables was potentially useful for distinguishing between adjacent performance levels such as *Basic* and *Below Basic* (e.g., having some high school education vs. none at all; reporting that one reads well vs. not well; reading a newspaper sometimes vs. never reading a newspaper; reading at work sometimes or more often vs. never reading at work).

In each case, the midpoint between the average scores of the two adjacent performance levels (*Below Basic* and *Basic*; *Basic* and *Intermediate*; *Intermediate* and *Proficient*) was calculated and averaged across the variables that provided contrasts between the groups. The committee developed a set of rules and procedures for deciding when and how to make adjustments to the bookmark cut scores when the cut scores associated with the selected background variables were different from the bookmark-based scores.

Furthermore, the NRC committee recommended that NCES distinguish a fifth group of adults with special importance to literacy policy—those who are nonliterate in English. As originally defined by the committee, the category "Nonliterate in English" consisted of adults who performed poorly on a set of easy screening tasks in 2003 and therefore were routed to an alternative assessment for the least-literate adults. Because the 1992 assessment included neither the alternative assessment nor the 2003 screening tasks, adults in this category cannot be identified for 1992.

To provide a more complete representation of the adult population who are nonliterate in English, NCES expanded the category to include not only the 3 percent of adults who took the alternative assessment, but also the 2 percent who were unable to be tested at all because they knew neither English nor Spanish (the other language spoken by interviewers). Thus, as defined by NCES, the category included about 5 percent of adults in 2003.

The new performance levels were presented to NCES as recommendations. Having accepted the general recommendations, NCES incorporated a few refinements before using the levels to report results. First, NCES changed the label of the top category from *Advanced* to *Proficient* because the term *proficient* better conveys how well the upper category of adults performs. Second, NCES added sample tasks from the 2003 assessment to illustrate the full range of tasks that adults at each level can perform, as well as a brief (one sentence) summary description for each level to enhance public understanding. Third, as outlined in the previous paragraph, NCES included additional adults in the *Nonliterate in English* category.

## 1.3 COMPONENTS OF THE NATIONAL ASSESSMENT OF ADULT LITERACY

The 2003 National Assessment of Adult Literacy comprised the background questionnaire, the main household assessment, the prison survey, the Fluency Addition to the NAAL (FAN), the Adult Literacy Supplemental Assessment (ALSA), and the health literacy component.

### 1.3.1 Background Questionnaire

The 2003 National Assessment of Adult Literacy household background questionnaire was used to collect data about various demographic and background characteristics. A primary goal of the assessment was to measure literacy trends between 1992 and 2003, so many of the questions on the 2003 background questionnaire were identical to questions on the 1992 background questionnaire. The 2003 background questionnaire also included some new questions that were added in response to input from stakeholders and users of the 1992 data.

A separate background questionnaire was developed for the prison survey. The prison background questionnaire was used to collect demographic data on inmates and provided contextual data on their experiences in prison that were related to literacy, including participation in classes, job training, and prison work assignments.

Both the household and the prison background questionnaire were administered orally in either English or Spanish. The demographic questions were identical on the prison and household background questionnaires.

### 1.3.2 Main Household Assessment

The main NAAL assessment, as distinct from the other NAAL components described in this section, measures how well Americans perform tasks with printed materials similar to those they encounter in their daily lives at work, at home, and in the community. Such tasks might include, for example, balancing a checkbook (quantitative literacy), filling out a job application (document literacy), or finding information in a news article (prose literacy). NAAL provides separate prose, document, and quantitative literacy scores.

### 1.3.3 Prison Survey

The NAAL prison survey is a nationally representative assessment of the English literacy skills of adult inmates in state and federal prisons in the United States. The assessment compares results for the prison population with those of the general U.S. adult population and reports changes in performance since the 1992 prison component of the National Adult Literacy Survey (NALS). Prison inmates were asked to complete the same tasks as adults living in households.

### 1.3.4 Fluency Addition to the NAAL (FAN)

In November 2001, a panel of experts recommended that the government provide, for the first time, a clearer picture of the basic reading skills of low-performing adults by examining their oral reading fluency. In response to this recommendation, an oral reading component for the NAAL, the Fluency Addition to the NAAL (FAN), was designed. FAN assessed the ability of adults to decode, recognize words and numbers, and read with fluency.

The tasks included on the oral reading fluency assessment were designed to be most sensitive to differences among readers with low proficiency instead of discriminating among highly proficient readers. Consistent with this approach, the word lists comprised frequent, common English words, and the reading passages were written at the elementary and middle school levels. Most proficient readers would not find the tasks particularly challenging, though they might differ in how efficiently they could complete them. In contrast, low proficiency readers might find the English words and passages (as well as the decoding tasks) challenging.

Four components were included in the oral reading fluency assessment (digit and letter reading, word reading, decoding, and passage reading). The components were measured as follows:

- Digit and letter reading
  - Respondents read a list of 35 letters and a list of 35 single-digit numbers.
- Word reading
  - Three word lists of varying difficulty were included on the assessment.
- Decoding
  - Decoding was measured through three lists of nonsense words.

- Passage reading
    - Eight passages were included on the oral reading fluency assessment.

### 1.3.5 Adult Literacy Supplemental Assessment (ALSA)

One of the limitations most often cited about the 1992 NALS was a lack of information about the literacy abilities of adults performing at the lowest levels on the assessment. The 2003 NAAL sought to address this problem by including a supplemental assessment given only to those adults who could not successfully complete the easiest prose, document, and quantitative items that appeared at the beginning of the assessment.

For these reasons, the ALSA is an interactive and adaptive assessment that uses authentic, highly contextualized materials commonly found in environments such as the home, the workplace, or a community agency. Although everything respondents were asked to read was written in English, the questions could be asked by the interviewers in either English or Spanish, and respondents were permitted to answer orally in either English or Spanish.

Some of the items presented to respondents in the ALSA include the following:

- Carbonated beverage can
- Television program schedule
- Utility bill
- Grocery advertisement

The items used in ALSA were not representations but were the actual items that respondents would encounter in everyday life. They also increased in difficulty as the administration progressed, an approach consistent with the cognitive demands in the main assessment. They included symbols that are found throughout the world and are recognizable to virtually anyone from any culture so that they are familiar to non-native English speakers and adults with only the most basic literacy skills. Instead of simply labeling a significant portion of the population as unable to read, the ALSA provides data on what skills low-literate adults do have that will allow policymakers and practitioners to adapt their curriculums, instructional materials, and professional development activities.

### 1.3.6 Health Literacy Component

The 2003 adult literacy assessment included a health literacy scale that consisted of 12 prose, 12 document, and 4 quantitative NAAL items. The health literacy items reflected the definition of health literacy used by the Institute of Medicine and by Healthy People 2010 (a set of national disease prevention and health promotion objectives led by the U.S. Department of Health and Human Services):

***The degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions.*** (U.S. Department of Health and Human Services 2000; Institute of Medicine 2004)

Tasks used to measure health literacy were organized around three domains of health and health care information and services: *clinical, prevention,* and *navigation of the health care system.* The stimulus materials and the 28 health literacy tasks were designed to assess respondents' skills for locating and understanding health-related information and services, and to represent the three general literacy scales—prose, document, and quantitative—developed to report the NAAL results.

The materials were selected to be representative of real-world health-related information, including insurance information, medicine directions, and preventive care information. Of the 28 health literacy tasks, 3 represented the *clinical* domain, 14 represented the *prevention* domain, and 11 represented the *navigation of the health care system* domain. The domains are defined as follows:

- The clinical domain encompasses those activities associated with the health care provider-patient interaction, clinical encounters, diagnosis and treatment of illness, and medication.
- The prevention domain encompasses those activities associated with maintaining and improving health, preventing disease, intervening early in emerging health problems, and engaging in self-care and self-management of illness.
- The navigation of the health care system domain encompasses those activities related to understanding how the health care system works and individual rights and responsibilities.

The NAAL health literacy scale did not include tasks that did not fit the definitions of prose, document, or quantitative literacy even if they were consistent with the definition of health literacy used by Healthy People 2010. For example, none of the NAAL health tasks required knowledge of specialized health terminology. The assessment also did not measure the ability to obtain information from nonprint sources, although questions about the use of all sources of health information—both written and oral—were included on the background questionnaire and are included in the report.

## 1.4 CONDUCTING THE SURVEY

### 1.4.1 Field Test

From April through August 2001, staff from the National Center for Education Statistics (NCES) and its contractors worked collaboratively to prepare for the National Assessment of Adult Literacy (NAAL) field test. During this period, a fully automated field test system was developed, and data delivery systems and procedures were implemented. Publicity materials for improving study cooperation rates were designed, and instructional manuals and training programs for supervisors and interviewers were developed.

Following the conclusion of the field test, the field-tested cognitive items were scored and the results were analyzed to determine which items to retain for the operational assessment. The background questionnaire (BQ) data obtained during the field test were analyzed, and changes were made to the BQ on the basis of the field-test data. The field-test results were also used to select the core items for the operational assessment and to develop the algorithm for selecting Adult Literacy Supplemental Assessment (ALSA) respondents.

### 1.4.2 Data Collection

Household data collection was conducted from March 2003 through February 2004; prison data collection was conducted from March through July 2004. Although data collection extended into 2004, the study is referred to as the 2003 National Assessment of Adult Literacy throughout this report and other reports, which follows the convention of the 1992 National Adult Literacy Survey, for which data collection extended into 1993.

Household interviews were conducted in respondents' homes; prison interviews usually took place in a classroom or library in the prison. Whenever possible, interviewers administered the background questionnaire and assessment in a private setting. Assessments were administered one-on-one using a computer-assisted personal interviewing system (CAPI) programmed into laptop computers. Respondents were encouraged to use whatever aids they normally used when reading and when performing quantitative tasks, including eyeglasses, magnifying glasses, rulers, and calculators.

The background questionnaire was administered orally, with interviewers reading questions from the computer screen and entering responses directly into the computer. Skip patterns and follow-up probes for contradictory or out-of-range responses were programmed into the computer. After completing the background questionnaire, respondents were handed a booklet with the assessment questions. The interviewers followed a script that introduced the assessment booklet and guided the respondent through the assessment.

Each assessment booklet began with the same seven questions (known as the assessment's core items). These seven questions required the respondents to read materials written in English, but the questions were presented in either English or Spanish. After the respondent completed those seven questions, the interviewer asked the respondent for the book and used an algorithm to determine on the basis of the responses to the questions whether the respondent should continue in the main assessment or be placed in the Adult Literacy Supplemental Assessment (ALSA). Three percent of adults weighted (5 percent unweighted) were placed in the ALSA.

A respondent who continued in the main assessment was given back the assessment booklet, and the interviewer asked the respondent to complete the tasks in the booklet and guided the respondent through the tasks. The main assessment consisted of 12 blocks of tasks with approximately 11 questions in each block, but each assessment booklet included only 3 blocks of questions. The blocks were spiraled so that across the 26 different configurations of the assessment booklet, each block was paired with every other block and each block appeared in each of the three positions (first, middle, last) in a booklet.

For ALSA interviews, the interviewer read the ALSA script from a printed booklet and classified the respondent's answers into the response categories in the printed booklet. ALSA respondents were handed the materials they were asked to read. Following the main assessment or ALSA, all respondents were administered the oral fluency assessment (FAN). Respondents were handed a booklet with passages, number lists, letter lists, word lists, and pseudoword lists to read orally. Respondents read into a microphone that recorded their responses on the laptop computer.

## 1.5 SAMPLE DESIGN

The 2003 National Assessment of Adult Literacy included two samples: (1) adults ages 16 and older living in households (99 percent of the sample weighted) and (2) inmates ages 16 and older in federal and state prisons (1 percent of the sample weighted). Each sample was weighted to represent its share of the total population of the United States, and the samples were combined for reporting.

### 1.5.1 Household Sample

The 2003 National Assessment of Adult Literacy household sample included a nationally representative probability sample of households. The household sample was selected on the basis of a four-stage, stratified area sample: (1) primary sampling units (PSUs) consisting of counties or groups of contiguous counties; (2) secondary sampling units (referred to as segments) consisting of area blocks; (3) housing units containing households; and (4) eligible persons within households. Person-level data were collected through a screener, a background questionnaire, the literacy assessment, and the oral module. To increase the number of Black and Hispanic adults in the NAAL sample, segments with moderate to high concentrations of Black and Hispanic adults were given a higher selection probability. Segments in which Blacks or Hispanics accounted for 25 percent or more of the population were oversampled at a rate up to three times that of the remainder of the segments. The final household reporting sample consisted of 18,102 respondents. The final weighted response rate for the household sample was 62.1 percent.

### 1.5.2 Prison Sample

The 2003 assessment also included a nationally representative probability sample of inmates in federal and state prisons. This sample was selected in two stages: (1) the selection of primary sampling units (PSUs) made up of federal and state prisons and (2) the selection of inmates within each PSU. The final prison reporting sample consisted of 1,156 respondents. The final weighted response rate for the prison sample was 88.3 percent.

## 1.6 REDUCING THE RISK OF DATA DISCLOSURE

Over the past decade, concerns about the disclosure of information related to individual survey respondents have increased. New laws have been put in place since the Privacy Act of 1974 to further ensure the protection of confidential data. The National Center for Education Statistics (NCES) and data contractors pledge confidentiality to respondents. The recently passed Education Sciences Reform Act of 2002 explicitly requires that NCES protect the confidentiality of all those responding to NCES-sponsored surveys so that no individual or facility can be identified. More specifically, NCES Standard 4-2, *Maintaining Confidentiality* (NCES 2002), provides guidelines for limiting the risk of data disclosure for data released by NCES. NAAL staff took careful measures to comply with these standards.

## 1.7 RESPONSE RATES AND NONRESPONSE BIAS ANALYSIS

NCES statistical standards require a nonresponse bias analysis when the unit response rate for a sample is less than 85 percent. The nonresponse bias analysis of the household sample revealed differences in the background characteristics of respondents who participated in the assessment compared with those who refused. A series of nonresponse bias analyses revealed that the potential amount of nonresponse bias attributable to unit nonresponse at the screener and background questionnaire stages was likely to be negligible.

## 1.8 WEIGHTING AND VARIANCE ESTIMATION

A complex sample design was used to select assessment respondents. The properties of a sample selected through a complex design might be very different from those of a simple random sample in which every individual in the target population has an equal chance of selection and in which the observations from different sampled individuals can be considered to be statistically independent of one another. Therefore, the properties of the sample for the complex data collection design must be taken into account during the analysis of the data. Standard errors calculated as though the data had been collected from a random sample would generally underestimate sampling errors. The NAAL uses sampling weights to account for the fact that the probabilities of selection were not identical for all respondents. Because the assessment used clustered sampling, conventional formulas for estimating sampling variability that assume random sampling and hence independence of observations are inappropriate. For this reason, all analyses done using the NAAL assessment data use a Taylor series procedure or another procedure that can incorporate the weights and account for the complex sample design.

## 1.9 SCORING

Different procedures were employed for scoring the three main components of the 2003 assessment: the cognitive items, the Fluency Addition to NAAL (FAN), and the Adult Literacy Supplemental Assessment (ALSA). For the cognitive items and the ALSA, the scoring procedures used were similar to the procedures implemented for scoring the NAAL field test. Scoring the FAN was more complex because the scores were generated by an automatic speech recognition (ASR) system. To ensure the validity of the FAN data, a sample of tasks scored by the ASR were compared to a sample of tasks scored by human scorers.

## 1.10 ITEM ANALYSIS, SCALING, AND ESTIMATES OF SUBPOPULATION PROFICIENCIES

Each respondent to the NAAL received a booklet that included 3 of the 13 assessments blocks. Because each respondent did not answer all the NAAL items, item response theory (IRT) methods were used to estimate average scores on the prose, document, and quantitative literacy scales; a simple average percent correct would not allow reporting results that are comparable for all respondents. IRT models the probability of answering a question correctly as a mathematical function of proficiency or skill. The main purpose of IRT analysis is to provide a common scale on which performance on some latent trait can be compared across groups, such as those defined by sex, race/ethnicity, or place of birth (Hambleton and Swaminathan 1985).

## 1.11 THE LITERACY OF ADULTS WITHOUT COGNITIVE DATA

Missing data are always expected in any large-scale assessment. Sampled individuals may not respond to an assessment for many reasons. A number of alternative methods are available to deal with missing data. The least desirable way is simply to ignore the missing data. This practice assumes that the data are missing at random and that the remaining observed samples are representative of the target population. However, if the pattern of missing data is correlated to the outcome of the study, this practice would yield both biased and inaccurate estimates of proficiency distributions for some subpopulations and consequently for the total population as well. For those reasons, NAAL analysts made sure to confront potential nonresponse biases that may develop from missing data.

## 1.12 VARIABLE CONSTRUCTION AND FILE DEVELOPMENT

NAAL staff conformed to National Center for Education Statistics (NCES) guidelines while documenting their procedures of variable construction and file development. The processes involved various steps, including the construction of the NAAL public use data for the household study and the prison survey as well as the NAAL item parameter files, followed by construction of the derived variables. In addition, NAAL staff documented how to analyze NAAL data by using AM software, and how to use the electronic codebooks.

*This page is intentionally left blank.*

# CHAPTER 2

# DEVELOPMENT OF THE SURVEY INSTRUMENTS

Elizabeth Greenberg, American Institutes for Research[1]

One of the goals of the 2003 National Assessment of Adult Literacy (NAAL) was to relate the literacy skills of the nation's adults to a variety of demographic characteristics and to other variables measuring how adults use their literacy skills in workplace, family, and community settings. To accomplish this goal, the assessment included a background questionnaire (administered in English or Spanish), as well as literacy tasks. This chapter summarizes the conceptual framework for the literacy assessment and discusses the development of the instruments administered in the assessment, including both the household and the prison background questionnaires, the 2003 cognitive items, and the two instruments that were newly developed for the 2003 assessment: the Fluency Addition to the NAAL (FAN) and the Adult Literacy Supplemental Assessment (ALSA). This chapter also describes bias and sensitivity review of the NAAL items, block assembly, and booklet design.

## 2.1 CONCEPTUAL FRAMEWORK OF THE LITERACY ASSESSMENT

The conceptual framework for the 2003 National Assessment of Adult Literacy was based on the framework developed for the 1992 National Adult Literacy Survey (NALS) and used the same definition of literacy:

> *Using printed and written information to function in society, to achieve one's goals, and to develop one's knowledge and potential.*

This definition characterizes literacy by focusing on what adults do with printed and written information. The definition goes beyond simply decoding and comprehending text and implies that the information-processing skills that adults use to think about content are part of the concept of literacy.

As in 1992, three literacy scales—prose literacy, document literacy, and quantitative literacy—were used in the 2003 assessment. The three scales represent distinct and important aspects of the ability to use printed and written information. These scales are discussed in more detail in section 2.4.1 of this chapter.

---

[1]The text describing the development of the FAN word lists and pseudoword lists (sections 2.5.2.1 and 2.5.2.2) was written by John Sabatini and Richard L. Venezky. The text describing the conceptual framework for the ALSA (the introduction to section 2.6) was written by Heide Spruck Wrigley.

To measure trend to the 1992 assessment, 6 of the 13 blocks of items used in 1992 were reused in 2003. Seven blocks of items were newly developed for the 2003 assessment. In addition, a core of seven easy items—administered at the beginning of the cognitive assessment—was developed for the 2003 assessment to replace the six-item core used in 1992.

Both the 1992 and the 2003 assessments measured English literacy, and all texts that respondents were asked to read were presented in English only. However, the 2003 assessment differed from the 1992 assessment by offering the option of administering the seven core items in Spanish for respondents whose English skills were not adequate for comprehending the instructions or the questions in English. The texts on which the core questions were based were presented in English only.

The 1992 and 2003 assessments also differed in the guidelines concerning calculator use for quantitative tasks. In 1992, calculator use was limited to one block of items, and adults participating in the assessment were required to use calculators for the quantitative items in that block. In 2003, adults were told that they could use a calculator for any of the quantitative items if they wished to do so (either their own calculator or one provided by the interviewer), but they were not required to use a calculator for any of the items. Allowing respondents to use calculators is consistent with the assessment's functional definition of literacy.

Because of these changes, some caution in interpreting changes in literacy from 1992 to 2003 is advised. However, the changes result in the ability to provide more accurate data about the English literacy of adults.

The 2003 assessment included three components that were not part of the 1992 assessment:

- a health literacy scale;
- the Fluency Addition to the NAAL (FAN); and
- the Adult Literacy Supplemental Assessment (ALSA).

Enough items with a health-related focus were developed to allow a health literacy scale in addition to the prose, document, and quantitative scales. These items provided a measure of how well respondents could read material that presented specific information about health-related topics, that is, the skills and strategies called "health literacy." The health literacy items were also classified as prose, document, or quantitative and were reported on those scales. The only difference between the health literacy items and the items on the three other NAAL scales was the context of the items. Although health was also one of the content areas of the 1992 assessment, that assessment did not include enough health items to create a separate scale. The 2003 assessment included 28 tasks based on 14 stimulus materials with health content—enough items to create a separate scale. The health literacy scale is described in more detail in section 2.4.3 of this chapter.

The Fluency Addition to the NAAL (FAN) was developed to assess the basic reading skills of adults as a complement to the functional literacy focus of the NAAL. The FAN consisted of a series of oral reading tasks. Respondents were asked to read aloud lists of digits, letters, words, pseudowords (nonsense words spelled phonetically), and passages. Their reading was recorded and then analyzed by computer for speed and accuracy. The FAN assessment was administered after the main NAAL to avoid interfering with measuring trend between 1992 and 2003. The FAN is discussed in more detail in section 2.5 of this chapter.

The Adult Literacy Supplemental Assessment (ALSA) was an alternative, performance-based assessment that allowed adults with marginal literacy skills to demonstrate what they could and could not do when asked to make sense of various forms of print. Respondents were screened into the ALSA on the basis of their responses to the seven core questions administered at the beginning of the cognitive assessment. Unlike the main assessment, for which respondents had to read the questions and instructions for the tasks, all ALSA tasks were administered orally in either English or Spanish, but the materials respondents were asked to read were provided in English only. All ALSA respondents were included on the main NAAL scale on the basis of their responses to the core questions; they also received separate ALSA scores. The ALSA is discussed in more detail in section 2.6 of this chapter.

## 2.2 HOUSEHOLD BACKGROUND QUESTIONNAIRE

The NAAL background questionnaire (BQ) collected data to give policymakers, program administrators, and researchers current information about the relationship between literacy and various demographic and background variables. The information collected on the BQ can be used to describe the literacy levels of demographic groups, identify target populations for literacy services, and describe the relationship between literacy level and social and economic outcomes. A primary goal of the NAAL was to maintain comparability of the prose, document, and quantitative scales between 1992 and 2003, so many of the questions on the NAAL BQ were identical to questions on the 1992 NALS BQ. The 2003 BQ also included some new questions that were added to collect data on policy and program concerns that the 1992 survey did not address. The BQ was available in Spanish and English, and bilingual interviewers were employed in areas with large Spanish-speaking populations. The BQ was administered with a computer-assisted personal interviewing (CAPI) system built into the laptop computers that the interviewers carried with them. The CAPI system allowed complex skip patterns to be automated so that sections of the questionnaire could be better targeted at specific populations (e.g., nonnative English speakers, older Americans, parents, people who had received welfare). The 2003 BQ was approximately 10 minutes longer than the 1992 BQ. The extra time enabled interviewers to collect more information of interest to stakeholders.

### 2.2.1 Development Procedures

Prior to awarding the contract for the development of the 2003 NAAL, the National Center for Education Statistics (NCES) sponsored several studies evaluating the content of the 1992 BQ. Smith and Sheehan-Holt (2000) surveyed secondary users of the 1992 NALS data and obtained their recommendations for modifying the information collected on the BQ. Reder and Edmonston (2000) analyzed demographic changes in the population over the decade from 1992 to 2002 and recommended changes to the BQ to address the changing demographics. Sherman, Condelli, and Koloski (1999) held focus groups with stakeholders and gathered their recommendations for the type of information that should be collected on the BQ. On the basis of the information collected in these studies, the 1992 BQ was modified to better serve the needs of NAAL data users. Items that were not useful to stakeholders in 1992 were dropped from the 2003 BQ and new items were added at the suggestion of stakeholders.

The following NAAL stakeholders reviewed the draft of the BQ for issues of content coverage, burden, and bias and sensitivity. If a reviewer was asked to comment on only certain sections of the BQ, those sections are indicated in parentheses. Following these reviews, the response options for specific questions were changed and some additional questions were added to the BQ.

- David W. Baker, M.D., M.P.H., Center for Healthcare Research and Policy, Case Western Reserve University (health)
- Dian Bates, Manager, Bureau of Adult Education, New Jersey State Department of Education
- Patricia Bennett, Program Manager, Maryland State Department of Education
- Jim Bowling, State Director of Adult Education, Ohio Department of Education
- James Conley, U.S. Department of Labor, Employment and Training Administration (job training and skills)
- Mary Craigle, Research Manager, Montana Office of Public Instruction
- Robert Crotzer, Adult Basic Education Coordinator, Maine Department of Education

- Mary Jo Deering, U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion (health)

- Debbie Faucette, Adult Education Section Leader, Louisiana Department of Education, Division of Adult Education and Training

- Jack Fyock, Health Care Financing Administration (health)

- Tom Grinde, Education Program Specialist, Wisconsin Technical College System Board

- Mark Haskins, Associate, New York State Department of Education, Office of Adult Literacy and Workforce Preparation

- Jeff Jagnow, Policy and Evaluation Branch Supervisor, Kentucky Department for Adult Education and Literacy

- Cheryl Keenan, State Director, Pennsylvania Department of Education

- Inaam Mansor, Arlington (Virginia) Public Schools (general and language background, education)

- Dan Miller, State Director of Adult Education, Illinois State Board of Education

- Ruth Parker, Emory University (health)

- Ron Pugsley, U.S. Department of Education (adult education, family literacy)

- Pavlos Roussos, Senior Director of Adult Education, Texas Education Agency

- Rima Rudd, Harvard University (health, political and social participation, literacy practices)

- Jon F. Warren, Director, Adult Education and Family Literacy, Missouri State Department of Education

- Jon Weintraub, U.S. Department of Education, Office of Vocational and Adult Education (OVAE), Office of Policy Analysis (labor force participation)

- Phil White, Director, Office of Adult Education, Tennessee Department of Education

- Tom White, Assistant State Superintendent, Oklahoma

- Mark V. Williams, M.D., Emory University School of Medicine (health)

- Heide Spruck Wrigley, Aguirre International (general and language background, education, political and social participation, literacy practices, job training and skills, family literacy, household income and welfare participation)

### 2.2.2 Cognitive Laboratory Analyses

All new NAAL background questions were evaluated in cognitive laboratories. Cognitive laboratories are structured, one-on-one interviews that use a think-aloud procedure originally developed by Ericsson and Simon (1980) to study problem solving by college students. In cognitive laboratory settings, respondents are taught how to think aloud, or express their thoughts orally, as they interpret or respond to a question. After the interviewer is satisfied that respondents understand what is expected, the interviewer proceeds through the questions, asking respondents to think-aloud as they respond to the survey items. The interviewer follows a script that includes probes that can be administered to elicit more information about the respondents' thought processes as they respond to a question. Some probes are administered to all respondents; other probes are administered at the discretion of the interviewer if respondents do not volunteer specific information during the think-aloud process.

With insights into the respondents' cognitive processes provided by the think-aloud interview and the follow-up probes, and with an understanding of the question's intent (as described by the question writer in a rationale statement), the analyst reviewing the cognitive laboratory results can usually determine whether the item is being interpreted and answered as the item writer intended. In addition, if the question is not functioning as intended, the think-aloud procedure and probes suggest the reasons the question may not work properly. This information is used to revise the question.

Twenty-one adults participated in the NAAL cognitive laboratories to evaluate the BQ. They were paid $50 and local travel expenses (public transit, taxi, or parking and mileage). Participants were recruited through a variety of channels, including ads in a local newspaper, flyers distributed at churches and community centers, and personal connections. Potential participants were screened to ensure that the sample of cognitive laboratory participants was diverse in terms of demographic characteristics, including gender, race/ethnicity, parenting status, native language, and welfare participation. These demographic characteristics were of interest because many of the new NAAL background questions focused on family literacy, language experiences of nonnative English speakers, and welfare participation. Each participant was asked to complete approximately half the BQ. Because of the skip patterns in the BQ, no question was answered by more than nine participants. Table 2-1 shows the demographic breakdown of the cognitive laboratory participants. The interviews were videotaped so that they could be reviewed later to identify item problems that might not have been apparent to the interviewer at the time of the interview.

**Table 2-1. Number of background questionnaire cognitive laboratory participants, by selected characteristics: 2003**

| Demographic characteristic | Number of participants |
|---|---|
| **Gender** | |
| Male | 9 |
| Female | 12 |
| **Race/ethnicity** | |
| White | 3 |
| Black | 12 |
| Hispanic | 5 |
| Multiracial | 1 |
| **Parental status**[1] | |
| Parents with children under 5 | 3 |
| Parents with children between 5 and 18 | 9 |
| Participants with no children or with adult children | 13 |
| **Language** | |
| Nonnative English speakers | 7 |
| Native English speakers | 14 |
| **Welfare status** | |
| On welfare during past year | 9 |
| On welfare, but not during past year | 1 |
| Never on welfare | 11 |

[1] Parental status categories are not mutually exclusive. Parents can have children under 5 and between 5 and 18.
NOTE: Black includes African American, and Hispanic includes Latino. Race categories exclude Hispanic origin unless specified.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 2.2.3 Key Constructs and Variables

The NAAL BQ covered the following areas:

- general and language background;
- educational background and experiences;
- political and social participation;
- labor force participation;
- literacy practices;
- job training and skills;
- demographic information;

- family literacy;
- household income and welfare participation;
- health; and
- additional demographics.

#### 2.2.3.1 General and Language Background

Section A of the BQ included questions on the following topics related to demographics and language background:

- age;
- country of birth;
- years living in the United States;
- age moved to the United States;
- education completed before moving to the United States;
- language(s) spoken before starting school;
- language(s) spoken by others in the home while growing up;
- language(s) currently spoken;
- self-evaluation of proficiency in English and other language(s); and
- participation in an English-as-a-second language (ESL) course.

Most of the questions in section A were trend questions originally used in 1992. The question about self-evaluation of proficiency in language(s) currently spoken was expanded to include all languages the respondent knew, not just English and language(s) learned before starting school. The question about ESL classes was modified from 1992, and the question about age when moving to the United States was new in 2003. Additional questions were added to address in more depth the topic of self-evaluation of proficiency in English.

#### 2.2.3.2 Educational Background and Experiences

Section B of the BQ included questions on the following topics related to educational background and experience:

- highest level of education completed;
- reasons for not completing high school or college;
- year graduated from high school and college (or, for nongraduates, year stopped education);
- type of high school diploma received;
- state in which high school diploma was obtained (or, for nongraduates, last year of high school was completed);
- state in which college diploma was obtained;
- number of years living in current state;
- enrollment in basic skills classes;
- receipt of information technology skill certification; and
- receipt of other skill certification.

The question on highest level of education completed was a trend question to 1992. The question on reasons for stopping schooling was modified from a 1992 question (additional categories were added), and the question was extended so that it was asked of individuals who did not complete college and those who did not complete high school. All other questions in this section were either completely new in 2003 or substantially revised from the 1992 wording.

### 2.2.3.3 Political and Social Participation

Section C of the BQ included questions on the following topics related to political and social participation:

- sources of information about public affairs (both English and non-English);
- volunteering;
- television viewing;
- library use;
- citizenship;
- voting; and
- veteran's status.

The question about sources of information in English about public affairs was a trend question from 1992, but response options were added to gather information about the use of the Internet and books and brochures. The question about television viewing was reworded from the question used in 1992 to include DVDs and videotapes, as well as broadcast and cable television. All other questions in this section were either completely new in 2003 or substantially revised from the 1992 wording.

#### 2.2.3.4 Labor Force Participation

Section D of the BQ included questions on the following topics related to labor force participation:

- employment status during the past week, the past year, and the past three years;
- hours worked in previous week;
- reason for not working;
- type of employer (government, private, self-employed, family business);
- income from employment (past week and past year);
- occupation; and
- industry.

Most of the questions in this section were trend questions repeated from 1992. The question about type of employer was new, and the questions concerning wages were modified to determine whether the respondent was reporting gross pay or take-home pay and also to determine whether the respondent was reporting total pay for the year or just part of the year.

#### 2.2.3.5 Literacy Practices

Section E of the BQ included questions on the following topics related to literacy practices:

- frequency of reading various types of materials in English and other languages;
- frequency of reading various types of materials at work;
- frequency of different types of computer use; and
- frequency of receiving assistance from family members or friends with various types of literacy-related activities.

The questions in this section were based on questions from 1992. However, all the questions except for the one asking about receiving assistance from family members or friends were modified or extended to include different categories.

#### 2.2.3.6 Job Training and Skills

Section F of the BQ included questions on the following topics related to job training and skills:

- participation in work-related job training;
- employer sponsorship of job training;
- content of job training; and
- self-assessment of literacy and computer skills.

All questions in this section were new in 2003.

#### 2.2.3.7 Demographic Information

Section G of the BQ asked about

- country of birth of respondent's parents; and
- educational attainment of respondent's parents.

The questions asking about the country of birth of the respondent's parents were new. The other questions in this section were trend questions to 1992.

#### 2.2.3.8 Family Literacy

Section H of the BQ included questions on the following topics related to family literacy:

- age of children living in the household;
- respondent's relationship to the children;
- interactive literacy activities between parents and children;

- training for parents regarding how to be the primary teacher for their children and full partners in their children's education; and
- computers in the home.

All questions in this section were new in 2003.

### 2.2.3.9 Household Income and Welfare Participation

Section I of the BQ included questions on the following topics related to household income and welfare participation:

- sources of household income;
- history of welfare participation (length of time, when participated, reasons for ending participation); and
- participation in classes to get off welfare.

The questions asking about sources of income were expanded from the version that appeared on the 1992 BQ. The other questions in this section were new in 2003.

### 2.2.3.10 Health

Section J of the BQ included questions on the following topics related to health:

- self-reported health status;
- self-reported disabilities (vision, hearing, learning, other);
- health insurance status for both the respondent and children living in the household;
- sources of information about health; and
- health screening/disease prevention activities.

The questions asking about vision and hearing difficulties were trend questions to 1992. All other questions in this section were new in 2003.

#### 2.2.3.11 Additional Demographics

Questions in this section asked about

- individual personal income from all sources;
- total family income from all sources; and
- race/ethnicity.

The response categories for personal and family income were changed from 1992. However, the wording of the questions remained the same. The questions asking about race and ethnicity were changed to reflect new Office of Management and Budget (OMB) requirements (U.S. Office of Management and Budget, 1997). In 2003, the question about Hispanic ethnicity was moved so that it was asked before the question about race; in 1992, the question about Hispanic ethnicity was asked after the question about race. Additionally, in 2003, respondents were given the option to choose as many categories as applied for both the question about Hispanic ethnicity and the question about race. In 1992, respondents could choose only one category in response to these questions.

### 2.2.4 Spanish Version

The BQ was translated into Spanish and the translation was reviewed by native Spanish speakers from Puerto Rican, Cuban, Argentinean, and Mexican backgrounds to ensure that the language used was comprehensible across a variety of Spanish cultures. The reviewers met and worked out agreements on language usage that could be understood by Spanish speakers from a variety of different backgrounds.

## 2.3 PRISON BACKGROUND QUESTIONNAIRE

A separate BQ was developed for the NAAL prison study. The prison BQ collected demographic data on inmates and provided contextual data on their experiences in prison that were related to literacy, including participation in classes, job training, and prison work assignments. The BQ was available in Spanish and English, and bilingual interviewers were employed in prisons with large Spanish-speaking populations. The prison BQ was administered with a CAPI built into the laptop computers the interviewers carried with them. The CAPI system allowed complex skip patterns to be automated so that sections of the questionnaire could be better targeted at specific populations (e.g., nonnative English speakers, inmates with low levels of formal education).

### 2.3.1 Development Procedures

The prison BQ was based on the 1992 prison BQ, with changes made to reflect changes in the 2003 household questionnaire. A few questions were also added to the prison BQ from the Survey of Inmates in State Correctional Facilities administered by the Bureau of Justice Statistics.

The following people reviewed the draft of the BQ for issues of content coverage, burden, and bias and sensitivity:

- Kay Britt, Roxbury Correctional Institution, Maryland (teacher)
- Vernell Doyle, Roxbury Correctional Institution, Maryland (teacher)
- Former inmate, Lorton Prison, Virginia
- Robert Johnson, Chair of Department of Justice, Law, and Society, American University
- John Linton, U.S. Department of Education and formerly with the Maryland State Correctional system
- Patricia O'Connor, Georgetown University
- Caroline Wolf Harlow, Bureau of Justice Statistics

### 2.3.2 Cognitive Laboratory Analyses

All questions on the prison BQ were evaluated in cognitive laboratories through one-on-one interviews as described in section 2.2.2. The interviews were conducted with inmates in the areas of the prisons used for educational classes. Because videotaping was forbidden in the prison environment, two NAAL staff members were present at each interview: one to conduct the interview and the other to take notes. To encourage prisoners to be open when responding to the questions about their experiences in prison related to literacy (including questions on topics such as ease of accessing the prison library), no guards or other prison officials were in the rooms at the time of the interview. Prisoners were not paid, but participating prisons were given a gift certificate to buy books for the prison library.

Nine interviews (with six men and three women) were conducted with inmates at three state facilities in Maryland and Virginia. NAAL staff were unable to do the same screening of prisoners to ensure demographic diversity that was done with household cognitive laboratory respondents. However, NAAL staff worked with the prison officials to ensure that the inmates participating in the cognitive laboratories included some inmates who had been in the prison long-term (over 5 years) and others who had been in the prison short-term (less than 2 years) because it seemed likely that experiences and the ability to answer the questions might differ by length of incarceration. NAAL staff also asked the prison officials to diversify the inmates participating in the cognitive laboratories by education levels so that some participants already had a GED or a high school diploma and others were in basic skills or GED classes.

### 2.3.3 Key Constructs and Variables Different From Household Questionnaire

The prison BQ included the following sections that were also on the household BQ:

- general and language background;
- educational background and experiences;
- political and social participation;
- literacy practices;
- demographic information;
- household income and welfare participation;
- health; and
- additional demographics.

The prison BQ also had some sections that were not on the household BQ:

- prison experiences (such as participation in classes and vocational training, and history of prior incarcerations); and
- prison work assignments and labor force participation (substituted for the labor force participation section on the household questionnaire).

The following household BQ sections were not on the prison BQ:

- labor force participation (this section was changed to prison work assignments and labor force participation);
- job training and skills (a few questions from this section were added to the prison work assignments and labor force participation); and
- family literacy.

The following sections describe more specifically how the prison BQ and the household BQ differed.

#### 2.3.3.1 General and Language Background

With one exception, all questions in this section were taken from the 2003 NAAL Household Background Questionnaire. One question concerning the location of an ESL class completed by the sampled prisoner was added.

#### 2.3.3.2 Educational Background and Experiences

With two exceptions, all questions in this section were taken from the 2003 NAAL Household Background Questionnaire. A question concerning the location of a basic skills class completed by the sampled prisoner was added, along with another question asking whether the inmate was on a waiting list for any academic classes. An additional response category was added to three questions drawn from the household questionnaire. Incarceration in a jail, prison, or detention center was added as reason for stopping schooling. The completion of a test preparation course while incarcerated was added as a response option to two questions about test preparation for technical and skills certification. The question on educational attainment was broken into two questions to obtain information about educational attainment prior to the current incarceration and additional education obtained in prison.

#### 2.3.3.3 Prison Experiences

This section drew on items from three questionnaires: the 1992 NALS Prison Background Questionnaire, the 2003 NAAL Household Background Questionnaire, and the Survey of Inmates in State Correctional Facilities administered by the Bureau of Justice Statistics. The questions in this section provided information about educational experiences unique to prison, prison social and community activities, and the inmate's criminal history. The two questions added from the Survey of Inmates in State Correctional Facilities provided information about participation in nonformal education while incarcerated (e.g., employment counseling, parenting skills) as well as in social and community activities.

#### 2.3.3.4 Prison Work Assignments and Labor Force Participation

The questions in this section were taken primarily from the 1992 NALS Prison Background Questionnaire. Two new questions about the frequency with which inmates read and wrote as part of their prison work assignment(s) were added. One response category for the question concerning preincarceration sources of income was added. The question was part of the1992 NALS Prison Background Questionnaire and was changed to include “Pay from jobs or wages” as a response option.

#### 2.3.3.5 Political and Social Participation

This section captured the political and social participation of prisoners, drawing on the 2003 NAAL Household Background Questionnaire. Three questions about inmates’ access to libraries were added. Two response categories for the question concerning prisoners’ sources of information for current events, public affairs, and the government were edited. The Internet was eliminated as a response option because prisoners do not have Internet access while incarcerated. Further, inmates and staff were included in the response category that named family members and friends as a source of information.

#### 2.3.3.6 Literacy Practices

This section provided data about a variety of literacy practices of inmates, drawing on the 2003 NAAL Household Background Questionnaire. Because of restrictions on the number and type of computer programs available to inmates, several response options were eliminated in the computer use question. Additionally, the wording for the question about getting help when completing basic literacy tasks was broadened to include other inmates and prison staff as a source of assistance.

#### 2.3.3.7 Demographic Information

The questions in this section were taken from the 2003 NAAL Household Background Questionnaire.

#### 2.3.3.8 Household Income and Welfare Participation

This section drew on the 2003 NAAL Household Background Questionnaire to collect data about inmates’ household income and welfare participation prior to incarceration. The response options on the marital status question were revised to eliminate responses inappropriate for prisoners. Incarceration was added to the question inquiring about why an inmate stopped receiving welfare payments.

#### 2.3.3.9 Health

The questions on this section were all taken from the 2003 Household Background Questionnaire. The Internet was eliminated as a response option for the question about sources of information about health because prisoners do not have Internet access while incarcerated.

#### 2.3.3.10 Additional Demographics

This section was taken from the 2003 Household Background Questionnaire.

## 2.4 2003 NATIONAL ASSESSMENT OF ADULT LITERACY COGNITIVE ITEMS

The 1992 National Adult Literacy Survey cognitive assessment included a core (six fairly easy items divided across the three scales at the beginning of the assessment that all respondents completed) and 13 additional blocks of items that were spiraled so that each respondent completed 3 blocks. Seven of the blocks in the 1992 assessment were new; 6 of the blocks were originally developed for the 1985 young adult literacy survey. The original goal for the 2003 assessment was to replace the core and the 6 blocks originally developed for the 1985 assessment and then reused in the 1992 assessment. However, because so many items performed well in the field test, the decision was made to also replace 1 of the blocks that was originally used in 1992. Therefore, the operational assessment included 7 new blocks plus 6 blocks from the 1992 assessment. The core items were also new.

The 1992 blocks that were replaced for the 2003 assessment were heavily weighted toward document items, and the decision was made to strive for more of a balance among the three scales in the newly developed blocks. Therefore, the replacement tasks could not simply mirror the tasks in the replaced blocks. Instead, the replaced tasks were analyzed to determine the range of items along the following dimensions:

- the structure of the stimulus materials (exposition, narrative, table, graph, map, etc.);
- the processes and strategies required to perform the task;
- the content represented or the context from which the stimulus was drawn (work, home, community, etc.); and
- the difficulty level of the task (percentage of respondents answering each item correctly).

Because of the change in the distribution of items across scales, the new items were not one-to-one matches with the old items along these dimensions. However, items were developed that reflected the range of the items being replaced along these dimensions.

In addition, the 2003 assessment included a new health literacy scale embedded in the prose, document, and quantitative items. Therefore, the 2003 tasks included more questions with health content than the tasks being replaced.

### 2.4.1 The Prose, Document, and Quantitative Scales

All items in the 2003 assessment were classified into the prose, document, or quantitative scales by using definitions similar to the definitions used in 1992. Specifically, the scales were defined as follows:

**Prose literacy.** The knowledge and skills needed to perform prose tasks (i.e., to search, comprehend, and use information from continuous texts). Prose examples included editorials, news stories, brochures, and instructional materials. Prose texts were further broken down as expository, narrative, procedural, or persuasive.

**Document literacy.** The knowledge and skills needed to perform document tasks (i.e., to search, comprehend, and use information from noncontinuous texts in various formats). Document examples included job applications, payroll forms, transportation schedules, maps, tables, and drug or food labels.

**Quantitative literacy.** The knowledge and skills required to perform quantitative tasks (i.e., to identify and perform computations, either alone or sequentially, using numbers embedded in printed materials). Examples included balancing a checkbook, figuring out a tip, completing an order form, or determining the amount of interest on a loan from an advertisement.

Some stimulus materials included both prose and document features. For example, a table (which is classified as a document) may have prose text around it describing the contents of the table. In these instances, tasks were classified according to where in the stimulus material the response to the task was located. If the response was in the table (the document part of the stimulus material), the task was classified as document. If the response was in the text around the table (the prose part of the stimulus material), the task was classified as prose.

### 2.4.2 Development of Items and Scoring Rubrics

The framework used to develop the prose, document, and quantitative items was the same framework used for the 1992 assessment (Campbell, Kirsch, and Kolstad 1992). This framework posited that literacy tasks vary along the following dimensions:

- materials/structure;
- processes/strategies;
- adult context/content; and
- task difficulty.

**Materials/structure.** Materials that adults read vary in the way the information in the materials is structured. At the highest level of aggregation, written materials can be classified as prose or document, or as a hybrid that includes characteristics of both prose and documents. Prose texts are organized in sentences and paragraphs, and the content may be narrative, expository, procedural, or persuasive. Document texts may be organized in matrix structures (i.e., rows and columns) or in an almost infinite variety of other formats, including maps, graphs, forms, indexes, bills, checks, coupons, and schedules. Documents are often designed to be skimmed, rather than read word for word. With the widespread availability of graphics and word processing software, hybrid texts, which combine features of prose and document texts, are becoming more common. Hybrid texts include graphs with prose explaining how to interpret the graph appearing underneath and informational articles in which the information is organized with subheadings and bullets.

The NAAL stimulus materials were selected to represent the structural variety of texts that adults encounter. To measure changes between 1992 and 2003, the 2003 NAAL pool included items structured similarly to items in the 1992 survey.

**Processes/strategies.** The processes or strategies required to perform the different types of tasks that adults perform with written materials vary across the materials and structures into which the materials can be classified. After stimulus materials were selected that represented the different types of materials and structure that adults regularly encounter, tasks were developed that reflected the different processes and strategies that adults use when they encounter written materials. The adult literacy framework posited that four basic processes/strategies characterize the prose and document literacy tasks: locate, cycle, integrate, and generate. For *locate* tasks, readers must match information given in the question with either literal or synonymous information in the text. *Cycle* tasks require readers to repeat the matching process multiple times. For 2003, *integrate* and *generate* tasks were combined into one category: higher-order thinking tasks. *Higher-order thinking* tasks require readers to do such things as pull together two or more pieces of information located at different points in a text or to go beyond the information in a text and make broad text-based inferences.

Quantitative tasks require different types of processes and strategies to complete. Although readers must obtain information from a written text to answer the quantitative questions (using locate or cycle strategies), completing the tasks requires performing arithmetical operations. The quantitative tasks were coded on the basis of whether one or more than one arithmetical operation was required to complete the task, as well as on the type(s) of arithmetical operation required (addition, subtraction, multiplication, division).

For a more detailed discussion of processes/strategies see Campbell, Kirsch, and Kolstad (1992).

**Context/content.** The substantive purposes for which adults read vary widely, and the NAAL cognitive tasks were developed to represent a wide variety of contexts in which adults might look for written information. Content areas represented in the pool of assessment tasks include

- community and citizenship: community resources and being informed;
- consumer economics: credit and banking, savings, advertising, making purchases, and maintaining personal possessions;
- health and safety: drugs and alcohol, disease prevention and treatment, safety and accident prevention, first aid, emergencies, staying healthy, and navigating the health system;
- home and family: interpersonal relationships, personal finance, housing, and insurance;
- leisure and recreation: travel, recreational activities, and restaurants; and
- work: occupations, finding employment, finance, and being on the job.

Because literacy practices vary so widely across the population, no single NAAL respondent was expected to regularly read materials in all the content areas covered on the assessment; some were expected to regularly read materials in only one or two of the content areas. Adults who dislike reading, or who read poorly, may avoid printed information as much as possible and read only the minimum amount necessary to complete their jobs and manage their finances. These infrequent readers may acquire other information through nonprint sources (radio, television, talking to friends or relatives, etc.). Other adults may read more broadly, using printed information to educate themselves about health issues, pursue their recreational interests, and so on.

Context/content proved to be more difficult to code than either materials/structure or processes/strategies. Many tasks overlapped two or more content areas. For example, a task asking a respondent to figure out the least expensive way to join a health club—using information presented in a table and accompanying text—could be classified as consumer economics (making purchases), health and safety (staying healthy), or leisure and recreation (recreational activities).

**Task difficulty.** The 1992 framework posited that the difficulty of a particular task was a result of the interaction of the type of process or strategy required by the task with other features of the task. NAAL staff were not able to consistently code all the variables influencing difficulty that were discussed in the 1992 framework. NCES has developed a framework for the adult literacy assessment that builds on the features of tasks identified in 1992 related to difficulty (White and McCloskey forthcoming). While developing the items, project staff focused primarily on the reading level of texts for prose items (measured with Lexile[2]) and the complexity of documents as reported by participants in cognitive laboratory interviews. As discussed above, quantitative tasks were coded for the number of mathematical operations required to complete the task (one or more than one) and for the type of mathematical operation required to complete the task (addition, subtraction, multiplication, division). Final determination of task difficulty when assembling the forms for the operational assessment was based on the field-test data.

Tables 2-2, 2-3, and 2-4 show the coding for the prose, document, and quantitative tasks included in the core and the seven new blocks of the 2003 adult literacy assessment.

#### 2.4.2.1 Development of Scoring Rubrics

The scoring rubrics were developed at the same time as the items by the item writers. The goal when developing the scoring rubrics was to determine whether respondents could accomplish the tasks posed in the items in real life. Thus, the level of detail and accuracy required in a response varied depending on the level of detail and accuracy that would be expected for a similar task in real life. Partial-credit points were included if there were substantively meaningful ways to accomplish part of task. Scoring rubrics were modified on the basis of the responses received in cognitive laboratories and were

[2] Lexile measures the complexity of a text based on semantic difficulty (vocabulary) and syntactic complexity (sentence length). See http://www.lexile.com for more information.

then reviewed by the expert panelists who reviewed the NAAL items. As described in chapter 4 of this report, the scoring rubrics were further refined on the basis of field test data.

**Table 2-2. Coding of the 2003 prose tasks, by type of text content, type of process/strategy, and Lexile score: 2003**

| Item number | Type of text content | Type of process/strategy | Lexile score |
|---|---|---|---|
| CC003 | Expository | Locate | 900 |
| CC004 | Expository | Locate | 900 |
| C020401 | Narrative | Locate | 790 |
| C020501 | Narrative | Higher-order thinking | 790 |
| C020901 | Expository | Locate | 1130 |
| C030101 | Expository | Cycle | 610 |
| C030201 | Expository | Locate | 610 |
| C030301 | Expository | Higher-order thinking | 610 |
| C040101 | Expository | Locate | 1030 |
| C040201 | Expository | Locate | 1030 |
| C040301 | Expository | Cycle | 1030 |
| C040701 | Expository | Higher-order thinking | 1240 |
| C050401 | Expository | Cycle | Hybrid/could not compute |
| C050801 | Expository | Locate | 1220 |
| C050901 | Expository | Locate | 1220 |
| C051001 | Expository | Higher-order thinking | 1220 |
| C051101 | Expository | Higher-order thinking | 1220 |
| C060101 | Narrative | Locate | 1030 |
| C060201 | Narrative | Higher-order thinking | 1030 |
| C061001 | Narrative | Higher-order thinking | 1130 |
| C061101 | Narrative | Higher-order thinking | 1130 |
| C070101 | Procedural | Locate | 460 |
| C070201 | Expository | Locate | 1200 |
| C070401 | Expository | Locate | 1200 |
| C070701 | Procedural | Locate | 700 |
| C070901 | Procedural | Locate | 620 |
| C071101 | Expository | Cycle | Hybrid/could not compute |
| C080301 | Narrative | Locate | 870 |
| C080401 | Narrative | Locate | 870 |
| C080601 | Persuasive | Locate | 1280 |
| C080701 | Persuasive | Higher-order thinking | 1280 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 2-3. Coding of the 2003 document tasks, by type of text content, and type of process/strategy: 2003**

| Item number | Type of text content | Type of process/strategy |
|---|---|---|
| CC001 | Form | Locate |
| CC002 | Other | Locate |
| CC007 | Table | Locate |
| C020101 | Map | Cycle |
| C020201 | Map | Cycle |
| C021001 | Graph | Cycle |
| C021101 | Graph | Cycle |
| C030501 | List | Cycle |
| C030601 | List | Cycle |
| C030701 | Form | Cycle |
| C030702 | Form | Cycle |
| C030703 | Form | Cycle |
| C030705 | Form | Cycle |
| C030708 | Form | Cycle |
| C040501 | Form | Cycle |
| C040502 | Form | Cycle |
| C040503 | Form | Cycle |
| C040504 | Form | Cycle |
| C050101 | Table | Cycle |
| C050201 | Table | Cycle |
| C050501 | Other | Locate |
| C060301 | Table | Cycle |
| C060501 | Table | Cycle |
| C060601 | Table | Cycle |
| C060901 | Table | Higher-order thinking |
| C070501 | Graph | Cycle |
| C071001 | List | Locate |
| C080201 | Other | Locate |
| C080501 | Bill/Form | Locate |
| C080502 | Bill/Form | Locate |
| C080503 | Bill/Form | Locate |
| C080504 | Bill/Form | Locate |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 2-4. Coding of the 2003 quantitative tasks, by type of text content, type of operation, and number of operations: 2003**

| Item number | Type of text content | Type of operation | Number of operations |
|---|---|---|---|
| CC005 | Form | Addition | One |
| CC006 | Other | Subtraction | One |
| C020301 | Persuasive/expository | Addition/multiplication | More than one |
| C020601 | Table | Addition | One |
| C020701 | Table | Addition or multiplication | One |
| C020801 | Table | Multiplication | One |
| C030401 | Other | Subtraction | One |
| C030704 | Form | Division | One |
| C030706 | Form | Multiplication | More than one |
| C030707 | Form | Addition | One |
| C030709 | Form | Addition | One |
| C040401 | Table | Subtraction/multiplication | More than one |
| C040601 | Table | Addition | More than one |
| C040801 | Table | Multiplication | One |
| C050301 | List | Multiplication | One |
| C050601 | Form | Addition | One |
| C050701 | Table | Addition/subtraction | More than one |
| C060701 | Other | Subtraction/multiplication | More than one |
| C060801 | Table | Subtraction | One |
| C070301 | Expository | Addition | One |
| C070601 | Graph | Subtraction | One |
| C070801 | Form | Addition/multiplication | More than one |
| C080101 | Procedural | Addition | One |
| C080801 | Table | Addition or subtraction | One |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 2.4.3 Health Literacy

One of the goals of the 2003 National Assessment of Adult Literacy was to determine the health literacy of the nation's adults and relate their health literacy skills to a variety of demographic characteristics and explanatory variables. Another goal was to directly compare the measures of health literacy with the measures of the general literacy of the population. To accomplish these goals, the assessment included a BQ as well as a set of tasks to simulate real-world decisions about health and health care information and services. The assessment used 14 health stimulus materials and 28 health tasks. All health tasks were also classified as prose, document, or quantitative tasks and were incorporated into those scales as well as into the health scale.

#### 2.4.3.1 Definition of Health Literacy and Purpose of the Health Literacy Assessment

The content of the Health Literacy Component (HLC) was determined by the U.S. Department of Health and Human Services (HHS) in accordance with the public health priorities represented in Healthy People 2010, the disease prevention and health promotion agenda for the nation, and in consultation with HHS staff and external health literacy experts. The Office of Disease Prevention and Health Promotion (ODPHP) of the HHS is the coordinating office for Healthy People 2010 and the lead agency for the Health Communication Focus Area in Healthy People 2010. The Health Communication Focus Area includes a national objective to improve the health literacy of those with marginal or inadequate literacy skills. As the lead agency, ODPHP organized a multiagency collaboration to identify topic domains, stimulus materials, and items that should be included in the HLC. In addition, ODPHP consulted with established health literacy experts outside the federal government about the appropriateness of the selected stimulus materials and items.

The goals of Healthy People 2010 are to increase the quality and years of healthy life and to eliminate health disparities. The measurement of the population's health literacy is key to understanding the methods and interventions that will be necessary to achieve these goals. The HLC of the NAAL offers a vehicle by which HHS can obtain a baseline measurement of the U.S. population's health literacy skills and project a target for improvement by the end of the decade.

The Institute of Medicine and Healthy People 2010 defines health literacy as

> *the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions.* (U.S. Department of Health and Human Services 2000 and Institute of Medicine 2004)

Some studies have suggested that low health literacy can lead to poor communication between patients and health care providers and, ultimately, to poor health outcomes, including increased hospitalization rates, less frequent screening for diseases such as cancer, and disproportionately high rates of disease and mortality (Baker et al. 1998; Gordon et al. 2002; Lindau et al. 2001; Williams et al. 2002). Patients with low health literacy may also be more likely to visit hospital emergency rooms for their care than patients with higher levels of health literacy (Baker et al. 2004). These findings have implications for the costs of caring for patients with low health literacy.

As the Committee on Health Literacy of the Institute of Medicine wrote,

> Health literacy is of concern to everyone involved in health promotion and protection, disease prevention and early screening, health care maintenance, and policy making. Health literacy skills are needed for dialogue and discussion, reading health information, interpreting charts, making decisions about participating in research studies, using medical tools for personal or family health care—such as a peak flow meter or thermometer—calculating timing or dosage of medicine, or voting on health or environment issues. (Institute of Medicine 2004, p. 31)

#### 2.4.3.2 Health Item Development

The health items were developed at the same time as the other prose, document, and quantitative items, following the same guidelines. Stimulus materials for the health items were suggested by HHS, and the development of the health tasks was a cooperative venture between HHS and the NAAL staff.

#### 2.4.3.3 Types of Health Literacy Tasks

The Health Literacy Component of the 2003 NAAL was organized around three domains of health and health care information and services: clinical, prevention, and navigation of the health care system. The domains represent clusters of key types of health and health care information and services that the general population in the United States might be likely to encounter. The stimulus materials and the associated tasks in the HLC were selected to cover these three domains. The tasks were designed to elicit respondents' knowledge and skills for locating and understanding health-related information and services and to represent the three general literacy scales―prose, document, and quantitative―developed to report the results of the NAAL.

The clinical domain encompasses those activities associated with the health care provider-patient interaction, clinical encounters, diagnosis and treatment of illness, and medication. Examples are filling out a patient information form for an office visit, understanding dosing instructions for medication, and following a health care provider's recommendation for a diagnostic test.

The prevention domain encompasses those activities associated with maintaining and improving health, preventing disease, intervening early in emerging health problems, and engaging in self-care and self-management of illness. Examples are following guidelines for age-appropriate preventive health services, identifying signs and symptoms of health problems that should be addressed with a health professional, and changing eating and exercise habits to decrease the risks for developing serious illness.

The navigation of the health care system domain encompasses those activities related to understanding how the health care system works and individual rights and responsibilities. Examples are understanding covered and noncovered benefits for health insurance plans, determining eligibility for public assistance programs, and being able to give informed consent for a health care service.

The distribution of the health literacy items across the three domains of health literacy is summarized in table 2-5.

**Table 2-5. NAAL health items, by distribution across the clinical, prevention, and navigation health domains: 2003**

| Item number | Clinical | Prevention | Navigation of the health care system |
|---|---|---|---|
| CC002 | | | X |
| CC007 | | X | |
| C020901 | | X | |
| C021001 | | X | |
| C021101 | | X | |
| C030101 | | X | |
| C030201 | | X | |
| C030301 | | X | |
| C040501 | | | X |
| C040502 | | | X |
| C040503 | | | X |
| C040504 | | | X |
| C040601 | | X | |
| C040801 | | | X |
| C050801 | | | X |
| C050901 | | | X |
| C051001 | | | X |
| C051101 | | | X |
| C060501 | | X | |
| C060601 | | X | |
| C070101 | X | | |
| C070901 | | X | |
| C071001 | | X | |
| C071101 | | | X |
| C080101 | X | | |
| C080201 | X | | |
| N110101 | | X | |
| N110201 | | X | |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 2.4.4 Cognitive Laboratory Analyses

A total of 66 cognitive laboratory interviews were completed over 4 months to evaluate cognitive questions newly developed for the 2003 assessment. These one-on-one interviews followed a format similar to that described for the BQ cognitive laboratory interviews in section 2.2.2. Participants were taught to think-aloud, then were asked to work through a cognitive question while expressing their thoughts orally. Interviewers administered probes after each question to further elicit how participants worked through each question. When participants got an item wrong, interviewers probed to determine whether they could not do the item or they misunderstood the item. Interviewers also asked participants whether the stimulus material associated with the question was something that they had encountered before and whether it was similar to the types of things they regularly read. At the end of the interview, participants were also asked about other types of materials they regularly read. These questions ensured that the assessment included some stimulus materials that were familiar to respondents from a wide variety of different backgrounds.

The cognitive laboratory interviews were split into two rounds. Thirty-six interviews were conducted during round 1, and 30 interviews were conducted during round 2. A total of 98 stimulus materials and 271 items were evaluated during the cognitive laboratories.

Presented below is a brief description of the protocol makeup for each round of cognitive laboratory sessions along with the demographic breakdown of the cognitive laboratory participants.

**Round 1.** During the round 1 interviews, 6 protocols were used. Each protocol consisted of 8 stimulus materials with 3 to 5 items per stimulus material. Because of the length of 2 of the stimulus materials and the number of associated questions, the questions for these stimulus materials were split in half. A total of 46 stimulus materials and 137 items were tested. Six adults were interviewed per protocol for a total of 36 interviews.

Recruitment was done as described in section 2.2.2. Respondents for the cognitive laboratory interviews were recruited to ensure diversity on the following demographic characteristics:

- age (over 55, 55 or younger);
- native language (English, non-English);
- educational attainment (high school student, GED/high school grad or lower, some college, college graduate);

- Race/Ethnicity (White, Black, Asian/Pacific Islander, Hispanic, Multiracial, Other); and
- Parenthood Status (Parent, Not a parent).

Table 2-6 shows the demographic breakdown of the 36 cognitive laboratory participants for round 1.

**Table 2-6. Number of NAAL assessment cognitive laboratory participants – round 1, by selected characteristics: 2003**

| Demographic characteristic | Number of participants |
|---|---|
| **Gender** | |
| Male | 17 |
| Female | 19 |
| **Race/ethnicity** | |
| White | 10 |
| Black | 16 |
| Hispanic | 2 |
| Asian/Pacific Islander | 2 |
| Multiracial | 3 |
| Other | 3 |
| **Age** | |
| 16–18 | 8 |
| 19–25 | 8 |
| 25–55 | 13 |
| 55 and over | 7 |
| **Language** | |
| Nonnative English speakers | 7 |
| Native English speakers | 29 |
| **Education** | |
| Currently high school student | 6 |
| College degree | 9 |
| No college degree | 15 |
| No high school diploma | 6 |
| **Parent** | |
| Yes | 15 |
| No | 21 |

NOTE: Black includes African American, Pacific Islander includes Native Hawaiian, and Hispanic includes Latino. Race categories exclude Hispanic origin unless specified.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Round 2.** During the round 2 interviews, 6 protocols were administered. Each protocol consisted of 9 stimulus materials with 3 to 5 items per stimulus material. Two stimulus materials from round 1 were modified and included in round 2. A total of 54 stimulus materials (2 from round 1) and 139 items (5 from round 1) were tested. The number of stimulus materials increased by 1 in this round because all the core stimulus materials and items were tested at this time; these stimulus materials were easier to process and had fewer and easier items associated with them than the noncore stimulus materials and items. Five adults were interviewed per protocol for a total of 30 interviews.

Table 2-7 shows the demographic breakdown of the 30 cognitive laboratory participants for round 2.

**Table 2-7. Number of NAAL assessment cognitive laboratory participants – round 2, by selected characteristics: 2003**

| Demographic characteristic | Number of participants |
|---|---|
| **Gender** | |
| Male | 13 |
| Female | 17 |
| **Race/ethnicity** | |
| White | 13 |
| Black | 9 |
| Hispanic | 3 |
| Asian/Pacific Islander | 2 |
| **Age** | |
| 16–18 | 6 |
| 19–25 | 9 |
| 26–55 | 7 |
| Over 55 | 8 |
| **Language** | |
| Nonnative English speakers | 7 |
| Native English speakers | 23 |
| **Education** | |
| Currently high school student | 7 |
| College degree | 9 |
| No college degree | 12 |
| No high school diploma | 2 |
| **Parent** | |
| Yes | 9 |
| No | 21 |

Note: All numbers do not sum to 30 because of missing data on some of the characteristics for some of the participants. Black includes African American, Pacific Islander includes Native Hawaiian, and Hispanic includes Latino. Race categories exclude Hispanic origin unless specified.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

## 2.5 FLUENCY ADDITION TO THE 2003 NATIONAL ASSESSMENT OF ADULT LITERACY

The Fluency Addition to the NAAL (FAN) comprised five separate exercises that measured a variety of aspects of a respondent's oral reading fluency and basic reading skills. Each exercise was timed, which allowed both accuracy and processing efficiency (i.e., rate), a key characteristic of proficiency in skilled, fluent reading, to be measured. These measures follow:

- digit reading and processing rate: expressed as the number of numbers correctly read per minute from a list of one-digit numbers in random order;
- letter reading and processing rate: expressed as the number of letters read correctly per minute from a list of letters in random order;
- decoding: expressed as the number of pseudowords read correctly per minute from a list of pseudowords;
- word recognition: expressed as the numbers of words read correctly per minute from a list of words; and
- passage reading: measured by the numbers of words read correctly per minute from a text passage.

### 2.5.1 Purpose

The FAN was developed to assess the basic reading skills of adults as a complement to the functional literacy focus of the NAAL. Analyses of FAN will relate the basic literacy skills of adults to NAAL scale scores and identify the point at which improvements in the basic skills of America's adults reach a plateau on the NAAL scales. At the point on the NAAL scale where the basic literacy skills of adults level off, it can be assumed that factors other than basic reading skills, such as critical thinking skills, contribute to higher performance on the NAAL. An important benefit of the FAN is that it will help ensure that the 2003 assessment provides meaningful information about the basic reading skills of adults at the lower levels of the prose, document, and quantitative scales.

### 2.5.2 Development Procedures

The FAN digit and letter lists were developed with a random number and letter generator. The development procedures for the FAN word lists, pseudoword lists, and passages are described in the next sections of this chapter. The word lists were based on similar lists used by Richard L. Venezky and John Sabatini in the Study of Adult Reading Acquisition (SARA) (Sabatini et al. 2000a; Sabatini et al. 2000b).

### 2.5.2.1 Word Lists[3]

Three classes of real words were selected from the Kucera and Francis (K-F) (1967) corpus: (1) two- to five-letter, one-syllable words, (2) two-syllable words, and (3) three- to five-syllable words. These words were then combined to form the three lists. The goal was to construct three lists in which the structure of words became progressively more complex while maintaining a relatively high word frequency and familiarity for a general population. The numbers of letters and syllables per word were the primary indices of complexity that were varied because these have been repeatedly shown to be good indicators of word-naming accuracy and response rates. In selecting new words for the lists, NAAL staff began with a subset of words that occur 95 times per million or higher. Whenever lower-frequency words were chosen for inclusion, they were selected from the SARA lists to allow comparisons with other data available for these words.

Proper names were excluded, as were words ending in -s or –ed (with two exceptions, *news* and *needs*);[4] however, no attempt was made to restrict parts of speech of word types (e.g., adjectives, nouns, verbs). An attempt was made to avoid any words that may have common alternate pronunciations (e.g., *read, wind*). In forming parallel lists, an attempt was made to separate any words that appeared closely related with respect to phonology, orthography, or semantics and therefore might cause confusion for respondents or scorers (*it/at, then/than, yes/no, more/most*).

**List 1.** An initial set of 403 two- to five-letter, one-syllable words with K-F frequencies of 100 words/million or higher was identified. K-F 100 (log 2.0) corresponds to the lowest frequency band used to construct the SARA word lists. Error rates on words from the first band were very low, even for participants with the lowest literacy levels in that study. Thirty additional words from higher bands used on the SARA study were also included for consideration because prior data on adult learner performance on these words are available.

To form lists 1a and 1b, 20 two-letter, 32 three-letter, and 32 four-letter words were randomly selected; 20 of the 30 SARA words were selected as part of this set. These words were randomly assigned to list 1a or list 1b. Five-letter words were excluded from list 1 because SARA study results indicate that

[3] This section was written by John Sabatini and Richard L. Venezky.

[4] Words ending in "s" or "ed" were dropped for two reasons. The first is that frequency tables often code the inflectional form of words with different frequencies than the base form. The second reason is that it would have complicated scoring to have to make decisions about how to treat respondents who dropped the "s" or "ed" from the end of a word when reading orally because some groups of respondents may routinely drop these sounds even if they recognize a word. *News* and *needs* were retained because they had been used in previous studies and comparability was deemed to be important. In addition, with regard to *news,* there was a desire to use the noun form of the word (which requires having an "s" at the end).

this additional letter may substantially increase the information processing and decoding complexity. Each list began with 5 two-letter words. The next 13 words were a random mix of two- and three-letter words; the remaining words were a mix of three- and four-letter words. Some reordering of words was then done to reduce possible phonological, orthographic, or semantic confusion (e.g., *in/is/it; bad boy*).

**List 2.** List 2 comprised one- and two-syllable words with a range of four to eight letters. Eighteen four-letter and 20 five-letter, one-syllable words were randomly selected from the initial list of 403 words. Nineteen two-syllable words from the SARA word lists were then identified. Finally, an initial set of 339 three- to eight-letter, two-syllable words with K-F frequencies of 95 words/million or higher were identified, then the 23 most frequent words from this two-syllable list were selected. Lists 2a and 2b were sequenced in random order.

**List 3.** List 3 comprised two- to five-syllable words with a range of four to eleven letters. From the set of 339 two-syllable words previously discussed, the next 22 most frequent words were selected. Eighteen two- or three-syllable words from the SARA word lists were also selected. Finally, the 44 most frequent words in the three- to five-syllable range were selected (only one five-syllable word, *university,* was selected by this procedure). List 3 was sequenced in ascending order first by letters and then by K-F frequency.

### 2.5.2.2 Pseudoword Lists[5]

Three classes of pseudowords were developed, with subclasses within each: (1) simple, invariant; (2) simple, variant; and (3) multisyllabics. In all of these classes, a variety of pseudowords were constructed to test decoding ability. All the pseudowords followed strict structural rules for English words. For example, no single-vowel item ended in a single <s, l, f> because these letters usually double in such positions (e.g., *class, call, off*). That is, the consonant-vowel-consonant (CVC) class with final <s, l, f,> was restricted to a small group of mostly function words and shortened forms: *of, Al, is, as, el, us.* Less frequently occurring letters (e.g., <j, x, z>) were used sparingly, and no items had common pronunciations that sounded like common English words. (It is nearly impossible to totally avoid pseudowords that sound like rarer dictionary entries.)

**Simple, invariant class.** The simple, invariant class had primarily CVC items for which the consonant portion varied from a simple consonant to a digraph (e.g., <ch>) to a cluster of consonants (e.g., <sp>). The vowels varied from simple vowels (<a, e, i, o, u, y>) to digraph vowels that are invariant

[5] This section was written by John Sabatini and Richard L. Venezky.

or nearly so (e.g., <oa, oi>). These items were constructed to avoid patterns that vary by dialect (e.g., C+<og>) or that have more than one common pronunciation (e.g., <gi-> as in *girl* and *giant*). A few items with <ch> and <th> were included, however, even though they could prove to be unreliable.

Initial <th> is pronounced as in *then* and *the* in initial position in function words; otherwise it is pronounced as in *thin*. (That is, in all pseudowords in initial position, it should be pronounced as in *thin*.) In final position, <th> is voiced in only a small number of verbs. Decoding studies show that respondents rarely give the voiced pronunciation to unfamiliar words with initial <th>. For <ch>, three pronunciations are possible, as in *cheese, chalet,* and *chord*. The first of these, however, is by far the most common.

The digraph vowels selected, <ai/ay, au/aw, ee> and so on except for <ea>, have a single, common pronunciation each. However, exceptions occur for all of them (e.g., *coyote, broad, aisle, been*). Nevertheless, the exceptions were not found to serve as models for pronouncing pseudowords with these spellings. Where variant pronunciations might occur, they were indicated in the pronunciation key.

**Simple, variant class.** The simple, variant class contained items for variant pronunciations of <c>; the final <e> vowel pattern; and a special class of vowel+<r> pronunciations. Both the hard (/k/) and soft (/s/) pronunciations of <c> were tested. For <g>, however, only the hard pronunciation was tested because a large number of exceptions exist for what should be the soft pronunciation, and many of these are common words (e.g., *get, gear, girl*).

A large number of items were constructed for the final <e> pattern because this is probably the most important variant decoding pattern learned in the primary grades. For each of the main vowels (<a, e, i, o, u>), two test items were constructed. For<y>, only one item was included.

Two items tested a vowel+<r> pattern in which the vowel and the following <r> coalesce to a single, <r>-colored vowel, which is represented in dictionaries as if it were a sequence of /U/+/r/, as in *her, fir,* and *burn*.

**Multisyllabics.** As a group, the multisyllabic pseudowords posed the largest challenge to testing because of the potential variation in pronunciation of each. Nevertheless, patterns with minimal potential variability were selected. These patterns varied from two to four syllables and assessed stress placement, long-short vowel shifts before specific suffixes and before doubled letters or consonant clusters, and the <-le> pattern. Although the major expected pronunciations were indicated in the pronunciation key, many other pronunciations of unstressed vowels were possible. For example, the first vowel in *decrift* and *recilf* could be schwa or any of the vowels pronounced in *rid, red,* or *bead*.

A number of words tested both stress and vowel quality before suffixes that condition stress placement. These included the <ic> and the <ity> pseudowords (*corelic, setric, tronic, vortastic, instamic, mertosity, lorsinity, contremity*). Other pseudowords that had predictable stress placement included those that end in <ing>, <ly>, <ious>, and a few other suffixes.

### 2.5.2.3 Passages

Twenty-seven passages were originally developed for the FAN assessment, and 16 of those passages were included in the field test. Passages were identified from a variety of sources—including texts for ESL classes and children's magazines—that fit into the content areas of the 2003 assessment. Texts were selected at two reading levels: one representing mid- to upper-elementary school (approximately grades 3, 4, and 5) and one representing middle school (approximately grades 7 and 8). Both Lexile and Fry readability procedures were used to rate texts for their grade level.[6] Two passages were written specifically for ALSA respondents at approximately a grade 2 level but were dropped after the field test. For the operational assessment, ALSA respondents were asked to read one passage at the easier level. All passages at the easier level were narrative text. At the more difficult level, both expository and narrative texts were included.

Unlike the main assessment, in which all texts were authentic and reproduced in their original format and presentation, FAN texts were all reformatted into a large, easy-to-read font with extra space between lines for ease of reading. Texts were also edited to revise complex clauses and other elements of sentence structure that made reading aloud difficult. This was done to get a measure of oral reading fluency that was not influenced by things unrelated to a respondent's ability to recognize and decode text.

For the field test, two comprehension questions were associated with each passage. These comprehension questions kept the respondents focused on the meaning of what they were reading. The instructions given to respondents told them to "read it out loud quickly but at a speed where you can understand what you are reading." They were told that they would be asked some questions after reading the passage. For the operational assessment, only one comprehension question was asked. The comprehension questions were not scored.

The properties of the passages in the operational assessment are summarized in table 2-8.

---

[6] Lexile measures the complexity of a text on the basis of semantic difficulty (vocabulary) and syntactic complexity (sentence length). See www.lexile.com for more information. Fry measures the difficulty of a text on the basis of the average number of syllables per word and the average numbers of words per sentence.

**Table 2-8. Oral reading fluency passages included in the Fluency Addition to the NAAL (FAN) operational assessment, by selected properties: 2003**

| Text title | Lexile score | Fry score | # Words | # Sentences | Average # words per sentence | # Simple verbs | # Com-pound verbs | # Coord-inate clauses | # Subord-inate clauses |
|---|---|---|---|---|---|---|---|---|---|
| Bigfoot | 1020 | 13 yrs | 186 | 12 | 16 | 26 | 4 | 3 | 13 |
| Chicken Soup | 1100 | 14 yrs | 153 | 10 | 15 | 12 | 6 | 0 | 9 |
| Curly | 380 | 7 yrs | 151 | 17 | 9 | 16 | 9 | 5 | 5 |
| Lori Goldberg | 1030 | 12 yrs | 156 | 8 | 20 | 13 | 9 | 4 | 7 |
| Exercise | 1020 | 15 yrs | 182 | 11 | 17 | 13 | 13 | 1 | 6 |
| Grand Canyon | 570 | 11 yrs | 166 | 17 | 10 | 17 | 1 | 0 | 1 |
| Guide Dogs | 700 | 9 yrs | 156 | 13 | 12 | 16 | 6 | 2 | 3 |
| My Friend Amanda | 700 | 11 yrs | 155 | 12 | 13 | 16 | 4 | 1 | 8 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 2.5.3 Cognitive Laboratory Procedures

The NAAL word lists, pseudoword lists, and passages were divided into three protocols for evaluation in cognitive laboratories. Twenty-seven cognitive laboratory interviews were conducted, three for each protocol. Participants for the FAN cognitive laboratories were recruited through a variety of sources, including ads in a local newspaper, flyers distributed at churches and community centers, and word of mouth. In addition, some participants from earlier rounds of cognitive laboratories on the background questionnaire and the cognitive assessment were contacted and asked whether they were interested in participating in another interview. Participants were paid $50 and transportation costs.

The FAN cognitive laboratory interviews had two primary purposes. The first was to identify any sequences of words on the word lists or pseudoword lists that were difficult for fluent respondents to read. The second purpose was to identify any words or pseudowords that were particularly problematic for nonnative English speakers. Given these goals, all participants in the FAN cognitive interviews were required to have a minimum of a high school education (or GED) to ensure that they were reasonably fluent in reading English. Interviews with nonfluent respondents would not identify word sequences that were likely to cause problems for a fluent reader. An effort was also made to include participants from a wide variety of language backgrounds.

The demographics of the FAN cognitive laboratory participants are shown in table 2-9.

**Table 2-9. Oral reading fluency cognitive laboratory participants, by selected characteristics: 2003**

| Demographic characteristic | Number of participants |
|---|---|
| **Gender** | |
| Male | 9 |
| Female | 18 |
| **Race/ethnicity** | |
| White | 8 |
| Black | 11 |
| Hispanic | 3 |
| Other | 5 |
| **Age** | |
| 16–25 | 7 |
| 26–55 | 11 |
| Over 55 | 5 |
| Declined to state | 4 |
| **Native language** | |
| English | 15 |
| Spanish | 3 |
| Other[1] | 9 |
| **Highest educational attainment** | |
| High school graduate/GED | 18 |
| Some college or associate's degree | 3 |
| College graduate | 6 |

[1] Includes Venda, Creole/French, Swahili, Turkish, Russian, Thai, Portuguese, Italian, and Chinese.
NOTE: Black includes African American and Hispanic includes Latino. Race categories exclude Hispanic origin unless specified.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 2.5.4 Automatic Scoring of FAN Lists and Passages

To automatically score oral reading performances, NAAL staff developed methods to produce base measures of oral reading accuracy and accurate oral reading rate. As part of the scoring process, NAAL used automatic speech recognition technologies. A speech recognition system has several components. One is an acoustic model. This is a representation of the sounds, or phonemes**,** produced when speaking the English language. To accommodate foreign accents, NAAL staff developed acoustic models trained on both native and nonnative speakers of English.

Another component of the speech recognition system is a dictionary. The dictionary lists the most common pronunciations for each word that the system should recognize. Every word that appeared in the

FAN materials was entered into the system's dictionary. In addition, entries were created for common substitutions of words in the source text.

A third component is the language model. This is a representation of the sequence of words the speaker is expected to say. For example, if the respondent is asked to read a passage that begins "Curly is my big black dog," then it is very likely that the reader will say the words "Curly is my big black dog." The high probability associated with this string of words is encoded in the language model for this passage. The language models contain not only the most likely strings of words that a reader is expected to say but also the types of mistakes and disfluencies readers are most likely to make. The reading errors can be represented as a list of "rules" (X goes to Y) with a probability associated with each one. For example, if the printed word is *a* and readers commonly say the word *the,* the rule for this reading error would be "a goes to the."

The acoustic models, dictionary, and language models are important inputs to statistical methods used by the speech recognition engine to formulate a hypothesis of what the speaker said. The speech recognition system identifies the string of words that best matches the respondent's speech, and this hypothesis is compared with the source text. Using a standard string alignment algorithm that minimizes the number of word deletions, substitutions, and insertions, the system aligns the respondent's response with the correct response. The reading errors are then tallied and weighted, and a final value of the number of words read correctly is generated.

Other information is also extracted from the respondent's utterance, such as the duration of speech, the rate of speech, and pause duration. These values are output as other data products in the machine score. The resulting machine scores provide the base measures for assessing the respondent's basic oral reading skills.

## 2.6 ADULT LITERACY SUPPLEMENTAL ASSESSMENT[7]

The Adult Literacy Supplemental Assessment (ALSA) was an alternative, performance-based assessment that allowed adults with marginal literacy skills to demonstrate what they could and could not do when asked to make sense of various forms of print. The ALSA assessment started with simple identification tasks and sight words and moved to connected texts, using authentic, highly contextualized materials commonly found at home, in workplaces, or in the community. The ALSA allowed low-literate adults to demonstrate to what extent they could navigate print materials by drawing heavily on visual

[7] The introduction to this section was written by Heide Spruck Wrigley.

information along with their knowledge of logos and sight words and their ability to process print that, although fairly simple, was largely text-based.

The assessment reflected the current perspective on reading that posits that the process of deriving meaning from even minimal print resides in the interaction among the reader, the text, and the context of the reading act. Readers make sense of print by drawing on their knowledge of the pragmatic, semantic, syntactic, and phonemic systems of the language to "make meaning" within a particular context. Readers' knowledge of the world and their experiences with certain text forms merge with alphabet knowledge and knowledge of sound-symbol relationships to allow meaning to emerge. Thus, readers access multiple knowledge sources in the brain (linguistic as well as world knowledge) to use literacy in meaningful ways.

The conceptual framework guiding ALSA also drew from studies showing the significant difference between language that is context-embedded (context cues are transparent, concepts are fairly concrete, and information is familiar to the reader) and language that is context-reduced (context is abstracted or must be derived and a series of inferences may have to be made) (Cummins 1979). The distinction between context-embedded and context-reduced texts applies to all forms of reading, including texts that challenge adults. Adults who are new to literacy tend to do much better with texts that are highly embedded in contexts and tasks. These adults depend largely on background knowledge that has been acquired through interaction with high-frequency, everyday print (a Coca-Cola can) or with print commonly found in the home (the local electricity bill). Interest may play a role as well. Stories about people in similar circumstances or about disasters that have been talked about in the news may provide the impetus to engage with print at more than just a surface level. Although low-literate adults may be successful in deriving print from high-interest, context-embedded, everyday print that is supported by visual information, they may still encounter a great deal of difficulty if similar information is presented in more abstract forms, as part of an article or in a newsletter. Unlike the main portion of the NAAL, which relied on print for which context was reduced, the ALSA used highly context-embedded forms of print.

Finally, the ALSA drew on work being done in sociolinguistics in the area of literacy practices. It supported the view that literacy does not consist solely of a set of skills that an individual does or does not have. Rather, it sees reading occurring as a part of a sociocultural context that either inhibits or facilitates understanding (Barton and Hamilton 1998; Hamilton 2000; Hill and Parry 1992; Street 1998, 2001). The sociolinguistic perspective underlying the ALSA also highlights the pragmatics of literacy, illustrating that much of adult reading happens through "literacy events" (e.g., everyday interactions with literacy) as flyers are read and shopping lists are made, magazines or manuals are flipped through, headlines are glanced at, directories are consulted, and posters and signs are noticed (Halliday and Hasan 1985;

Halliday 2002; Widdowson 1983, 1990). For adults, reading is a purposeful act, attempted and accomplished within a specific context. In asking adults to engage in meaningful texts that remain whole and that reflect reading tasks common in everyday life, the ALSA sought to determine to what extent low-literate adults manage to derive meaning from print and to what extent they are able to use decoding skills as aids in that process.

### 2.6.1 Purpose

Data from the 1992 NALS suggested that up to 10 percent of the 2003 sample of adults would not be able to take the 2003 assessment. Rather than have no information about the literacy abilities of that group of adults, NAAL staff developed the Adult Literacy Supplemental Assessment (ALSA) for those respondents. The goal of the ALSA was to provide descriptive information on the literacy skills of this segment of the population by assessing their ability to read common, everyday materials they would be likely to encounter and use in daily life. Instead of respondents being asked to read questions and answer in writing, ALSA questions were read to respondents and they answered orally. Questions were read in either English or Spanish, and respondents answered in either language, but all the stimulus materials they were asked to read were in English. Responses on the seven core questions were used to screen respondents into either the main assessment or the ALSA.

### 2.6.2 Development Procedures

The ALSA was based on an assessment that AIR administered to more than 400 adult ESL literacy students as part of the What Works Study for Adult ESL Literacy Students, a national study of ESL instructional practices. Development work for adapting this assessment to the 2003 NAAL focused on choosing appropriate stimulus materials for a national assessment, refining the questions, standardizing the administration of the assessment, and standardizing and simplifying the scoring rubrics so that responses could be scored by field interviewers. Standardization of administration was done through an interactive process involving iterative cognitive laboratory interviews followed by more formal pilot testing.

The ALSA consisted of nine stimulus materials with 7 to 11 questions associated with each stimulus material. The first 2 questions associated with each stimulus material were familiarity questions, designed to both determine whether the respondent knew what the stimulus material was and to ease the respondent into the assessment with simple questions. These 2 questions asked the respondent what the stimulus material was and where one would be likely to see, purchase, or use it. These questions were followed by assessment questions to measure literacy. The assessment questions can be classified as letter identification (pointing to a letter read by the interviewer), word identification, word reading, and comprehension. The next-to-last question associated with each stimulus material asked the respondent whether he or she ever used or saw things similar to the stimulus material. If the respondent replied yes, a follow-up question was posed, asking whether the respondent had ever read the stimulus material before this assessment.

The distribution of the different types of assessment questions is shown in table 2-10.

**Table 2-10. Number of Adult Literacy Supplemental Assessment (ALSA) questions, by type of question: 2003**

| Type | Number of questions |
|---|---|
| Letter identification | 5 |
| Word identification | 9 |
| Word reading | 12 |
| Comprehension | 19 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 2.6.3 Cognitive Laboratory Procedures

Cognitive laboratory interviews to evaluate the ALSA assessment were held at centers teaching adult basic education classes in the Washington, D.C., area. All ALSA cognitive laboratory participants were enrolled in adult basic education classes, indicating they had literacy levels that matched the ALSA target population. Flyers were given to teachers to distribute to their students. Participants were paid $20. See section 2.2.2 for a discussion of cognitive laboratory procedures.

## 2.7 BIAS AND SENSITIVITY REVIEW OF NAAL ITEMS

AIR project staff performed the initial review of the NAAL items for bias and sensitivity. All items were also reviewed by panelists who did not work on the development of the items. Bias and sensitivity reviews are intended to identify items that include material that is not related to the construct being measured but that may interfere with a respondent's performance on an item.

### 2.7.1 Expert Panels

Prior to the field test, NAAL cognitive items were reviewed by the following panelists:

- Vivian Gadsden, National Center on Fathers and Families, University of Pennsylvania
- Peggy McGuire, Equipped for the Future
- Emily Miller Payne, Southwest Texas State University
- Carlos Rodriguez, American Institutes for Research
- John Sabatini, National Center on Adult Literacy at the Graduate School of Education, University of Pennsylvania
- Mary Dunn Siedow, North Carolina Literacy Resource Center
- Sondra Stein, Equipped for the Future, National Institute for Literacy
- Heide Spruck Wrigley, Aguirre International

Final blocks of NAAL items for the operational assessment were reviewed by the following panelists:

- Peter Afflerbach, University of Maryland
- Miriam Burt, Center for Applied Linguistics
- Michael Kamil, Stanford University
- John Sabatini, University of Pennsylvania

Final scoring rubrics were reviewed by the following panelists:

- Peter Afflerbach, University of Maryland
- Charles Peters, University of Michigan

FAN passages were reviewed by the following panelists:

- Scott Baker, University of Oregon
- Lynn Fuchs, Vanderbilt University
- Michael Kamil, Stanford University
- John Sabatini, University of Pennsylvania

- Richard Venezky, University of Delaware
- Joanna Williams, Columbia University

ALSA items were reviewed by the following panelists:

- Michael Kamil, Stanford University
- Pardee Lowe, U.S. Department of Defense
- Emily Miller Payne, Southwest Texas State University
- Victoria Purcell-Gates, Michigan State University
- Mary Dunn Siedow, North Carolina Literacy Resource Center
- Elvira Swender, American Council on the Teaching of Foreign Languages
- Heide Spruck Wrigley, Aguirre International

### 2.7.2 Identifying Biased and Sensitive Items

To determine whether a stimulus material or an item was biased, reviewers independently reviewed each stimulus material and item for

- stereotypes;
- recognition of population diversity; and
- familiarity and accessibility.

Panelists individually rated each stimulus material and question as either "Accepted" or "Rejected" on the basis of the presence of bias. For every rejected stimulus material and item, panelists were asked to explain why they believed that the item was biased and to suggest revisions. A group consensus was reached on whether each item should be accepted, rejected, or revised.

## 2.8 FIELD-TEST BOOKLET DESIGN

NAAL field-test booklets used a partial spiral design so that each block appeared in each position in the block, but every block did not appear with every other block. Table 2-11 presents the spiral patterns for the test booklets.

**Table 2-11. NAAL field-test booklet design: 2003**

| Booklet # | Position 1 | Position 2 | Position 3 | Position 4 |
|---|---|---|---|---|
| 1 | Core 1 | 1.1 | 1.2 | 1.3 |
| 2 | Core 2 | 2.1 | 2.2 | 2.3 |
| 3 | Core 3 | 3.1 | 3.2 | 3.3 |
| 4 | Core 4 | 4.1 | 4.2 | 4.3 |
| 5 | Core 1 | 1.2 | 1.3 | 1.1 |
| 6 | Core 2 | 2.2 | 2.3 | 2.1 |
| 7 | Core 3 | 3.2 | 3.3 | 3.1 |
| 8 | Core 4 | 4.2 | 4.3 | 4.1 |
| 9 | Core 1 | 1.3 | 1.1 | 1.2 |
| 10 | Core 2 | 2.3 | 2.1 | 2.2 |
| 11 | Core 3 | 3.3 | 3.1 | 3.2 |
| 12 | Core 4 | 4.3 | 4.1 | 4.2 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

## 2.9 MAIN ASSESSMENT BOOKLET DESIGN

This section describes the booklet design for the main assessment and the FAN.

### 2.9.1 2003 National Assessment of Adult Literacy Cognitive Items

The goal in assembling cognitive blocks for the operations assessment was to have

- seven blocks;
- 11 questions per block balanced between scales;
- four to six stimulus materials per block;
- a range of difficulty corresponding to the items that were being replaced (match top and bottom of distribution, range of items in the middle); and
- a distribution of processes/strategies corresponding to the items that were being replaced.

The steps for selecting items and assembling blocks follow:

1. Delete from the pool all items with differential item functioning (DIF; see chapter 4 for an explanation of DIF), interrater reliability problems, and discrimination problems (based on an analysis of field-test data).
2. Delete from the pool all noncore items with field test p-values below .20 and above .90 (outside the range of items that were being replaced).

3. Begin creating a pool of items that will be used to assemble blocks by selecting seven health stimulus materials (one for each of the seven blocks). Select items on the basis of coverage of the three health areas (prevention, clinical, navigation of the health system), p-value (items wanted with a range of difficulty to construct a health scale), distribution across literacy scales (prose, document, quantitative), and distribution by type of process/strategy.

4. Add all items based on the almanac, the Medicare and You brochure, the colon cancer pamphlet, and the NAAL newspaper that were not eliminated in step 1 to the pool of items used to assemble blocks.

5. Add stimulus materials that were developed to replace specific items that were being replaced (map, graph, check, order form) to the pool of items for block assembly.

6. Analyze the distribution of items that were selected for block assembly in steps 3, 4, and 5 in terms of scale, p-value, and type of match.

7. From the remaining items, select items as needed (including additional health items) to balance the pool of items in terms of the characteristics listed in step 6 and to create a total pool of 77 questions balanced across the prose, document, and quantitative scales.

8. Assemble from the pool blocks of selected items so that each block has 11 questions distributed across the three scales, has a range of p-values (some easy and some difficult items), and can be completed in approximately 15 minutes.

After the new items were distributed among the seven blocks, the blocks were spiraled into booklets as shown in table 2-12. The 2003 NAAL used the same spiral design as the 1992 adult literacy assessment. See the *Technical Report and Data File User's Manual for the 1992 National Adult Literacy Survey,* pages 90 to 91, for a discussion of the BIB spiral design.

**Table 2-12. NAAL operational booklet design: 2003**

| Booklet number | Block numbers contained in booklet: Position 1 | Position 2 | Position 3 | Position 4 |
|---|---|---|---|---|
| 1 | Core | 1 | 2 | 13 |
| 2 | Core | 2 | 3 | 9 |
| 3 | Core | 3 | 4 | 7 |
| 4 | Core | 4 | 13 | 8 |
| 5 | Core | 13 | 9 | 6 |
| 6 | Core | 9 | 7 | 10 |
| 7 | Core | 7 | 8 | 11 |
| 8 | Core | 8 | 6 | 12 |
| 9 | Core | 6 | 10 | 5 |
| 10 | Core | 10 | 11 | 1 |
| 11 | Core | 11 | 12 | 2 |
| 12 | Core | 12 | 5 | 3 |
| 13 | Core | 5 | 1 | 4 |
| 14 | Core | 1 | 3 | 8 |
| 15 | Core | 2 | 4 | 6 |
| 16 | Core | 3 | 13 | 10 |
| 17 | Core | 4 | 9 | 11 |
| 18 | Core | 13 | 7 | 12 |
| 19 | Core | 9 | 8 | 5 |
| 20 | Core | 7 | 6 | 1 |
| 21 | Core | 8 | 10 | 2 |
| 22 | Core | 6 | 11 | 3 |
| 23 | Core | 10 | 12 | 4 |
| 24 | Core | 11 | 5 | 13 |
| 25 | Core | 12 | 1 | 9 |
| 26 | Core | 5 | 2 | 7 |

NOTE: Block 1 and blocks 9 through 13 were originally used in the 1992 adult literacy assessment. Blocks 2 through 8 were newly developed for the 2003 assessment. This booklet design was used for both the 1993 NALS and the 2003 NAAL.
SOURCE: Campbell, A., Kirsch, I.S., and Kolstad, A. (1992). *Assessing Literacy: The Framework for the National Adult Literacy Survey.* Washington, D.C.: National Center for Education Statistics. U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 1992 National Adult Literacy Survey.

The FAN texts were spiraled into 16 groups with two texts in each group. Each group included an easy text and a difficult text. In addition to the texts, each respondent was asked to read number lists, letter lists, word lists, and pseudoword lists. Table 2-13 shows how the texts were combined.

**Table 2-13. Booklet layout for operational Fluency Addition to the NAAL (FAN) administration: 2003**

| Tab | Page | Stimulus |
|---|---|---|
| 1 | 1-1 | Curly |
| 1 | 1-2 | Exercise |
| 2 | 2-1 | Guide Dogs |
| 2 | 2-2 | Lori Goldberg |
| 3 | 3-1 | Grand Canyon |
| 3 | 3-2 | Bigfoot |
| 4 | 4-1 | Amanda and I |
| 4 | 4-2 | Chicken Soup |
| 5 | 5-1 | Curly |
| 5 | 5-2 | Lori Goldberg |
| 6 | 6-1 | Guide Dogs |
| 6 | 6-2 | Bigfoot |
| 7 | 7-1 | Grand Canyon |
| 7 | 7-2 | Chicken Soup |
| 8 | 8-1 | Amanda and I |
| 8 | 8-2 | Exercise |
| 9 | 9-1 | Curly |
| 9 | 9-2 | Bigfoot |
| 10 | 10-1 | Guide Dogs |
| 10 | 10-2 | Chicken Soup |
| 11 | 11-1 | Grand Canyon |
| 11 | 11-2 | Exercise |
| 12 | 12-1 | Amanda and I |
| 12 | 12-2 | Lori Goldberg |
| 13 | 13-1 | Curly |
| 13 | 13-2 | Chicken Soup |
| 14 | 14-1 | Guide Dogs |
| 14 | 14-2 | Exercise |
| 15 | 15-1 | Grand Canyon |
| 15 | 15-2 | Lori Goldberg |
| 16 | 16-1 | Amanda and I |
| 16 | 16-2 | Bigfoot |
| 17 | 17-1 | Practice numbers |
| 17 | 17-2 | Speeded numbers |
| 17 | 17-3 | Practice letters |
| 17 | 17-4 | Speeded letters |
| 17 | 17-5 | Practice words |
| 17 | 17-6 | Word list 1 |
| 17 | 17-7 | Word list 2 |
| 17 | 17-8 | Word list 3 |
| 17 | 17-9 | Practice pseudowords |
| 17 | 17-10 | Pseudoword list 1 |
| 17 | 17-11 | Pseudoword list 2 |
| 17 | 17-12 | Pseudoword list 3 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 2.9.2 2003 National Assessment of Adult Literacy Fluency Items

The FAN passages were matched with assessment booklets as shown in table 2-14. ALSA respondents were administered FAN passages on the basis of the number of the booklet they used for completing the core items. ALSA respondents were asked to read only the first (easier) passage associated with their assigned FAN tab.

**Table 2-14. Mapping between assessment booklets and Fluency Addition to the NAAL (FAN) passages: 2003**

| Booklet number | FAN tab | FAN tab |
|---|---|---|
| 1 | 1 | 17 |
| 2 | 2 | 17 |
| 3 | 3 | 17 |
| 4 | 4 | 17 |
| 5 | 5 | 17 |
| 6 | 6 | 17 |
| 7 | 7 | 17 |
| 8 | 8 | 17 |
| 9 | 9 | 17 |
| 10 | 10 | 17 |
| 11 | 11 | 17 |
| 12 | 12 | 17 |
| 13 | 13 | 17 |
| 14 | 14 | 17 |
| 15 | 15 | 17 |
| 16 | 16 | 17 |
| 17 | 1 | 17 |
| 18 | 2 | 17 |
| 19 | 3 | 17 |
| 20 | 4 | 17 |
| 21 | 5 | 17 |
| 22 | 6 | 17 |
| 23 | 7 | 17 |
| 24 | 8 | 17 |
| 25 | 9 | 17 |
| 26 | 10 | 17 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

# CHAPTER 3

# FIELD TEST

Michelle Amsbary, Martha Berlin, Thomas Krenzke, Leyla Mohadjer,
Lyn Clark (retired), and Jacqueline Hogan, Westat

## 3.1 INTRODUCTION

From April through August 2001, staff from the National Center for Education Statistics (NCES) and its contractors worked collaboratively to prepare for the National Assessment of Adult Literacy (NAAL) field test. During this period, a fully automated field test system was developed, and data delivery systems and procedures were implemented. Publicity materials for improving study cooperation rates were designed, and instructional manuals and training programs for supervisors and interviewers were developed.

The field test yielded information about study participation and cooperation rates, the automation and interaction of intricate instruments and management systems, and the training and management of field staff. The evaluation of the field test informed development efforts for the main study, including refinements to the survey instruments and aspects of the CAPI system, the design of the assessment, and the data collection procedures for the main study.

Significant effort was required to convert the screener, background questionnaire, and interviewer guide component of the assessment from paper and pencil to an effective Blaise computer-assisted personal interviewing (CAPI) application.[1] Each instrument was tested and approved before a fully integrated, automated CAPI system for the field test was prepared. In addition to the data collection instruments, the fully automated system included the Interviewer Management System and the Data Management System. These systems were designed to guide the flow of the interviews, allow interviewers and supervisors to manage their case assignments, implement the NAAL contact procedures and rules, generate production reports, transmit data, send e-mail, and provide receipt control functions.

The field test data collection was divided into two separate efforts: the household sample and the volunteer sample. The two-sample design was implemented to achieve the target sample sizes of at least 400 completed assessments in each of the race/ethnicity categories of interest (Blacks, Hispanics, and others), with about 33 percent of the assessments completed by people aged 60 years or older. The

[1]Blaise is a computer-assisted interviewing (CAI) system and survey processing tool developed by Statistics Netherlands.

subgroup sample sizes were necessary for the Differential Item Functioning (DIF) analysis (discussed in section 3.2.2). It was determined that this target could not be reached cost-efficiently through the household sample alone. Therefore, the data collection contractor worked with focus group services in the same locations as the household sample to recruit volunteers to supplement the sample. This enabled the specification of sampling quotas necessary to reach the target sample sizes. The sample design and the data collection instruments, systems, and procedures are described in sections 3.2 and 3.3, respectively.

Interviewers attended a 5-day in-person training on the screener, background questionnaire, and assessment administration, as well as on general contact and administrative procedures. During data collection, two regional supervisors, a field manager, and additional staff regularly monitored production.

Interviewing for the household component of the field test lasted 4 months, beginning in September 2001. Interviews were conducted in 10 primary sampling units (PSUs) across the country by 41 skilled interviewers. The interviewers screened 1,189 of a possible 1,686 occupied households for basic demographic information about household members, for a screener cooperation rate of 70.5 percent. In the screened households with eligible sample persons, 1,176 sample persons were selected for the background questionnaire and assessment. Interviewers were able to complete background questionnaire interviews with approximately 979 sample persons, for a background questionnaire cooperation rate of 83.2 percent. Finally, 972 assessments were completed, for an assessment cooperation rate of 99.3 percent.

The volunteer component of the field test was conducted in 9 of the 10 PSUs from which the household sample was selected. No volunteers were sought in the other PSU because the PSU minority population was small and the sampling quotas would have been very difficult to achieve. Respondents were recruited by focus group services, according to specifications outlined by the data collection contractor's statistical staff. Interviewers and respondents traveled to the focus group facility where interviews were conducted. The volunteer data collection was conducted for approximately 6 weeks, beginning in early November 2001. The interviewers conducted 439 assessments with volunteer respondents.

Respondents from both the household and volunteer components were provided with a monetary incentive to increase cooperation rates. See table 3-13 for the total sample yield from the household and volunteer samples.

As described in section 3.4, data preparation and processing systems and procedures were developed to support the conversion of screener, background questionnaire, and interviewer observation data from Blaise into files for delivery to AIR. The hard-copy assessment booklets were tracked through the Data Management System and sent for scoring. Electronic scoring data were received from the scoring contractor, reviewed by the data collection contractor, and delivered to the contractor responsible for analysis.

## 3.2 SAMPLE DESIGN

The field test target sample sizes were approximately 400 completed assessments from each of the race/ethnicity categories of interest (Blacks, Hispanics, and others), with about 33 percent of the assessments completed by persons aged 60 years or older. To achieve these targets, the field test sample was divided into two components: a household sample and a volunteer sample.

The household sample was selected through a multistage area sample, with a goal of completing approximately 900 assessments within 10 PSUs. Within participating households, eligible respondents were selected with a sampling algorithm programmed into the CAPI screener. The volunteer component included a target sample of 425 assessments from a volunteer sample selected by focus group facilities in 9 of the 10 PSUs.

### 3.2.1 Target Population

The target population for the NAAL field test was persons aged 16 and older who, at the time of the field test, resided in households in the 10 sampled PSUs. The target population included members of dwelling units who lived in college dormitories, but excluded those in group quarters, military barracks, and halfway houses.

### 3.2.2 Target Sample Sizes

The sampling targets for the field test called for at least 100 completed assessments per booklet group for each of the race/ethnicity groups (Blacks, Hispanics, and others), as well as at least 400 completed assessments from persons aged 60 and older. Four distinct booklet groups were used in the field test. To reduce the possibility of question-order effects,[2] the items in these booklets were spiraled and presented in a different order, resulting in a total of 12 booklet types. Table 3-1 provides the

[2]The order in which questions are presented can affect the answers given to subsequent items.

minimum expected sample sizes (by race/ethnicity and age) for one of four booklet groups required for the DIF analysis.[3] The target sample sizes presented in table 3-1 were spread equally across the three spiraled versions of each of the four booklet groups for the sole purpose of the DIF analysis. For example, for booklet group 1, the target sample sizes were distributed among booklets 1, 5, and 9. (See section 2.8 for a discussion of the field test assessment booklet design.)

**Table 3-1. Minimum expected sample sizes per booklet group, by age and race/ethnicity: 2001**

| Race/ethnicity | Total | Age | |
|---|---|---|---|
| | | < 60 years | 60+ years |
| Total | 300 | 201 | 99 |
| | | | |
| Hispanic | 100 | 67 | 33 |
| Black | 100 | 67 | 33 |
| Other | 100 | 67 | 33 |

NOTE: All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as other include non-Hispanics of all other races, including multiracial.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2001 National Assessment of Adult Literacy field test.

The household sample was designed to yield approximately 900 completed assessments. The size of the volunteer sample was derived by subtracting the expected yield from the household sample (by race/ethnicity and age) from the minimum sample sizes (presented in table 3-1) per booklet group, times 4 (for 4 booklet groups). For example, there were only 16 Hispanics age 60+ expected in the household sample. From table 3-1 we see that 33 is the minimum sample size per booklet group (or 33 x 4 = 132 cases across the four booklet groups). Therefore, the volunteer sample was specified to result in 132 – 16 = 116 Hispanics age 60+.

Table 3-2 presents the expected yield by race/ethnicity and age for a target sample of 900 in the household sample. The expected yield by race/ethnicity and age for the volunteer sample is provided in table 3-3. The expected yield by race/ethnicity and age for both the household and volunteer samples is shown in table 3-4.

[3]The DIF analysis was conducted to identify items that should be examined more closely by subject-matter specialists for possible bias and subsequent omission from the testing instrument. The items with potential bias were differentially difficult for members across subgroups with comparable scores.

**Table 3-2. Expected sample yield from the household sample, by age and race/ethnicity: 2001**

| Race/ethnicity[1] | Total (percent) | Age[2] (percent) <60 years | 60+ years |
|---|---|---|---|
| Total | 900 (100) | 717 (80) | 183 (20) |
| Hispanic | 153 (17) | 137 (15) | 16 (2) |
| Black | 221 (25) | 190 (21) | 31 (3) |
| Other | 526 (58) | 390 (43) | 136 (15) |

[1] The distribution is based on 1990 census data for the sampled segments.
[2] The distribution is based on the 2000 Statistical Abstract of the United States, tables 14, 17, and 18.
NOTE: All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as other include non-Hispanics of all other races, including multiracial. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2001 National Assessment of Adult Literacy field test.

**Table 3-3. Expected sample yield from the volunteer sample, by age and race/ethnicity: 2001**

| Race/ethnicity | Total (percent) | Age (percent) <60 years | 60+ years |
|---|---|---|---|
| Total | 425 (100) | 208 (49) | 217 (51) |
| Hispanic | 246 (58) | 130 (31) | 116 (27) |
| Black | 179 (42) | 78 (18) | 101 (24) |
| Other | 0 (40) | 0 (0) | 0 (0) |

NOTE: All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as other include non-Hispanics of all other races, including multiracial.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2001 National Assessment of Adult Literacy field test.

**Table 3-4. Expected sample yield from the household and volunteer samples, by age and race/ethnicity: 2001**

| Race/ethnicity | Total (percent) | Age (percent) <60 years | 60+ years |
|---|---|---|---|
| Total | 1,325 (100) | 925 (70) | 400 (30) |
| Hispanic | 399 (30) | 267 (20) | 132 (10) |
| Black | 400 (30) | 268 (20) | 132 (10) |
| Other | 526 (40) | 390 (29) | 136 (10) |

NOTE: All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as other include non-Hispanics of all other races, including multiracial. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2001 National Assessment of Adult Literacy field test.

As shown in table 3-4, the field test required 1,325 completed assessments to ensure that the minimum sample sizes per booklet group in table 3-1 were met by race/ethnicity and age categories.

### 3.2.3 Household Sample Design

#### 3.2.3.1 Primary Sampling Unit and Segment Sample

The first stage of sampling for the household component of the field test was to select PSUs. In most surveys, segments are formed from blocks within the selected PSUs. However, in an effort to contain resources, the field test used previously sampled segments and unused listings of housing units from an earlier study conducted by the data collection contractor. From the listed sample, 10 PSUs were chosen that yielded 93 sampled segments and approximately 4,700 dwelling units within those segments. The 10 PSUs were selected to be diverse in region, urbanicity, and minority composition.

#### 3.2.3.2 Sample of Households Within Segments

Within the 10 sample PSUs and the 93 sample segments, a sample of households was randomly selected from every household listed in the 93 segments ($N$ = 4,691 dwelling units). The field test implemented the within-segment sampling procedures planned for the main survey. That is, segments were classified as high or low minority on the basis of 1990 census data, and minority households were oversampled in high-minority segments.

To test the sampling procedures to be used in the main study (as described in section 7.1.3.3), a household was defined as a minority household if the person owning or renting the dwelling unit (the reference person) was a member of a minority group. The nonminority households in high-minority segments were subsampled in order to simulate the subsampling procedures planned for the main study. The automated screener administered to the sampled households implemented these sampling and subsampling algorithms and procedures.

#### 3.2.3.3 Sample of Eligible Persons Within Selected Households

The last stage of sampling was the selection of persons 16 years and older from each sample-eligible household. For each screened household, the CAPI screener program determined the age-eligible persons from the household roster provided by the screener respondent. If the household had three or fewer age-eligible persons, one person was selected for the background questionnaire and assessment. If the household had more than three age-eligible persons, two persons were selected. In order to select the person(s), random numbers were generated for each eligible person in the CAPI system; then the random numbers were sorted, and the person with the lowest random number was selected. If there were more than three eligible persons, then the persons with the lowest two random numbers were selected. These sampling procedures were programmed into the CAPI screener.

### 3.2.4 Volunteer Sample Design

The NAAL field test had two competing objectives: to test the data collection and sampling procedures and to achieve the number of completed cases required to finalize the assessment booklet design for the main study. For testing the sampling and data collection procedures, it was appealing to have a balanced-representative sample. For the DIF analysis, it was important to achieve approximately equal sample sizes across race/ethnicity categories. However, given the 93 sample segments selected for the field test, the procedures were not expected to yield the desired minimum sample sizes, even when oversampling procedures were implemented for Blacks and Hispanics. Therefore, to augment the expected household sample yield of 900 completed assessments, a volunteer target sample of 425 respondents was selected to help meet the minimum sample size per booklet group (shown in table 3-1) by race/ethnicity and age.

The volunteers were drawn from 9 of the 10 field test PSUs. The sampling quotas required for the volunteer sample are as follows: within each of the 9 PSUs, the total target sample size was equal to 47 assessments divided between 27 Hispanics (of whom 13 were aged 60 or older) and 20 Blacks (of whom 11 were aged 60 or older).

## 3.3 DATA COLLECTION

### 3.3.1 Data Collection Instruments

In preparation for the field test, a CAPI system was developed. The advantages of CAPI over paper-and-pencil methods arise primarily from the electronic capture of data during the interview, the resolution of inconsistencies at the point of data collection, and the ability of electronic instruments to support more complex designs and methods. CAPI relieves the interviewer from applying complicated sample selection procedures to determine eligible households and respondents. Additional advantages for the NAAL included allowing the computer assignment of assessment booklets to sample persons to better ensure the integrity of the design, improving the quality of background questionnaire data by eliminating most interviewer error, and providing internal edit checks and improved timeliness of data delivery.

A CAPI system also provides enormous benefits by electronically linking the home office, supervisors, and interviewers. An automated case management system allows closer monitoring of cooperation rates, production, and costs that are critical to the management of the fieldwork.

Blaise, a commercial off-the-shelf product of Statistics Netherlands, was determined to be the best product available to implement the interviewing instruments. The Blaise instruments are designed for ease of interviewer usability. The instruments facilitate interviewing by presenting the text in a clear and legible manner and by providing the interviewer with visual cues to the type of text or screen.

Several edits were built into the CAPI system. These included soft edits (those that are verified with the respondent and can be suppressed by the interviewer) and hard edits (those that the interviewer must reconcile before continuing with the interview). Edits were developed as range checks, designed to ensure that a numerical response fell within reasonable limits, as well as those that checked for consistency among multiple data items within the instrument.

The NAAL design specified that the interviewers administer the following instruments to household respondents or eligible sample persons:

- The screener was administered to identify eligible sample persons in the sampled households.
- The background questionnaire collected various demographic data.
- The core and main assessments were administered to each sample person through a combination of a hard-copy assessment booklet and a CAPI interviewer guide. This interviewer guide included the interviewer observation items.

The CAPI instrumentation developed for NAAL is described in sections 3.3.1.1 through 3.3.1.3. The interviewer and supervisor management systems are discussed in sections 3.3.2.1 and 3.3.2.2, respectively.

#### 3.3.1.1 Screener

The screener instrument was administered once the interviewer had established contact with the household, verified the household’s street address, and introduced the study, and took approximately 5 minutes to administer. The CAPI screener contained the elements described here:

- a household enumeration that obtained relationship to the head of household (reference person), gender, and age for all household members;

- questions to collect race and ethnicity for those aged 16 and older; and
- questions on home and alternate telephone numbers.

The CAPI system used household information collected in the screener to implement the sampling procedures for determining whether the household was eligible and which sample person(s) should be interviewed for the background questionnaire and assessment. See appendix X for the exact questions included in the screener.

#### 3.3.1.2 Background Questionnaire

The background questionnaire was an approximately 25-minute instrument that included questions on a variety of topics, including general and language background; educational background and experience; political and social participation; labor force participation; literacy practices; job training and skills; demographic information; family literacy; household income and welfare participation; health questions; and additional demographics.

Demographic information collected during the screener interview was directly imported into the background questionnaire. The CAPI program controlled the background questionnaire instrument flow by determining, on the basis of age and gender, which questions should be asked and which should be skipped. Additionally, questions that were determined to be inapplicable based on responses to earlier items were not asked.

#### 3.3.1.3 Core and Main Assessments

The field test used four unique assessment booklet groups. Each booklet began with a core assessment, followed by three blocks of main assessment items. (See section 2.8 for a discussion of the field test design.) The blocks in the four booklets were then spiraled to reduce the possibility of question order effects. That is, the three blocks in each of the four booklets appeared in each position (first, second and third) in a separate booklet. Therefore, the three booklets in each booklet group contained the same items, organized in the same blocks, but presented in a different order. This resulted in a total of 12 booklets.

The average administration time for each booklet group varied. As table 3-5 shows, booklet group 3 (blocks 3.1, 3.2, and 3.3) had the shortest average administration times, and booklet group 1 (blocks 1.1, 1.2, and 1.3) had the longest times. Block 1.2 was particularly long because of the taped oral response item.[4]

**Table 3-5. Average administration times for item blocks: 2001**

| Booklet group | Block | Average administration time (minutes) |
|---|---|---|
| 1 | 1.1 | 8.7 |
| | 1.2 | 12.3 |
| | 1.3 | 9.5 |
| 2 | 2.1 | 11.8 |
| | 2.2 | 8.0 |
| | 2.3 | 10.4 |
| 3 | 3.1 | 7.5 |
| | 3.2 | 7.2 |
| | 3.3 | 7.9 |
| 4 | 4.1 | 8.4 |
| | 4.2 | 8.4 |
| | 4.3 | 9.6 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2001 National Assessment of Adult Literacy field test.

The assessment administration also required the use of several stimulus materials, including an almanac, a colon cancer pamphlet, a *"Medicare and You"* brochure, a newspaper, a tape recorder, and a calculator.

The interviewer guide was the CAPI accompaniment to the hard-copy assessment booklets and was programmed to accommodate instructions for all 12 versions of the assessment booklet in one CAPI program. The CAPI interviewer guide was used in conjunction with the assessment booklet to help the interviewer guide the sample person through the assessment and record interview times, observations, and status codes.

The CAPI interviewer guide consisted of the following components:

- assessment booklet double-blind identification (ID) number entry (this module contained numerous edits to verify that a valid booklet was used for each case);
- core instructions in English and Spanish;
- core scoring procedures;
- general assessment directions;

[4] See section X.X for a discussion of this item, which required the interviewer to tape record the sample person's response to the assessment item.

- item-by-item main assessment task instructions in English; and
- interviewer observation questions.

The interviewer guide also indicated which booklet number (1-12) should be administered to each sample person #1 in the household. In households where two sample persons were selected, interviewers selected the booklet from a spare bundle of booklets. The spare bundles were spiraled to ensure a random distribution of booklet groups.

#### 3.3.1.4 Language of Administration

The CAPI screener and background questionnaire were translated into Spanish, as were the instructions (but not stimulus materials) for the core assessment items. The main assessment was available only in English. See exhibit 3-1 for a complete description of the rules of administration for the NAAL instruments.

**Exhibit 3-1. Translation guide, NAAL data collection: 2001 field test**

| Instrument | Interviewer administers instrument in English | Bilingual interviewer administers Spanish-translated instrument | Bilingual neighbor or household member administers Spanish-translated instrument | Bilingual neighbor or household member translates instrument into language other than English or Spanish |
|---|---|---|---|---|
| Screener | X | X | X | X |
| Background questionnaire | X | X | X | |
| Core exercise | X | X | | |
| Main exercise | X | | | |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2001 National Assessment of Adult Literacy field test.

### 3.3.2 General Features of the Computer-Assisted Interviewing Systems

#### 3.3.2.1 Interviewer Management System

The interviewers' assigned cases and other study activities were managed with an integrated software system, the Interviewer Management System. This system had various features that were accessed by using one of two laptop modes of operation: stand-alone mode or online mode.

The Interviewer Management System stand-alone mode provided the following capabilities:

- case browse, for an interviewer to review assignments;
- status review for the case, as well as its individual tasks;
- the ability to launch and conduct CAPI instruments; and
- entry of status codes and other information on an Electronic Record of Calls.

The Interviewer Management System online mode provided the following capabilities:

- time and expense reporting;
- shipping of case materials;
- data transmission; and
- e-mail.

#### 3.3.2.2 Supervisor Management System

The Supervisor Management System was designed as a component of the Data Management System. Supervisors used this system to manage the case work within their region. The Supervisor Management System was designed with a look and feel very similar to the Interviewer Management System. The numerous functions of the Supervisor Management System are described below:

- The Assignments link was used to review, assign, reassign, and unassign cases to interviewers.
- The Case Browse navigation link was designed to search for cases by using various criteria.
- The Case Details function allowed supervisors to set the final status of cases.

- The T&E link was used to review a subset of time and expense data that interviewers recorded.

- The Reports navigation link was used to produce all study reports. (See section 3.3.5.3 for a discussion of the reports.)

- The Receipt Control link was used to track and monitor the distribution and use of the assessment booklets within the system.

### 3.3.3 Interviewer Materials

#### 3.3.3.1 Advance Materials

Considerable effort was made to develop introductory materials that would convince respondents of the study's legitimacy and importance. Appearance, content, and word choice were all important considerations. The introductory materials included a toll-free telephone number (the hotline), which respondents could call to obtain more information about the study or verify the study's legitimacy. Spanish-language versions of all introductory materials were produced as well.

**Advance letter.** Prior to the interviewer's first contact with the household, an advance letter was mailed with a brochure to all sampled households for which a complete address was available. The letter introduced the study, identified the sponsor, stated the study's purpose, and asked for cooperation.

**Brochure.** The informative and attractive brochure included with the letter explained the study in detail and underlined the importance of participation.

**U.S. Department of Education letter of introduction.** This letter was signed by the NCES Project Officer and verified that the interviewer was an authorized representative of the U.S. Department of Education.

**Community authorization letter.** This general letter was designed to be shown to apartment managers, postal employees, police departments, or other professional people whom interviewers might encounter in the community. It provided assurances that the interviewer was not selling or soliciting but was a trained professional working on a government-sponsored education study.

**Sorry-I-missed-you card.** This card was left when the interviewer visited a household and no one was home.

**Nonresponse letters.** Several versions of nonresponse letters were developed and sent to households that refused or were otherwise reluctant to participate in the study.

#### 3.3.3.2 Household Folder

One household call record folder was produced for each case ID in the sample. The household folder helped interviewers keep track of the status of all cases in their assignment. The Record of Actions was located on the back cover of the household folder. Interviewers used this grid to record the status and outcome of every contact attempt with the household.

#### 3.3.3.3 Handcards

Interviewers received sets of bound handcards designed to facilitate the flow of the interview and improve the efficiency of the sample persons' reporting. The handcards were used during the screener and background questionnaire interviews. Two handcards, which listed race and ethnicity response options, were used for the screener. The background questionnaire had 10 handcards, which listed the sets of response categories used most frequently throughout the interview, such as "a lot, some, a little, or none." The content of all 12 handcards is located in appendix X. The handcards were also translated into Spanish.

### 3.3.4 Field Staff Recruitment and Training

The following sections describe the recruitment of field staff, the overall training approach, the supervisor and interviewer training sessions, and general interviewing techniques (GIT) and CAPI Train.

#### 3.3.4.1 Field Staff Recruitment

Two regional supervisors participated in recruiting interviewers for their regions. A total of 42 interviewers were recruited in early August 2001. To maximize the efficiency of interviewer travel and ensure a knowledge of the local geography and population, interviewers were hired from the areas in which interviewing assignments were located.

All candidates were screened by telephone for availability, level of interest in the project, related job experience, and general ability to communicate personably and effectively. Those who met the basic qualifications were invited to a job interview, conducted by a NAAL regional supervisor or field manager. The job interview included a mock personal interview to assess the candidate's general ability to read aloud and follow directions.

#### 3.3.4.2 Approach to Training

The data collection contractor's basic approach to training is to maximize trainees' involvement and participation in the training, provide ample opportunity for supervisory staff to observe and evaluate trainee performance, and provide trainees with detailed reference documents.

Throughout training, interviewers received close attention and were given extensive hands-on experience with the CAPI system and interviewer materials. Training scripts and exercises build in complexity and were designed to address situations that interviewers were likely to encounter. The following techniques were used to train interviewers on all instruments:

- **Demonstration interview.** On the first day of project-specific training, two members of the training staff presented a demonstration of the entire interview, through the use of the CAPI programs and the appropriate interview and stimulus materials.
- **Interactive lectures.** The basic concepts of the instruments were taught through interactive lectures that increased in complexity as the training progressed. The trainees were led through the CAPI instruments and were called on to act the role of the interviewer while the trainer played the respondent and interjected relevant training points.
- **Role-playing exercises.** The trainers arranged the trainees in pairs, taking into consideration their strengths and weaknesses. Each trainee had the opportunity to play the role of both the interviewer and the respondent while the training staff observed and corrected the trainees as needed.
- **Exercises.** The trainees completed several practice exercises during training, including exercises on core assessment scoring procedures, noninterview report forms (NIRFs), entry of information into the Electronic Record of Calls, and collection of industry and occupation information.
- **Paid respondent practice.** The trainees conducted practice interviews with paid respondents, who were recruited by a local focus group facility.

Before interviewer training, project staff developed three reference manuals that documented the field test instruments and procedures. These materials are described below:

- **Supervisor manual.** This document, used exclusively by the two regional supervisors and the field manager, covered all the study procedures, as well as the Supervisor Management System and the use of reports for monitoring work in their regions.

- **Interviewer manual.** This document contained an overview of the study and detailed information on all data collection procedures. It documented all screening procedures and listed question-by-question specifications for each question in the screener, the background questionnaire, and the interviewer guide. Hard-copy versions of the CAPI screens were included for reference.

- **Trainer's guide.** The trainer's guide included all lecture scripts, role-playing exercises, and written exercises used in training. The trainers were required to use the trainer's guide to ensure standardization of the materials presented in the interviewer training sessions.

#### 3.3.4.3 Supervisor Training

The two regional supervisors, as well as the field manager, attended a 20-hour supervisor training in August 2001. The training included specifics on the administration of the interview instruments and supervisory responsibilities for the quality control of interviewer production. The supervisors were presented with an abbreviated version of the interviewer training session. They also received additional training in the supervisor-specific sections of the Supervisor Management System and in supervisory techniques.

#### 3.3.4.4 Interviewer Training

Forty-one interviewers completed the 28-hour project-specific training session in September 2001. A separate 1-day Spanish training session was held for the bilingual interviewers, nine of whom completed the Spanish training. This separate training session was required for the bilingual interviewers to train them on the administration of the screener, background questionnaire, and core assessment items. See section 3.3.1.4 for a discussion of the language administration procedures implemented for each NAAL instrument. In addition to the study instruments, interviewers were trained on the use of the advance materials, such as the brochure and study instruction, which were also translated into Spanish. Separate training materials were developed in Spanish for this session.

#### 3.3.4.5 General Interviewer Techniques and CAPI Train

Before project-specific training, the 11 interviewers new to the data collection contractor and social science interviewing attended a 5-hour GIT training session. These interviewers also completed a home study with written exercises that they turned in before the formal GIT training began. The in-person GIT training program included an audiovisual presentation, interactive participation, written exercises, and a question-and-answer period. Each interviewer received a GIT manual that documented the presentation.

A total of 41 interviewers attended the CAPI Train session. This session consisted of a self-administered tutorial that introduced the interviewers to the procedures for using Blaise software to conduct a CAPI interview. The tutorial instructed trainees on CAPI question types, function keys, and special commands.

### 3.3.5 Conduct of the Data Collection Effort

#### 3.3.5.1 Field Staff and Organization

The staff responsible for field test data collection included the project director, the field director, and two regional supervisors. The field director reported directly to the project director, the regional supervisors reported to the field director, and the field interviewers reported to one of the two supervisors.

#### 3.3.5.2 Management

For purposes of field operations, the 10 field test PSUs were divided into two regions, each headed by one regional supervisor. The supervisor's primary responsibility was day-to-day oversight of the production of about 20 interviewers each, including attention to cost, cooperation rates, shipment of closed-out work, and quality control tasks. The field director coordinated field operations for the household and volunteer samples and maintained close contact with the regional supervisors on issues of production, cost, cooperation rates, shipment of closed-out work, and other issues.

#### 3.3.5.3 Data Collection Monitoring With Computer-Generated Reports

As mentioned in section 3.3.2.2, the Supervisor Management System was created for the field test. One of the Supervisor Management System components was a Reports mechanism. Reports were updated in real time as an interviewer or supervisor entered information or transmitted data.

The reports allowed all levels of management to monitor daily the progression of completion rates, cooperation rates, and distribution of cases in pending codes by region, interviewer, PSU, and segment for each of the survey instruments (screener, background questionnaire, and assessment).

### 3.3.6 Data Collection Procedures

Data collection for the household component of the field test extended over 14 weeks, beginning September 13, 2001 and concluding December 22, 2001. Data collection for the volunteer sample extended over about 6 weeks, starting in November 2001. The volunteer sample enabled the data collection contractor to specify sampling quotas that were necessary for achieving the sample targets. Respondent contact and interview scheduling procedures for both the household and volunteer samples are discussed below.

### 3.3.6.1 Contact Protocol

#### 3.3.6.1.1 Household Sample

The interviewer's first task was to conduct an in-person screening interview with a respondent at each of the sampled dwelling units that was occupied. Interviewers were instructed to administer the screener to a household member aged 16 or older. The screener was conducted to identify eligible sample persons in the household.

When a sample person was selected during the screening interview and was available at that time, the interviewer attempted to complete the background questionnaire and the assessment during this same visit. For the remaining sample persons, the interviewer set an appointment time to return to complete the background questionnaire and the assessment. If a sample person demonstrated reluctance to participate, either through numerous broken appointments or a voiced refusal, the interviewer discussed further attempts and strategies with the supervisor.

A total of 1,176 sample persons were identified through completed screeners. Of these sample persons, 979 completed the background questionnaire and 972 completed the assessment.

#### 3.3.6.1.2 Volunteer Sample

Focus group facilities in 9 of the 10 field test PSUs located and scheduled interviews with participants who possessed the target sample characteristics (race/ethnicity and age).

The facilities were provided with appointment dates and times for scheduling participants, as well as a plan for reporting the progress of contact and scheduling efforts. Interviewers traveled to the facilities at the scheduled times and administered the screener, the background questionnaire, and the assessment at the facility.

A total of 441 volunteers completed the background questionnaire, and 439 completed the assessment.

#### 3.3.6.2 Scheduling Protocol

##### 3.3.6.2.1 Household Sample

Interviewers were required to complete the background questionnaire and the assessment with the identified sample person in the same visit. The interviewer informed the sample person of the amount of time needed to complete the two instruments and either transitioned directly into the interview after completing the screener or set up a future appointment.

##### 3.3.6.2.2 Volunteer Sample

All three instruments—the screener, the background questionnaire, and the assessment—were completed in one sitting with the volunteer respondents.

#### 3.3.6.3 NAAL Hotline

Along with a letter, each household received a brochure in the mail outlining the important points of the field test. The brochure included a toll-free hotline number for respondents to call with any additional questions. Respondents could call this number to verify the legitimacy of the survey, verify the identity of the interviewer, and reschedule appointments. If a respondent called outside normal business hours, a message (recorded in both English and Spanish) asked the respondent to leave his or her name and telephone number for a return call the next business day. A total of 20 calls were received by the hotline.

### 3.3.7 Quality Control Measures and Feedback to Staff

#### 3.3.7.1 Quality Control Measures

In addition to the CAPI range and logic checks and home office review of completed cases, the following quality control measures were used to ensure high-quality work in the data collection phase of the field test.

**Taped interviews.** Interviewers were instructed to tape their 3rd and 10th interviews for supervisory review. Respondent permission was secured in advance of taping. Listening to the tapes allowed the supervisor to evaluate the interviewer-respondent interaction, the flow of the interview, and general interviewer techniques. Forty-three interviews were successfully taped and reviewed.

**Observation.** Fifteen interviews were observed by the data collection contractor's home office staff, the field director, and regional supervisors. The observations were a valuable tool for evaluating interviewer performance, as well as the procedures and materials used to conduct the field test.

**Validation.** Validation was performed on both completed and nonresponse cases. Key information was asked of the respondent to confirm that the interview had taken place. A total of 112 validations were completed. The validation effort identified one occupied dwelling unit, incorrectly finalized as a non-dwelling unit address, in a segment populated mainly by seasonal dwelling units. All other validations were rated as acceptable.

### 3.3.7.2 Feedback to Staff

#### 3.3.7.2.1 Weekly Newsletter

The supervisors and interviewers received nine newsletters during the data collection period. The newsletters were distributed by e-mail as well as in hard-copy format. Prepared by home office staff, the newsletters clarified questions raised by the interviewers and supervisors, as well as situations observed by supervisors, project staff, and systems staff through the various quality control measures. The newsletter was also a motivational tool because it included cumulative completion goals and the number of completed interviews for the household and volunteer samples.

#### 3.3.7.2.2 Weekly Supervisor Calls

**Interviewer reports to the supervisor.** During scheduled weekly telephone calls, the supervisors and the interviewers discussed all aspects of the interviewer's work, including production, costs, performance and cooperation rates, quality control results, timeliness of the receipt of completed cases by the home office, survey materials needed by the interviewer, and any e-mails and hard-copy memos and newsletters outlining procedural changes.

**Supervisor reports to the home office.** Once a week, each supervisor held a scheduled telephone call with the field director to discuss progress and problems in the region, prospects and plans for completing the remaining work, and what help, if any, the supervisor needed to complete all work in the region by the end of the field period.

#### 3.3.7.2.3 E-mail

The e-mail system implemented for the field staff was Microsoft Outlook Express. The field director, home office staff, and supervisors and interviewers used e-mail to communicate quickly and effectively with one another. The ability to communicate electronically greatly facilitated the rapid exchange of ideas and information and, in many cases, replaced telephone calls and mailed correspondence.

### 3.3.7.3 CAPI Help Desk

A CAPI help desk was established and operated by staff specially trained in the NAAL instrumentation. If interviewers or supervisors experienced technical problems with the CAPI system during the field test, they could call the toll-free help desk number and receive assistance in resolving the problem. The help desk received 465 calls during the data collection period.

## 3.3.8 Cooperation Rates

### 3.3.8.1 Cooperation Rates for the Field Test

The field test was designed to evaluate data collection procedures, the automation and interaction of the instruments, and the training and management of field staff. It also was to provide information on study participation and cooperation rates but was not intended to be a robust test of achievable main study cooperation rates. Nonetheless, best practices to build cooperation rates were followed. The cooperation rate factors are described below and in considerably more detail in chapter 8.

### 3.3.8.2 General Cooperation Rate Factors

Cooperation rates on household studies are influenced by three broad categories of factors:

- the ability of the interviewers to obtain cooperation;
- the effectiveness of callback procedures; and
- the efforts made by interviewers and supervisors to convert initial nonresponse cases to completed interviews.

**Interviewers' ability to obtain cooperation.** The field test interviewers were trained in techniques for handling reluctant respondents, answering questions, and avoiding refusals. To assist the interviewers in gaining respondent cooperation, a cover letter and brochure were mailed to all sampled households before the interviewer attempted to complete a screener. Additionally, respondents who completed the background questionnaire and attempted the assessment were paid $30. This monetary assessment was provided upon completion of the interview.[5]

**Callback procedures.** Interviewers were trained to make trips at different times of day (morning, afternoon, or evening), taking into account that late afternoon and evenings would be the most productive hours in most cases. They were also trained to make trips on different days of the week and on weekends (Saturday or Sunday).

**Efforts to convert nonresponse.** Each type of nonresponse case required a different strategy for conversion. Refusals were the most difficult type of nonresponse case to convert. When a respondent refused or broke off an interview, the interviewer captured information about the reason for the refusal. Using this information, the interviewer and the supervisor determined the course of follow-up action.

Procedures were also designed to boost productivity when (1) a respondent was not at home (interviewers were supplied with sorry-I-missed-you cards that could be left on the doorstep), (2) the respondent had a language barrier (whenever possible, interviewers fluent in Spanish were sent to Spanish-speaking households), (3) the respondent was too ill to participate, or (4) the dwelling unit was vacant.

### 3.3.8.3 Field Test Results

Cooperation rate and sample monitoring tables were created and updated weekly to monitor the sample distribution and sample yield in the household sample, volunteer sample, and total sample. The actual distributions of the sample were compared with the expected distributions by race/ethnicity, age, and booklet group. Finally, sample validation, a standard quality control procedure, was conducted to verify the sampling operation used by the CAPI system.

During the household component of the field test, the number of assessed Blacks and the number of assessed persons aged 60 and older were lower than expected. After identifying the shortfall in respondents aged 60 and older, more emphasis was placed on collecting data from such respondents in the volunteer sample. The shortfall in the number of assessed Blacks resulted because outdated 1990 census block data (the most recent data available at the time) were used to calculate the expected yield for the

[5] Based on two incentive experiments conducted as part of the National Adult Literacy Survey in 1991 and 1992, it was confirmed that a $20 incentive increased the cooperation rate. For the NAAL, an inflation factor was added, resulting in a $30 incentive.

household sample (as shown in table 3-2), and the race/ethnicity distribution changed in the field test census blocks since 1990.

#### 3.3.8.4 Summary of Cooperation and Cooperation Rates and Actual Sample Sizes

Throughout the data collection period, sample sizes and cooperation rates were monitored. The following subsections summarize the information that was collected.

##### 3.3.8.4.1 Probability Sample

**Sample yield.** Table 3-6 provides the actual sample yield[6] by race/ethnicity and age for the household sample. The household sample yielded a much larger percentage of Hispanics than expected; 29 percent of the assessments were completed by Hispanics, although only 17 percent were expected (see table 3-2). The household sample yielded a smaller percentage of both Blacks and others than expected; 20 percent of the assessments were completed by Blacks and 52 percent were completed by others, compared with expected sample yields of 25 percent and 58 percent, respectively. This sample also yielded a smaller percentage of people aged 60 or older: 15 percent of assessments versus an expected yield of 20 percent.

**Table 3-6. Actual sample yield from the household sample, by age and race/ethnicity: 2001**

| Race/ethnicity | Total[1] (percent) | Age (percent) <60 years | Age (percent) 60+ years |
|---|---|---|---|
| Total | 967 (100) | 819 (85) | 148 (15) |
| | | | |
| Hispanic | 276 (29) | 261 (27) | 15 (2) |
| Black | 189 (20) | 162 (17) | 27 (3) |
| Other | 502 (52) | 396 (41) | 106 (11) |

[1] The total accounts for 967 of the 972 completed assessments for which race/ethnicity and age data were available.
NOTE: All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as other include non-Hispanics of all other races, including multiracial. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2001 National Assessment of Adult Literacy field test.

**Cooperation rates.** Table 3-7 compares the expected cooperation rates for the screener, the background questionnaire, and the assessment with the actual cooperation rates in the household sample. The target screener cooperation rate was not met because data collection was terminated soon after the target of approximately 900 assessments was reached. However, the background questionnaire and

[6] The total accounts for 967 of the 972 completed assessments for which race/ethnicity and age data were available.

assessment rates exceeded expectations. As the table shows, the screener cooperation rate was 71 percent. The overall cooperation rate, calculated as the screener cooperation rate multiplied by the background questionnaire/assessment cooperation rate, was 58 percent. The expected cooperation rates for the screener and background questionnaire were taken from the contractual rates required for these instruments in the main study. The expected cooperation rate for the assessment was based on the 1992 National Adult Literacy Study (NALS) assessment rate. The overall rate is the product of the screener, background questionnaire, and assessment rates. The screener cooperation rate was computed as the set of households that completed the screener, including those screened out due to subsampling rules, divided by the set of occupied dwelling units in the sample. The background questionnaire cooperation rate was computed as the set of completed background questionnaires, including those not completed due to language problems or mental disabilities,[7] divided by the number of sampled persons. The assessment cooperation rate was computed as the set of cases that were fully complete or partially complete (except due to refusal) or fully incomplete if due to reading/writing barrier, language problem, mental disability, physical disability, or lost in mail due to mail courier error, divided by the set of completed background questionnaires.

**Table 3-7. Expected and actual cooperation rates from the household sample, by survey component: 2001**

| Cooperation rate type | Expected cooperation rate (percent) | Actual cooperation rate (percent) |
|---|---|---|
| Screener | 85.0 | 70.5 |
| Background questionnaire | 80.0 | 83.2 |
| Assessment | 93.2 | 99.3 |
| Overall | 63.3 | 58.3 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2001 National Assessment of Adult Literacy field test.

The specific reasons for screener, background questionnaire, and assessment nonresponse during the field test are shown in tables 3-8 through 3-10. The tables show that the major reasons for nonresponse to the screener and background questionnaire are due to refusal, or not at home after maximum calls. Once the background questionnaire was completed, 99 percent of sampled persons cooperated in taking the assessment.

[7] Mental disability cases were not added to the background questionnaire numerator in the field test since disposition codes did not distinguish between mental and physical disabilities.

**Table 3-8. Screener nonresponse in the household sample, by reason for nonresponse: 2001**

| Reason | Household sample Number | Household sample Percent |
|---|---|---|
| All sampled dwelling units | 1,843 | † |
| Occupied dwelling units | 1,686 | † |
| | | |
| Incompletes | 497 | 29.5 |
| Refusal/breakoff | 245 | 14.5 |
| Not home after maximum calls | 200 | 11.9 |
| Illness/disability | 7 | 0.4 |
| Language problem/unavailable | 4 | 0.2 |
| Other[1] | 41 | 2.4 |
| | | |
| Completed screener interview | 1,189 | 70.5 |

† Not applicable.
[1] Locked buildings/gated communities.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2001 National Assessment of Adult Literacy field test.

**Table 3-9. Background questionnaire nonresponse in the household sample, by reason for nonresponse: 2001**

| Reason | Household sample Number | Household sample Percent |
|---|---|---|
| Sampled persons | 1,176 | † |
| | | |
| Incompletes | 197 | 16.7 |
| Refusal | 96 | 8.2 |
| Not home after maximum calls | 64 | 5.4 |
| Mental/physical disability | 7 | 0.6 |
| Other[1] | 30 | 2.6 |
| | | |
| Completed interviews | 979 | 83.3 |

† Not applicable.
[1] Unavailable during field period; field period ended.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2001 National Assessment of Adult Literacy field test.

**Table 3-10. Assessment nonresponse in the household sample, by reason for nonresponse: 2001**

| Reason | Household sample | |
|---|---|---|
| | Number | Percent |
| Completed interviews | 979 | † |
| Incompletes | 7 | 0.7 |
| Partial complete/refused | 4 | 0.4 |
| Refused | 3 | 0.3 |
| Completed exercises[1] | 972 | 99.3 |
| Partial complete | 25 | 2.6 |
| Reading/writing barrier | 7 | 0.7 |
| Language problem | 16 | 1.6 |
| Mental disability | 2 | 0.2 |
| Physical disability | 0 | 0.0 |
| Fully incomplete | 3 | 0.3 |
| Physical disability | 3 | 0.3 |

† Not applicable.
[1] The total of 972 completed assessments includes 6 cases lost in the mail due to mail courier error.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2001 National Assessment of Adult Literacy field test.

### 3.3.8.4.2 Volunteer Sample

The actual sample yield by race/ethnicity and age for the volunteer sample is provided in table 3-11. The volunteer sample yields by race/ethnicity very closely matched what was expected (see table 3-3). Of the 439 volunteer assessments completed, 56 percent were completed by Hispanics (58 percent were expected), 42 percent were completed by Blacks (42 percent were expected), and 2 percent were completed by others (0 percent were expected). However, the volunteer sample yielded a smaller than expected percentage of completed assessments by those aged 60 and older (47 percent actual versus 51 percent expected).

**Table 3-11. Actual sample yield from the volunteer sample, by age and race/ethnicity: 2001**

| Race/ethnicity | Total (percent) | Age (percent) <60 years | Age (percent) 60+ years |
|---|---|---|---|
| Total | 439 (100) | 234 (53) | 205 (47) |
| Hispanic | 247 (56) | 135 (31) | 112 (26) |
| Black | 184 (42) | 95 (22) | 89 (20) |
| Other | 8 (2) | 4 (1) | 4 (1) |

NOTE: All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as other include non-Hispanics of all other races, including multiracial. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2001 National Assessment of Adult Literacy field test.

### 3.3.8.4.3 Overall Sample

The overall size of the combined household and volunteer sample is provided in table 3-12. As in the household sample, Hispanics completed a larger portion of the assessments than expected (shown in table 3-4), and Blacks, others, and respondents aged 60 and older completed a smaller portion of assessments than expected. Hispanics completed 37 percent of the assessments (compared with an expected sample yield of 30 percent), Blacks completed 27 percent of the assessments (compared with an expected yield of 30 percent), and others completed 36 percent of the assessments (compared with an expected yield of 40 percent). Respondents aged 60 and older completed 25 percent of the assessments (compared with an expected yield of 30 percent).

**Table 3-12. Actual sample yield from the household and volunteer samples, by age and race/ethnicity: 2001**

| Race/ethnicity | Total (percent) | Age (percent) <60 years | Age (percent) 60+ years |
|---|---|---|---|
| Total | 1,406 (100) | 1,053 (75) | 353 (25) |
| Hispanic | 523 (37) | 396 (28) | 127 (9) |
| Black | 373 (27) | 257 (18) | 116 (8) |
| Other | 510 (36) | 400 (28) | 110 (8) |

NOTE: All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as other include non-Hispanics of all other races, including multiracial. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2001 National Assessment of Adult Literacy field test.

### 3.3.8.5 Demographic Profiles of Respondents

Table 3-13 provides the actual sample yield from both the household and volunteer samples. From this table, it is possible to see a detailed distribution of the sample by age and race/ethnicity for the household and volunteer samples.

**Table 3-13. Actual sample yield from the household and volunteer samples, by race/ethnicity and age: 2001**

| Age (years) | Grand total (percent) | Household sample (percent) | | | | Volunteer sample (percent) | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Total[1] | Hispanic | Black | Other | Total | Hispanic | Black | Other |
| Total | 1,406 (100) | 967 (69) | 276 (20) | 189 (13) | 502 (36) | 439 (31) | 247 (18) | 184 (13) | 8 (1) |
| < 60 | 1,053 (75) | 819 (58) | 261 (19) | 162 (12) | 396 (28) | 234 (17) | 135 (10) | 95 (7) | 4 (0) |
| 60+ | 353 (25) | 148 (11) | 15 (1) | 27 (2) | 106 (8) | 204 (15) | 112 (8) | 89 (6) | 4 (0) |

[1] The total accounts for 967 of the 972 completed assessments for which race/ethnicity and age data were available.
NOTE: All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as other include non-Hispanics of all other races, including multiracial. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2001 National Assessment of Adult Literacy field test.

## 3.4 DATA PREPARATION AND PROCESSING

During the field period, interviewers returned materials to the data collection contractor in two formats: electronic and hard copy. The electronic data consisted of screener, background questionnaire, and interviewer guide data, as well as status code updates. The hard-copy materials returned to the data collection contractor included assessment booklets, household folders, and non-interview report forms (NIRFs).

The Data Management System was used to support data processing activities, among other functions. The Data Management System consisted of three main components: the Supervisor Management System (see section 3.3.2.2), receipt control functions, and the reporting mechanism.

### 3.4.1 Overview of Data Preparation and Processing Activities for the Screener, the Background Questionnaire, and the Interviewer Guide

Data preparation and processing systems and procedures supported the extraction of screener, background questionnaire, and interviewer guide data out of Blaise and into SAS data files. As cases were received from interviewers, they were appended to a cumulative database. The cases were prepared for a review of interviewer comments and back-coding of other-specify and open-ended items. Comment review and coding of open-ended items were performed by using tools that applied edits consistent with the interviewing instruments. Data were then converted into the study database. Frequencies were run daily on the study database and reviewed for outliers and inconsistencies within the data. After the frequency review, the data were loaded into the Blaise Editing System, batched, and tracked.

### 3.4.2 Receipt and Batching of Cases

The Blaise Editing System used the concept of data batches to keep the information in the batch reports manageable. A tracking file contained all the records within a batch. A batch number was used to link tracking file records to batch file records.

The Blaise management report identified batches for which some data cleaning was required. The batches were then coded, verified, adjudicated, and readied for delivery. Although the majority of edit errors were resolved during the interview owing to edits built into the automated instrumentation, postcollection editing was also required. This editing included the edits executed in the instrumentation, as well as edits considered too time-consuming or complex to execute in an interview situation.

### 3.4.3 Processing of Hard-Copy Assessment Materials

#### 3.4.3.1 Procedures

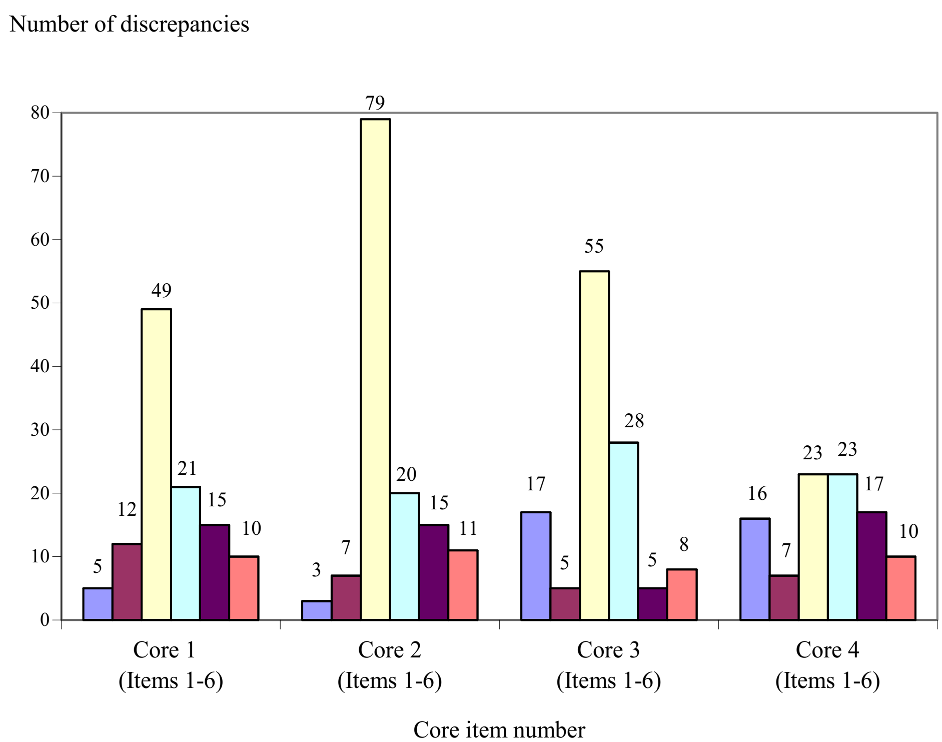
Interviewers were instructed to return completed assessment materials to the data collection contractor twice weekly. All shipments were tracked electronically through the Interviewer Management System and the Data Management System. Upon arrival, packages were checked into the Data Management System. Each item was compared with the interviewer's list of items included in the shipment and marked as received in the Data Management System. Any discrepancies were brought to the attention of the systems management team and, where needed, to the attention of the field director and supervisors.

#### 3.4.3.2 Core Assessment Validation

To evaluate interviewer accuracy in scoring the core assessment items, each core assessment item was rescored by a trained coder in the home office, and the validation score was compared with the score recorded by the interviewer. Of the 1,405 core assessments that were validated, approximately 360 had at least one scorer disagreement. Of the 360 cases, 76 had more than one scorer disagreement, for a total of 461 disagreements. See figure 3-1 for a distribution of these scoring disagreements across core items.

Item #3 in cores 1, 2, and 3 had the highest level of scoring disagreement between the field interviewer and the home office coder. Core 1, item #3 and core 3, item #3 were identical items; core 2, item #3 was a distinct item. Both of these items were prose tasks in which the sample persons were asked to provide an open-ended response based on reading a passage. It was determined during the field period that the answers to these core tasks were difficult to score consistently because of some ambiguity in the scoring rubric. It was unclear whether the sample person had to produce a response identical to that in the scoring rubric or if some leniency was permitted. Based on the levels of coder disagreement experienced in the field test, the scoring rubrics for the core items used in the main data collection effort were revised and improved to decrease the ambiguity and add some leniency in the application of the rubric. Additionally, the interviewer and home office training program was enhanced to give trainees more opportunities to practice coding the tasks, based on actual answers collected in the field test.

**Figure 3-1. Distribution of discrepancies in core assessment scoring results, by core item number: 2001**




SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2001 National Assessment of Adult Literacy field test.

### 3.4.4 Delivery of Assessment Booklets for Scoring

The assessment booklets were delivered to the assessment booklet printing and scoring subcontractor for scoring at four points during the field period.

### 3.4.5 Quality Control of Data in the Study Database

A detailed reporting system, integrated with the Data Management System, was used throughout the field test. Data in these reports were reviewed by the field supervisors, the field director, and the home office staff throughout the field period.

Receipt control reports were used to track and verify that assessment booklets were being sent to the data collection contractor's home office on a timely basis after completion. Additionally, the integrated receipt control system tracked the assessment booklets throughout the field period and documented the location of each booklet.

At the end of the field period, an extensive data validation process was completed to ensure that Blaise interview data existed for each completed case in the Data Management System.

In addition, the assessment scoring data were reviewed and the scores received from the scoring subcontractor were compared with the scoring data to ensure that a score was received for each case ID and NCS booklet barcode number that were originally sent.

*This page is intentionally left blank.*

# CHAPTER 4

# FIELD-TEST SCORING AND ANALYSIS OF FIELD-TEST DATA

Elizabeth Greenberg, American Institutes for Research

Following the conclusion of the field test, the field-tested cognitive items were scored and the results were analyzed to determine which items to retain for the operational assessment. The background questionnaire (BQ) data obtained during the field test were analyzed, and changes were made to the BQ on the basis of the field-test data. The field-test results were also used to select the core items for the operational assessment and to develop the algorithm for selecting Adult Literacy Supplemental Assessment (ALSA) respondents.

## 4.1 DEVELOPMENT OF SCORING PROCEDURES

This section discusses the refinement of the scoring rubrics on the basis of range-finding conducted using the field-test booklets, the development of scorer training materials, and the field-test scoring procedures.

### 4.1.1 Refinement of Scoring Rubrics

Draft scoring rubrics for the items in the National Assessment of Adult Literacy (NAAL) field test were developed by the item writers. Prior to the scoring of the NAAL field test, project staff conducted range-finding—reviewing a sample of assessment booklets to determine the range of responses received for a particular item—to identify the different types of responses received for the individual NAAL items. Approximately 100 responses to each item were reviewed as part of the range-finding process, and the range-finding staff photocopied all distinct responses to each item to use during scorer training. The scoring rubrics were refined and expanded to address alternative correct or partially correct responses identified during range-finding. For example, one assessment item asked respondents to contrast two different ideas. Although scorers will always need to exercise some judgment when evaluating the responses to this type of question, range-finding identified the most common responses. These responses were then summarized on the scoring rubrics, with the proper score assigned to each response.

### 4.1.2 Development of Scorer Training Materials

The responses collected during range-finding were used as the basis for the scorer training materials. Responses that closely matched each score point on the scoring rubric were identified during range-finding for use as anchor papers to illustrate each score point. Both straightforward and ambiguous responses for each score point were identified as training papers for the scorers to score individually and then discuss as a group. The number of responses used as training papers varied by item. For straightforward items, which required a number or short phrase as a response, five or six training papers were identified. For items for which the responses were longer and more complex to score, up to 20 training papers were identified.

Project staff developed a detailed script for each item for the scorer trainers to follow. The script for each item began by asking the scorers to read the item and formulate an answer. Next, the script walked the scorers through the rubric, explaining what a respondent had to do to receive credit for an item. Where applicable, the script instructed the trainers to remind the scorers to be flexible when interpreting responses and to allow variations in wording that still expressed the meaning of the original score point.

Next, the script directed the trainers to ask whether the scorers had any questions and then to walk the scorers through the anchor papers. The script included a table that listed each anchor paper and explained why it was chosen. After going through the anchor papers, the scorers were directed by the trainers to individually score the training papers for a particular item. Another table in the script listed the correct score point for each training paper and an explanation of why that score point was assigned. The script directed the trainers to discuss each training paper with the scorers and to not move on until all scorers indicated that they understood why each paper was assigned a particular score.

### 4.1.3 Field-Test Scoring Procedures

Field-test scoring was conducted in Iowa City, Iowa. Because the NAAL field-test scoring took place during January, which is normally a slow time for testing organizations, NAAL was able to use only scorers who had significant experience scoring open-ended language arts items. Twenty scorers were hired for the NAAL field test. They were divided into two groups, called tables, with nine scorers and a table leader at each table.

Item development staff traveled to Iowa to train the scorers. Each table was trained on one block at a time and completed the scoring of that block before moving on to score the next block. Because this was a field test and the first time these scoring rubrics and scorer training materials were used, one purpose of the scoring was to further refine both the rubrics and the training materials. Item development staff discussed the rubrics and the training materials with the scoring staff as they were scoring and noted the training papers for which more detailed explanations of the assigned scores would be helpful. They also noted the items for which additional training papers or more detailed rubrics were needed to address ambiguities. In a few cases, rubrics were edited during the field-test scoring process.

## 4.2 ANALYSIS OF FIELD-TEST DATA

This section describes the analysis of the NAAL field-test data, including background questionnaire responses, field-test cognitive data, the Fluency Addition to the NAAL (FAN), and the Adult Literacy Supplemental Data (ALSA).

### 4.2.1 Background Questionnaire Analyses

Project staff analyzed the NAAL field-test BQ data to identify

- questions that respondents refused to answer;
- questions for which the relationship to the respondent's literacy level was different than expected;
- questions with unexpected response patterns (e.g., large numbers of respondents picked a response that goes against known research);
- redundant questions (i.e., questions measuring the same construct with high response correlations);
- problems with skip patterns;
- places where look-up tables could replace open-ended questions to make the analysis easier; and
- any difficulties with the administration of the BQ in the field.

All field interviewers were asked to complete a questionnaire identifying any problems they encountered when administering the BQ in the field, and project staff then reviewed their responses. Additionally, approximately half a dozen field interviewers were asked to participate in a debriefing session to discuss any problems they encountered when administering the BQ in the field. Field staff all reported that the BQ was easy or very easy to administer.

Project staff computed frequencies for every question on the BQ and mean booklet scores for all response categories for each background question. Project staff also ran factor analyses for sets of questions that were designed to measure different aspects of the same underlying construct.

Almost all the questions on the BQ functioned as expected. As a result of the factor analyses, three questions that overlapped with other questions and were not necessary to measure an underlying construct were dropped. A question asking respondents to identify the city and county they lived in when they graduated high school was dropped because the results could not be matched to existing databases by using Federal Information Processing Standards (FIPS) codes or ZIP codes; however, the question asking respondents which state they lived in when they graduated high school was retained. Skips were added in a few places where questions were not relevant to some of the respondents, and one problem with the computer-assisted personal interviewing (CAPI) programming for a skip pattern was identified and corrected. Reference time periods were changed for a couple of questions for which interviewers indicated that respondents were confused or unable to answer a question.

Administration time for the field-test BQ averaged 26 minutes, which was 9 minutes less than the time allocated. This provided some flexibility, and five questions were added to the BQ, including one asking respondents when they completed their last year of college and one asking about veteran status.

### 4.2.2 Analysis of Field-Test Cognitive Data

The field test consisted of 12 blocks of items,[1] with a total of 141 noncore items, plus four versions of the core, with a total of 17 core items (see section 2.1 for a discussion of core and noncore items). The noncore blocks were assembled in a partial spiral so that each block appeared in each position in a test booklet, but not every block appeared with every other block. Table 4-1 shows the field-test booklet spiral.

[1]The 12 non-core blocks were numbered 1.1, 1.2, 1.3, 2.1, 2.2, 2.3, 3.1, 3.2, 3.3, 4.1, 4.2, and 4.3.

**Table 4-1. Field-test booklet spiral: 2003**

| Booklet # | Position 1 | Position 2 | Position 3 | Position 4 |
|---|---|---|---|---|
| 1 | Core 1 | 1.1 | 1.2 | 1.3 |
| 2 | Core 2 | 2.1 | 2.2 | 2.3 |
| 3 | Core 3 | 3.1 | 3.2 | 3.3 |
| 4 | Core 4 | 4.1 | 4.2 | 4.3 |
| 5 | Core 1 | 1.2 | 1.3 | 1.1 |
| 6 | Core 2 | 2.2 | 2.3 | 2.1 |
| 7 | Core 3 | 3.2 | 3.3 | 3.1 |
| 8 | Core 4 | 4.2 | 4.3 | 4.1 |
| 9 | Core 1 | 1.3 | 1.1 | 1.2 |
| 10 | Core 2 | 2.3 | 2.1 | 2.2 |
| 11 | Core 3 | 3.3 | 3.1 | 3.2 |
| 12 | Core 4 | 4.3 | 4.1 | 4.2 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

Project staff analyzed the NAAL field-test cognitive data to identify

- items with differential item functioning (DIF) in regard to female versus male, Hispanic versus other, Black versus other, or age 60+ versus under age 60;
- items with interrater reliability problems;
- items that did not discriminate well among respondents with different levels of literacy;
- noncore items with field-test p-values (percentage of respondents who answered each item correctly) outside the range of the 1992 noncore items that were being replaced (below .20 or above .90);
- items appropriate for partial credit; and
- a set of core items that could be used to screen respondents for inclusion in ALSA or the main NAAL.

Because the field test was designed to oversample Blacks and Hispanics, weights were created prior to the field-test data analysis so that the percentages of Blacks, Hispanics, and people age 60 or older corresponded to the percentages in the 2000 Census.

#### 4.2.2.1 Item Analyses

As shown in table 4-1, booklets 1, 5, and 9 included the same blocks and items. Similarly, booklets 2, 6, and 10 included the same blocks; booklets 3, 7, and 11 included the same blocks; and booklets 4, 8, and 12 included the same blocks. Item analysis was done separately for each of the four sets of field-test booklets, using weighted data. Because of the nature of the field-test sample, which combined respondents selected through a probability sample with a convenience sample of respondents selected through focus group centers (see chapter 3), the field-test data could not be weighted to accurately reflect population estimates. However, the data were weighted to roughly reflect the racial/ethnic breakdown of the population (Black, Hispanic, other) and the age breakdown of the population based on the 2000 Census.

Project staff computed the following statistics for each item:

- percentage of respondents who answered the item correctly (p-value);
- mean booklet score for each score point associated with an item;
- differential item functioning (DIF; female vs. male, Hispanic vs. other, Black vs. other, or age 60+ vs. under age 60);
- biserial correlation; and
- interrater reliability.

Each of these statistics is discussed below.

***Percentage of Respondents Who Answered Each Item Correctly (p-value)***

The size of the field test necessitated obtaining data from each respondent for all the items with which the respondent was presented. Therefore, field-test blocks were purposely kept short so that respondents could easily complete all the items in a block. Because of this decision, project staff did not distinguish between "not reached" and "missing/skipped" when analyzing the field-test data and treated all blank responses as "incorrect" when calculating *p*-values for the field-test items. Noncore items with field-test *p*-values outside the range of the 1992 items that were being replaced (below .20 or above .90) were flagged. Of the 141 noncore items in the field test, 3 had *p*-values below .20 and 3 had *p*-values above .90 (table 4-2).

**Table 4-2. Number of NAAL noncore and core field-test items, by percent correct (p-value): 2003**

| *p*-value | Number of noncore items | Number of core items |
|---|---|---|
| Below .20 | 3 | 0 |
| .20 to .90 | 135 | 8 |
| Above .90 | 3 | 9 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### *Mean Booklet Score*

The mean booklet score for a score point is the average number of items that respondents choosing that score point answer correctly in the remainder of the booklet in which the item of interest appears. Generally, respondents who receive a score of correct on an item should have a higher mean booklet score than respondents who do not answer the item correctly. If they do not, this is an indication that the item is not functioning as expected. Mean booklet scores for each score point—including "any other response," blank, off task, illegible, and "I don't know"—were computed for each item. No items had incorrect responses that had mean booklet scores equal to or higher than the correct response.

### *Differential Item Functioning*

Differential item functioning (DIF) analysis refers to procedures that assess whether specific items are differentially difficult for different groups of respondents after controlling for overall differences among groups. DIF procedures compare the performance of groups on each item within sets of respondents who have the same level of performance, usually measured by total test score. Items identified as having DIF are then evaluated to determine whether they are biased; that is, whether the DIF is related to a factor unrelated to what is being tested. A biased item is generally deleted from a test because the probability of doing well on the item depends in part on characteristics of the item that are not related to the construct being measured.[2]

Items were classified as A, B, or C in regard to DIF:

- A means the item has negligible DIF.
- B means the item has moderate DIF.

[2] This paragraph was adapted from the forthcoming National Assessment of Educational Progress (NAEP) Technical Report.

- C means the item has significant DIF.

DIF classifications were based on delta ( $\hat{\Delta}_{MH}$ ) computed as follows:

$$\hat{\Delta}_{MH} = -\frac{4}{1.7}\ln(\hat{\alpha}_{MH}) = -2.35\ln(\hat{\alpha}_{MH})$$

where $\hat{\alpha}_{MH}$ is the Mantel-Haenszel (MH)-odds ratio.[3] Statistical tests of the null hypothesis were performed at the .05 level using the Chi-square test statistics MH (chi-square), which is distributed as a chi-square with 1 degree of freedom (table 4-3).

**Table 4-3. Classification rules for differential item functioning (DIF) categories: 2003**

| Category | Classification rule |
|---|---|
| C | $\lvert\hat{\Delta}_{MH}\rvert$ is significantly greater than 1.0, and $\lvert\hat{\Delta}_{MH}\rvert > 1.5$. |
| B | $\lvert\hat{\Delta}_{MH}\rvert$ is significantly different from zero and > = 1.0, and either<br>a) $\lvert\hat{\Delta}_{MH}\rvert \leq 1.5$ or<br>b) $\lvert\hat{\Delta}_{MH}\rvert$ is not significantly greater than 1.0 |
| A | $\lvert\hat{\Delta}_{MH}\rvert$ is not significantly different from zero using $MH\chi^2$ , or $\lvert\hat{\Delta}_{MH}\rvert < 1.0$ |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

DIF was classified either + or – depending on whether the items favored the target group or the reference group. Items that favored the target group were classified +, and items that favored the reference group were classified –. DIF analysis was conducted for the following target groups: females versus males, Hispanics versus others (others included all adults who were not classified as Hispanic or Black), Blacks versus others (others included all adults who were not classified as Hispanic or Black), and adults age 60 or over versus under age 60. Items with B or C DIF were flagged for further examination. Table 4-4 shows the number of items that exhibited DIF in the field test.

[3] For a discussion of delta, see Holland, P.W., and Thayer, D.T. (1988). *Differential Item Performance and Mantel-Haenszel* in Wainer, H. & Braun, H. (Eds.). *Test Validity.* Hillsdale, NJ: Lawrence Erlbaum Associates.

**Table 4-4. Number of NAAL noncore and core field-test items, by differential item functioning (DIF) categories: 2003**

| Differential item functioning | Number of noncore items | Number of core items |
|---|---|---|
| Female vs. male: A | 128 | 17 |
| Female vs. male: B+ | 6 | 0 |
| Female vs. male: B– | 5 | 0 |
| Female vs. male: C+ | 0 | 0 |
| Female vs. male: C– | 2 | 0 |
| | | |
| Hispanic vs. other: A | 131 | 14 |
| Hispanic vs. other: B+ | 2 | 2 |
| Hispanic vs. other: B– | 7 | 1 |
| Hispanic vs. other: C+ | 0 | 0 |
| Hispanic vs. other: C– | 1 | 0 |
| | | |
| Black vs. other: A | 132 | 17 |
| Black vs. other: B+ | 5 | 0 |
| Black vs. other: B– | 4 | 0 |
| Black vs. other: C+ | 0 | 0 |
| Black vs. other: C– | 0 | 0 |
| | | |
| 60+ vs. under 60: A | 123 | 13 |
| 60+ vs. under 60: B+ | 11 | 2 |
| 60+ vs. under 60: B– | 4 | 2 |
| 60+ vs. under 60: C+ | 3 | 0 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### *Biserial Correlation*

The biserial correlation is an item discrimination index that is based on the weighted correlation between the item score and the booklet score, where the booklet score excludes the item being studied. The biserial correlation indicates the extent to which each item differentiates among those respondents who possess the skills being measured and those who do not. The higher the value of the biserial correlation, the more discriminating an item is. In general, a biserial correlation below .2 indicates an item problem. None of the field-test items was flagged because of issues related to biserial correlation. Among the field-test items, 136 had biserial correlations of .60 or higher, 21 had biserial correlations of .40 to .59, and 1 had a biserial correlation of .20 to .39 (table 4-5).

**Table 4-5. Number of NAAL noncore and core field-test items, by biserial correlation: 2003**

| Biserial correlation | Number of noncore items | Number of core items |
|---|---|---|
| Above .60 | 124 | 12 |
| .40 to .59 | 17 | 4 |
| .20 to .39 | 0 | 1 |
| Below .20 | 0 | 0 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

***Interrater Reliability***

Half the field-test booklets were scored by a second scorer, and interrater reliability statistics, comparing the score assigned by the first scorer with the score assigned by the second scorer, were computed for each field-test item. In addition to computing exact agreement between the first and second scorers, project staff computed the percentage of cases for which one scorer assigned full credit and the second scorer assigned partial credit, or one scorer assigned partial credit and the second scorer assigned no credit. Items with exact interrater reliability below .95 were flagged for further examination. For the field test, 140 items had interrater reliability of 95 percent or higher, 10 items had interrater reliability of 90 to 94.9 percent, 6 items had interrater reliability of 85 to 89.9 percent, 1 item had interrater reliability of 80 to 84.9 percent, and 1 item had interrater reliability below 80 percent (table 4-6).

**Table 4-6. Number of NAAL noncore and core field-test items, by interrater reliability (exact agreement): 2003**

| Interrater reliability | Number of noncore items | Number of core items |
|---|---|---|
| 95 or higher | 125 | 15 |
| 90 to 94.9 | 8 | 2 |
| 85 to 89.9 | 6 | 0 |
| 80 to 84.9 | 1 | 0 |
| Below 80 | 1 | 0 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 4.2.2.2 Flagged Items

Items that were flagged during the item analysis of the field-test data were examined by senior project staff. All noncore items with *p*-values outside the range of .20 to .90 and all noncore items with biserial correlations below .60 were eliminated from the item pool. Core items with *p*-values above .90 were not considered to be a problem because the core items were deliberately kept easy to ease participants into the main assessment blocks. Two core items with biserial correlations between .40 and .59 were also retained for the operational assessment.

Senior project staff also examined all items with scoring interrater reliability below 95 percent. The single item with interrater reliability below 80 percent was the only item on the field test that asked the respondent to give an oral response that was recorded for scoring later. That item was eliminated because of the difficulty in scoring the oral response. The other items with interrater reliability below 95 percent were among the more-difficult-to-score items, but it seemed likely that interrater reliability could be improved for these items with additional scorer training. These items were not automatically eliminated from the item pool, but—as discussed in chapter 13—additional range-finding was conducted for these items and additional scorer training materials were developed for the seven items with interrater reliability below .95 that were retained for the operational assessment.

All items with C DIF were eliminated from the pool for assembling the operational blocks, with the exception of one item that had C+ DIF in favor of respondents age 60 or over. This item was kept because the stimulus material and question were determined to be important to the assessment. The review panel that evaluated the operational blocks (see chapter 2) was asked to pay particular attention to this item, and the panel endorsed the decision to keep this item in the assessment.

B DIF was considered less of a problem than C DIF, and items with B DIF were not automatically eliminated from the item pool. In general, these items were balanced in terms of whether they favored the target group or the reference group. Among the items in the seven new blocks assembled from the field-test items, 3 items exhibited female versus male B DIF (1 in favor of females, 2 in favor of males), 3 items exhibited Hispanic versus others B DIF (all in favor of others), 3 items exhibited Black versus others B DIF (1 in favor of Blacks, 2 in favor of others), and 11 items exhibited 60+ versus under 60 DIF (8 in favor of 60+, 3 in favor of under 60). No items retained for the core exhibited any DIF.

#### 4.2.2.3 Partial Credit

For some of the NAAL field-test items, data were collected for potential partial-credit score points. For these items, the NAAL field-test data analysis was used to make preliminary decisions about whether the items supported awarding partial credit. These decisions were based on the weighted mean field-test booklet scores associated with the partial-credit score points. If the mean field-test booklet score associated with a partial-credit score point was between the mean booklet score associated with a correct response and a mean booklet score associated with an incorrect response, the partial-credit score point was retained for the operational scoring rubrics. For items for which the mean booklet score for the partial-credit score point did not fall between the mean booklet score for full credit and the mean booklet score for no credit, the partial-credit score point was dropped from the scoring rubrics for the operational assessment.

#### 4.2.2.4 Algorithm Used to Identify Adult Literacy Supplemental Assessment Respondents

One goal of the field-test data analysis was to select a set of questions that could accurately identify the bottom 10 percent of respondents as appropriate for the ALSA. The bottom 10 percent of respondents were targeted because on the basis of 1992 data, it was thought that they would be unlikely to complete enough NAAL items to provide meaningful data.[4] The analysis plan called for starting with one of the four versions of the core that was field-tested in fall 2001 and supplementing it as necessary with the fewest number of additional items that would accurately identify ALSA respondents. The additional items were chosen because of their ease of scoring (they needed to be accurately scored by field interviewers) and performance in the field test (items with DIF and items that did not discriminate well among respondents with different levels of literacy were eliminated from consideration). Because of DIF problems, the version of the core that appeared in booklets 1 and 3 was eliminated from consideration as the basis of the core for the operational assessment.

On the basis of the analysis of the field-test data, the core that appeared in field-test booklet 2, supplemented with question 9 from block 1 of that booklet, was identified as having the best combination of items to identify respondents who would likely fall in the bottom 10 percent (unweighted) of the population because of literacy—the screening criteria for the ALSA as discussed above. Question 9 in block 1 was a health question based on the *Medicare and You* booklet. The question directed respondents to a particular page in the booklet and asked them to locate information about how often people should get a flu shot.

In the NAAL field test, 350 respondents answered booklets 2, 6, or 10 (as shown in table 4-1, these booklets included the same set of cognitive items, with the three blocks presented in a different location in each book). A rank ordering of these respondents from lowest to highest score showed that the 35th respondent correctly answered 4.5 questions in the entire test booklet, including the core and

[4]The bottom 10 percent of respondents were identified by using unweighted data. Because of the nature of the National Assessment of Adult Literacy field-test sample, which combined respondents selected through a probability sample with a convenience sample of respondents selected through focus group centers, the field-test data cannot be weighted to accurately reflect population estimates. However, the field-test data were weighted to roughly reflect the racial/ethnic breakdown of the population (Black, Hispanic, other) and the age breakdown of the population (60 and over vs. under 60). After weighting by these characteristics (race/ethnicity and age), the bottom 10 percent of the field-test sample corresponded to approximately the bottom 5 to 7 percent of the population.

noncore blocks (the fraction comes from partial-credit items) and the 36th respondent correctly answered five questions in the entire test booklet. Thus, to have approximately 10 percent of respondents participate in ALSA, the goal was to identify items that would accurately predict which respondents would correctly answer 5 or fewer of the questions in the core plus the three additional blocks in books 2, 6, or 10.

The following algorithms were developed to discriminate between ALSA and main NAAL respondents by identifying those who would correctly answer five or fewer questions in books 2, 6, or 10:

Algorithm for selecting ALSA respondents for respondents who took the core in English:

- score of 0 on the original 6 core items; OR
- score of 0 on the additional core question (question 9 in block 1 of field-test booklet 2—the flu shot) AND did not attempt to answer core questions 3 and 4.

Algorithm for selecting ALSA respondents for respondents who took the core in Spanish:

- score of less than 5 on the Spanish version of the core; AND
- score of 0 on the additional core question (question 9 in block 1 of field-test booklet 2—the flu shot).

Table 4-7 shows the number of respondents who would be correctly and incorrectly classified into the ALSA and the main NAAL on the basis of this algorithm and the field-test data. Correct classification is defined as placement into the ALSA if the total number of items answered correctly in the entire book (the core plus the three additional blocks) was 5 or fewer and placement into the main NAAL if the total number of items answered correctly in the entire book (the core plus the three additional blocks) was more than 5.

**Table 4-7. Number of field-test respondents correctly and incorrectly classified into the main NAAL and the Adult Literacy Supplemental Assessment (ALSA) on the basis of the algorithm for selecting ALSA respondents: 2003**

| Responding category | NAAL | ALSA |
|---|---|---|
| Correctly classified respondents | 311 | 32 |
| Incorrectly classified respondents | 4 | 3 |
| Total respondents | 315 | 35 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

On the basis of the recommended criteria, the four respondents who would have been incorrectly screened into the NAAL had total scores of 2, 2.5, 4, and 4 for the cognitive items in the three blocks. The three respondents (0.9 percent of the total sample) misclassified into the ALSA had total scores of 5.5, 5.5, and 13.5. The effect of classifying the two respondents with scores of 5.5 into the ALSA was small because their scores were very close to the cut-point score. According to these criteria for screening respondents into the ALSA, only one of the field-test respondents (the person with a score of 13.5 on the entire booklet) would clearly be misclassified.

### 4.2.3 Analysis of Fluency Addition to the NAAL Field-Test Data

The Fluency Addition to the NAAL (FAN) field test is described in chapter 5. This section describes the analysis of the FAN field-test data.

Sixteen connected text passages were field-tested for the FAN (eight easy and eight difficult passages) with the goal of selecting eight passages for the operational assessment (four at the easier level and four at the more difficult level). The easier passages included both narratives and expository texts; the more difficult passages were all expository texts.

Table 4-8 shows the results from the FAN field test. Although 16 passages were field-tested, two passages were eliminated because copyright permission to use the passages in the main assessment could not be obtained. A third passage was eliminated because field-test respondents had a lot of difficulty responding to the comprehension question associated with the passage, indicating that the passage was more difficult than intended. Passages in table 4-8 are organized by respondents' average words per minute; easier texts are listed first, followed by more difficult texts. All selected texts were classified as either grade 3 or grade 8 on the basis of Lexile. The column labeled "status" indicates whether the text was kept for the operational assessment or dropped after the field test. The column labeled "explanation" gives the reason for the decision to keep a text. The decision with regard to which passages to include in the operational assessment was based on the following criteria:

- diversity among the passages in terms of structure (narrative vs. expository) and
- diversity among the passages in terms of reading speed.

**Table 4-8. Fluency Addition to the NAAL (FAN) passages, by field-test results: 2003**

| Title | Structure | Lexile | Words per minute | Status | Explanation |
|---|---|---|---|---|---|
| Curly | Narrative | 540 (gr. 3) | 186.4 | Keep | Fastest narrative, grade 3 |
| Walter | Narrative | 590 (gr. 3) | 178.1 | Drop | Dropped |
| Rainbow | Narrative | 640 (gr. 3) | 174.9 | Drop | Dropped |
| Guide Dogs | Expository | 700 (gr. 3) | 169.4 | Keep | Fastest expository text, grade 3 |
| Solar Eclipse | Expository | 750 (gr. 4) | 167.1 | Drop | Dropped |
| Grand Canyon | Expository | 570 (gr. 3) | 159.5 | Keep | Slow expository text, grade 3 |
| Physical Activity | Expository | 650 (gr. 3) | 162.5 | Drop | Very similar to exercise at grade 8; do not want to spiral two texts with similar topics |
| Amanda and I | Narrative | 700 (gr. 3) | 162.1 | Keep | Slowest narrative, grade 3 |
| Exercise | Expository (health) | 1020 (gr. 8) | 164.4 | Keep | Fastest passage, grade 8 |
| Lori Goldberg | Expository | 1030 (gr. 8) | 159.5 | Keep | Second fastest passage, grade 8 |
| Elk | Expository | 1020 (gr. 8) | 147.8 | Drop | Dropped |
| Bigfoot | Expository | 1020 (gr. 8) | 143.5 | Keep | Second slowest passage, grade 8 |
| Chicken Soup | Expository (health) | 1100 (gr. 8/9) | 136.6 | Keep | Slowest passage, grade 8 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

*This page is intentionally left blank.*

# CHAPTER 5

# FIELD TEST OF THE FLUENCY ADDITION TO NAAL

Michelle Amsbary, Westat

## 5.1 INTRODUCTION

In November 2001, a panel of experts recommended that the government provide, for the first time, a clearer picture of the basic reading skills of low-performing adults by examining their oral reading fluency. In response to this recommendation, an oral reading component for the National Assessment of Adult Literacy (NAAL): the Fluency Addition to the NAAL (FAN) was designed. The FAN field test was conducted to test this newly created oral module assessment and the associated procedures.

In preparation for the field test, the FAN software was developed and tested. The interviewing and data delivery systems and procedures developed for the 2001 NAAL field test were revised to incorporate the newly developed oral module instrument. Instructional manuals and training programs for supervisors and interviewers were developed.

The FAN field test was administered exclusively to volunteers recruited and screened by a focus group facility. The sample design is described in section 5.2. The interviews were conducted by trained interviewers in the respondents' homes.

Interviewers attended a 2-day in-person training that focused on the administration of the oral module but also covered the administration of the screener, the background questionnaire, and the assessment, as well as general contact and administrative procedures. During data collection, two regional supervisors, a field manager, and additional data collection contractor home office staff frequently monitored production. Section 5.3 discusses the data collection instruments and materials, field staff recruitment and training, and the data collection effort for the FAN field test.

Interviewing for the FAN field test lasted 5 weeks, beginning in late September 2002. Interviews were administered in 10 locations across the country by 21 skilled interviewers. A total of 520 interviews were completed.

As described in section 5.4, the data preparation and processing systems and procedures used for the 2001 NAAL field test were refined for the FAN field test. During data collection, data were reviewed for interviewer comments and "other-specify" responses and then converted into the study database. After frequency review, the data were loaded into the Blaise Editing System, batched, and tracked through delivery to the analysis contractor. The hard-copy assessment booklets were tracked through the Data Management System and sent to the analysis contractor as well. Electronic scoring data from the FAN voice recording files were received from the scoring contractor, reviewed, and delivered for analysis.

## 5.2 SAMPLE DESIGN

For this purposive field test sample, focus group facilities in the 10 selected locations were contracted to recruit 500 respondents according to specific sampling requirements. These sampling requirements were based on educational attainment, race/ethnicity, and native/nonnative English-speaking ability, as well as a balance of age and gender.

In selecting the sample, sites with high percentages of the target minority populations were chosen, particularly sites where high concentrations of Asian respondents were expected. Minority (especially Asian) populations were targeted to test the FAN instrument with diverse dialects and accents. The focus group services made extraordinary efforts to recruit Asian respondents, including advertising for participants through radio spots and in-person recruiting in public locations, such as shopping malls. Despite these efforts, the Asian respondents were the most difficult to recruit. Toward the end of the field period, the sample requirements were relaxed, allowing the recruitment of additional Asian respondents with a high school diploma in place of those without a high school diploma.

## 5.3 DATA COLLECTION

### 5.3.1 Data Collection Instruments

The FAN field test used the following computer-assisted personal interviewing (CAPI) instruments: an abbreviated screener, the background questionnaire, an interviewer guide, and the oral module. These CAPI components were integrated into the Interviewer Management System. A hard-copy assessment booklet containing the core and main assessments was included in the field test as well. These instruments are described below. All instruments were available in both English and Spanish, except for the main exercise, which was available only in English.

The oral module consisted of numerous passages and lists of numbers, letters, and words to be read aloud by the respondent and recorded directly onto the laptop computer. (See section 2.5 for a full discussion of the development and administration of the oral module assessment.) Prior to the field test, a small feasibility test was conducted with volunteers simply to test the stand-alone oral module. Disks of the recordings created as part of this feasibility test were reviewed.

The version of the background questionnaire used for the FAN field test was the same as that used in the 2001 NAAL field test. Unlike the field test, which used 12 versions of the assessment booklet, the FAN field test used only assessment booklet 2. Because the primary focus of the FAN field test was the oral module, and the items in the assessment booklets had been tested through the 2001 field test, only one type of booklet was used. Additionally, some modifications and enhancements were made to the overall Interviewer Management System to incorporate the new fluency task.

Changes to the Supervisor Management System were implemented to include details about the oral module task. Enhancements were made to the interviewer shipping module (to accommodate the shipping of zip disks), the receipt control system, the browse case functionality, and all study management reports. In addition, the interviewer data transmission process was enhanced to permit the transmission of the FAN administration file. This file collected timing and comment data captured by the oral module.

### 5.3.2 Interviewer Materials

The materials used in the FAN field test were a subset of those used in the 2001 NAAL field test. The respondents were recruited and screened through a focus group facility, so no advance letter or brochure was required. Additionally, noninterview report forms, segment folders, U.S. Department of Education letters of introduction, community authorization letters, nonresponse letters, and sorry-I-missed-you cards were not used. Handcards were used only for the background questionnaire, not the screener.

The oral module component required numerous materials, including the oral module booklet, the oral module interviewer guide, an interviewer headset, a respondent headset with a microphone, and a zip drive. Four versions of the oral module were tested during the FAN field test. Each version contained the same basic tasks, but the content of the items was slightly different or was presented in a different order. The interviewers' headsets enabled them to hear the signal that the time allotted had expired. (See section 5.3.4.4 for a more detailed discussion of the oral module interviewer guide and interviewer headsets.) The respondent headset with microphone allowed samples of the respondent's reading to be recorded and stored directly on the interviewer laptop. The zip drive was used to create zip disks of the voice recording files. Additionally, procedures were developed for a paper-based backup to the automated scoring produced from the oral module, in case of problems with the recordings. As part of this procedure, the interviewer used an oral module interviewer guide.

### 5.3.3 Field Staff Recruitment and Training

Field staff were recruited from among supervisors and interviewers who had experience with the NAAL instrument or a similar literacy instrument. A 2-day field supervisor training was held, followed by a 2-day interviewer training. These training sessions are detailed in section 5.3.3.2.

#### 5.3.3.1 Field Staff Recruitment

Two experienced supervisors were hired, including one who had supervised interviewers on the 2001 NAAL field test. Twenty-one interviewers were recruited, including several bilingual (Spanish and English) speakers. All interviewers had been employed on either the 2001 field test or a similar literacy-related study.

#### 5.3.3.2 Field Staff Training

A 2-day supervisor training session was held in mid-September 2002, followed immediately by a 2-day interviewer training session.

New training materials were created for the oral module component. The training materials created for the 2001 NAAL field test were revised to reflect changes resulting from the integration of the oral module, such as the use of the oral module CAPI application and associated booklets, use of the headsets and microphones, creation of zip disk recordings using the zip drive, and the revised contact protocol. A list of procedural and instrumentation changes between the 2001 field test and the FAN field test was developed and distributed to the field staff as well.

The two supervisors were trained on the Supervisor Management System, administrative procedures, use of field production reports, and the CAPI instrumentation. Twenty-one interviewers received extensive training on the administration of the screener, background questionnaire, core and main exercise, and new oral module component.

In addition to interactive and role-play exercises conducted with fellow trainees, interviewers had the opportunity to conduct an interview with a respondent recruited by a local focus group service and paid for his or her participation. This gave the interviewers exposure to a real interviewing environment before they conducted their first actual case assignment. Software was implemented to permit a quality control review of the audiorecordings captured during the live respondent interviews. All interviews were reviewed and deemed acceptable and usable for analysis, although they were not included in the field test analysis.

### 5.3.4 Data Collection Effort

The FAN field test data collection was conducted over 5 weeks in September and October 2002. The field staff structure, progress monitoring, contact protocol, and data collection procedures are presented in the following sections.

#### 5.3.4.1 Field Staff and Organization

Twenty of the 21 interviewers trained completed the FAN field test. The two supervisors guided and supported the field interviewers working in their region. The supervisors worked closely with interviewers to assign volunteer respondents and monitored progress in scheduling appointments and completing interviews.

The supervisors also assisted the home office study manager in monitoring the focus group service samples and in identifying volunteers requiring replacement. The field director oversaw the work of the supervisors.

#### 5.3.4.2 Monitoring of Data Collection Progress

Throughout the field period, the FAN sample yield was closely monitored in terms of the number of volunteers recruited and the number of completed background questionnaires. The actual number of completed interviews was compared weekly with the sample requirements for each focus group site.

#### 5.3.4.3 Contact Protocol

Interviews were conducted in respondents' homes to test the feasibility of administering the FAN in that setting. Interviewers were responsible for contacting the respondent, setting an appointment, getting directions to and locating the respondent's house, making a reminder call, and conducting the interview. Interviews were scheduled at least 3 hours apart, and interviewers were instructed to plan to complete at least two interviews a day.

The focus group facility provided the supervisors with information about the recruited respondents. The supervisors then gave the interviewers a spreadsheet containing information for the respondents in their area, including name, telephone number, address, language preference (English or Spanish), and the best time and day to call.

Using the script on the household folder, the interviewers introduced themselves, set an appointment date and time, and obtained directions to the volunteer's home. Before the appointment, the interviewer made a confirmation call as well. All call attempts were recorded on the Record of Actions on the back of the household folder, as well as in the Electronic Record of Calls in the Interviewer Management System.

#### 5.3.4.4 Data Collection Procedures

The instruments were administered in the following order: screener, background questionnaire, core and main exercises, and oral module. All instruments were administered in the same visit. Many of the procedures and study materials were similar to or unchanged from those used during the 2001 NAAL field test. Interviewers were not permitted to use bilingual neighbors or other household members to translate, administer, or complete any instruments used in the FAN field test.

In the oral module component, the CAPI system told the interviewer which version of the oral module to administer. The respondent simply read from the oral module booklet, as instructed by the interviewer. During the administration of the oral module, the respondent wore a headset with a microphone so that samples of his or her reading could be recorded. The interviewer wore a headset and followed along in a copy of the booklet used by the respondent, the oral module interviewer guide. The oral module interviewer guide was used as a backup to document the respondent's progress in case the recording equipment failed. Each task was timed, and when the time expired (as indicated by a beep in the interviewer's headset), the interviewer indicated the last word read by the respondent by circling it in the oral module interviewer guide.

### 5.3.5 Quality Control Measures and Feedback to Staff

The field manager and supervisors, in addition to the home office staff, monitored production daily. The home office staff also assessed the quality of the data received from the interviewers and provided feedback as necessary. The study hotline staff supported the FAN field test by responding to interviewer calls.

### 5.3.6 Summary of Data Collection Results

A total of 520 interviews were completed during the FAN field test. All but 13 interviews had an accompanying zip disk containing the respondent voice recordings; all interviews had a paper backup (oral module interviewer guide) available, as needed, for analysis.

The sample requirements were based on race/ethnicity, educational attainment, and the respondent's status as a native or nonnative English speaker.

For race/ethnicity, the goal was to complete interviews with approximately 125 respondents from each of the following groups: Black, Hispanic, Asian, and White. This goal was exceeded in all but the Asian category, which was slightly below the goal.

In the educational attainment category, the goal was to interview 250 respondents in each of two attainment groups (high and low). High educational attainment was defined as having a high school diploma (or general education diploma [GED]) or higher. Low educational attainment was defined as not having a high school diploma or GED. Achieving this goal was complicated by the difficulty in recruiting Asian respondents with less than a high school education until the standard was relaxed to permit recruiting additional Asian respondents with a high school education.

In the native/nonnative English speaker category, the goal was to interview approximately one-fourth of the Hispanic and Asian respondents in each of the following categories: native English speaker with high educational attainment, native English speaker with low educational attainment, nonnative English speaker with high educational attainment, and nonnative English speaker with low educational attainment. The field test came very close to meeting the goal for Hispanic respondents but fell short among Asian respondents, particularly in the case of native English speakers with low educational attainment. Table 5-1 provides further details on completed interviews.

**Table 5-1. Percent of interviews completed in the FAN field test, by race/ethnicity, language status, and educational attainment: 2002**

| | | Race/ethnicity and educational attainment[1] | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Black | | Hispanic | | Asian | | White | | | Other | |
| Language status | Total | High | Low | High | Low | High | Low | High | Low | Not specified | High | Low |
| Total | 10.0 | 13.6 | 11.0 | 14.2 | 12.1 | 14.6 | 7.3 | 13.5 | 11.3 | 0.2 | 0.8 | 1.3 |
| Native English speaker | 69.2 | 13.6 | 11.0 | 6.9 | 5.8 | 4.4 | 1.5 | 13.5 | 11.0 | 0.2 | 0.4 | 1.0 |
| Nonnative English speaker | 30.8 | 0 | 0 | 7.3 | 6.3 | 10.2 | 5.8 | 0 | 0.4 | 0 | 0.4 | 0.4 |

[1] High educational attainment was defined as having a high school diploma (or general education diploma [GED]) or higher. Low educational attainment was defined as not having a high school diploma or GED.
NOTE: All adults of Hispanic origin are classified as Hispanic, regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as Asian are non-Hispanic Asian only. Those classified as White are non-Hispanic White only. Those classified as other include non-Hispanics of all other races, including multiracial.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

## 5.4 DATA PREPARATION AND PROCESSING

The data preparation and processing activities and procedures implemented for the FAN CAPI instruments and hard-copy materials were very similar to those used for the 2001 NAAL field test. Small revisions were made to the receipt control and editing systems to accommodate the oral module instrument and associated materials. New systems and procedures were developed to support the processing of the zip disks containing the oral module recordings.

### 5.4.1 Receipt, Batching, and Editing of Cases

As cases were transmitted and processed, they were batched into groups of 10. Reports created for each batch provided information on the status of each case: dirty (hard edit encountered), suspect (soft edit encountered), or clean. The reports also included all other-specify entries and remarks entered by interviewers. These reports and the associated cases were reviewed, and data changes were made as appropriate.

The final editing model used for the 2001 NAAL field test, with one additional edit, was used to edit the background questionnaire for the FAN field test. The FAN background questionnaire editing process was primarily focused on reviewing the open-ended responses and interviewer comments.

Edited background questionnaire data were delivered in November 2002, along with the exercise and oral module status codes and the oral module timing data.

### 5.4.2 Processing and Delivery of Hard-Copy Assessment Materials

Completed assessment booklets were reviewed by the data collection contractor. They were then shipped to the analysis contractor twice a week throughout the data collection period.

### 5.4.3 Processing and Delivery of Oral Module Data

The oral module recordings were copied to zip disks twice a week by the interviewers and mailed to the data collection contractor's home office. Zip disks were receipted and backup copies of the data were made. Completed zip disks were shipped for scoring once a week.

Weekly verification reports were run to ensure that completed oral module files were received for all cases for which the interviewer transmitted a finalized case status. Staff reviewed the scoring contractor's website twice a week to resolve issues with oral module files that could not be processed.

## 5.5 REFINEMENTS FOR MAIN HOUSEHOLD STUDY

The FAN field test achieved many of its ambitious goals, particularly the development of an instrument to measure basic reading skills and the successful analysis of the related data.

An interviewer debriefing was conducted at the end of the field period to assess interviewer reaction to the training session, strategies to gain respondent cooperation, setup and administration of the oral module instrument, usability of the oral module materials and equipment, and respondent perception of the oral module tasks.

Interviewer input on the wording of the oral module interviewer instructions was obtained as well, in an effort to streamline the instructions, allow more flexibility for the interviewer, and reduce the overall oral module administration time.

These results were used to redesign the FAN, including adding functionality to skip the comprehension questions when the respondent was unable to read the passages and revising the interviewer instructions to make more of the text optional, to be read only when required by the respondent. NCES also identified steps to reduce the oral module administration time, including reducing the number of passages read by the respondent and the number of follow-up questions asked by the interviewer.

*This page is intentionally left blank.*

# CHAPTER 6

# FIELD TEST OF ADULT LITERACY SUPPLEMENTAL ASSESSMENT

Michelle Amsbary and Barbara Forsyth, Westat

The Adult Literacy Supplemental Assessment (ALSA) of the National Assessment of Adult Literacy (NAAL) was developed to gather as much information as possible about adults with limited literacy skills. The ALSA field test was designed to finalize both the ALSA questionnaire and scoring rubrics before their implementation in the NAAL main study. Its principal objective was to ensure that interviewers could score ALSA responses consistently. A three-phase field test for the ALSA was developed. The interviews took place over 3 months, starting in early September 2002. Phase 1 consisted of 10 ALSA interviews, all conducted in English. Phase 2 consisted of 36 ALSA interviews, roughly half conducted in English and the remainder in Spanish. Phase 3 consisted of four interviews, three of which were in Spanish. Six interviewers were recruited and trained for the ALSA field test, including two bilingual interviewers. Each phase consisted of recruiting respondents, training interviewers, and collecting data. Interviewer debriefing sessions were held after each phase.

## 6.1 PRETEST PHASE 1

The goal of the phase 1 test was to test the English version of the ALSA questionnaire and scoring rubrics developed through cognitive laboratory work. Two interviewers were trained to conduct the phase 1 interviews.

The sample for phase 1 comprised respondents with less than a fifth-grade education in the person's country of origin. Ten volunteers were recruited through adult basic education and literacy programs and literacy tutors in selected areas of Maryland, Virginia, and Washington, D.C. All interviews were conducted in English.

The ALSA stimulus materials included generic materials such as a Coke can or box of pancake mix, as well as materials such as grocery store ads and TV guides that are, by definition, local to only certain geographic areas. These types of materials show some similarities across locales, but also some differences. One of the goals of this pretest was to test whether the similarities were sufficient so that a single material could be used in multiple locales. If the similarities were not sufficient, these types of materials would have to be eliminated from the ALSA assessment.

In response to concerns that familiarity with the stimulus materials might bias the result, “local” and “nonlocal” versions of the questionnaire and stimulus materials were selected and tested for the field test. The local stimulus materials included three items that were specific to the geographic region in which the field test was conducted. The nonlocal stimulus materials included comparable items that were taken from some other geographic region. Each respondent saw only one of the two sets of stimulus materials. Five respondents completed the local version, and four received the nonlocal version. The pilot study was designed to provide evidence of differential performance relative to testing administered with the local versus nonlocal stimulus materials.

The seven core assessment items developed for the NAAL main study were also administered to the ALSA respondents. This approach allowed the survey designers to determine which respondents were true ALSA sample persons based on their performance on the core items.

Interviewers audiotaped all interviews for which respondents gave consent for taping. Observations of as many interviews as feasible were conducted. Written summaries of the observed interviews were produced. Both the audiotapes and written summaries were used to clarify procedures for subsequent testing.

Interviewers completed nine interviews, using the English language version of the questionnaire. The phase 1 results were used to refine the questionnaire and scoring rubrics and to revise training materials for field-test phases 2 and 3.

On the basis of performance on the core assessment, five of the nine respondents were determined to have been appropriate ALSA candidates; the remainder would not have been eligible for the ALSA.

## 6.2 PRETEST PHASE 2

Phase 2 was the first test with revised Spanish language materials after the cognitive laboratory work. It was anticipated that changes to the Spanish language materials and interview procedures would be greater than changes to the English language materials and procedures. Six interviewers were trained to conduct this pretest, two of whom were bilingual.

The data collection contractor visited adult basic education programs in Montgomery County, Maryland, to recruit adults known by teachers to have relatively low reading test scores appropriate for the ALSA assessment. Work establishments were also visited to recruit Spanish-speaking volunteers with less than a fifth-grade education. These visits were followed up with telephone calls to schedule interview appointments. Fifty adults were recruited; of these, 40 interviews were scheduled, including 20 with English-only speakers and 20 with Spanish-speaking volunteers.

On the basis of the results of the phase 1 test, the ALSA questionnaire and interviewing methods were revised slightly for phase 2. The changes were incorporated into both the English and Spanish versions of the instruments.

Thirty-six interviews were conducted, including 17 in Spanish. All interviews were successfully tape recorded. Sixteen interviews were observed.

The average administration length for the ALSA assessment in phase 2 was 50.9 minutes. The core assessment was administered in 34 of the 36 completed interviews. Of these 34 respondents, 15 passed the core assessment and 19 failed. Therefore, these 19 were true ALSA respondents and would have been classified into the ALSA according to the criteria used for the NAAL main study.

The phase 2 results were used to revise interview procedures and to further refine the scoring rubrics, interviewer materials, and training materials.

## 6.3 PRETEST PHASE 3

The goal of the small phase 3 study was to test any final revisions to interview materials and procedures before finalizing them for the NAAL main study. One bilingual interviewer with experience in the previous phases of the pretests conducted these final interviews.

In preparation for phase 3, the ALSA referral sources were reviewed to determine which were most effective in providing truly eligible ALSA respondents. It was determined that in phases 1 and 2, literacy programs provided the highest respondent eligibility rates (71 percent), followed by sources such as internal recruiting efforts and nearby business establishments (69 percent), and Montgomery County, Maryland, adult basic education programs (36 percent). This information was used to guide the recruitment effort for phase 3. Five respondents were recruited for this phase.

On the basis of phase 2 field test observations and the interviewer debriefing, several refinements were made to the ALSA questionnaire. The largest questionnaire revision involved adding scripted probes to the vocabulary items to assist interviewers when respondents gave no response. Minor revisions were made to clarify some of the rubrics and the Spanish language translations.

Four interviews (three in Spanish and one in English) were conducted during this final phase.

## 6.4 RESULTS OF THE FIELD TEST

A total of 49 interviews were completed during the three phases of the ALSA data collection. The data collection contractor observed the interviews and submitted written summaries of the observations, along with copies of the completed questionnaires and interview audiotapes, for review and analysis.

Results from all three phases were used to determine whether nonlocal materials could be effective in the main data collection effort. On the basis of the three phases of interviewing, it was decided to use nonlocal stimulus materials in the main data collection effort. The questionnaire content was revised to accommodate the selected stimulus materials, and preparations were made for the NAAL main study.

Additionally, the timing data indicated that the ALSA instrument needed to be shortened so that respondents would not be forced to struggle for an extended time, possibly becoming frustrated before completing the Fluency Addition to NAAL. A recommendation was made to shorten the ALSA instrument by eliminating six vocabulary items, five items asking respondents to read connected text, and the telephone book stimulus.

# CHAPTER 7

# SAMPLE DESIGN

Leyla Mohadjer and Thomas Krenzke, Westat

## 7.1 INTRODUCTION

The purpose of the National Assessment of Adult Literacy (NAAL) was to estimate the literacy levels of the adult population in the United States. To adequately represent the target population, the NAAL included both a household component and a prison component. The NAAL household study included two sets of household samples: a national household sample and household samples from six states. The target population for the national and state household samples consisted of adults 16 years or older who resided in housing units at the time of interview. This population of 221,020,000 adults in 2003 included persons who resided in college dormitories but excluded adults in military barracks, halfway houses, and other group quarters.[1] Each household sample was selected on the basis of a four-stage, stratified cluster sample.

As in the 1992 National Adult Literacy Survey (NALS), all states were given an opportunity to explore the skill levels of their adults by participating in the State Assessment of Adult Literacy (SAAL) part of the NAAL. (Participating states paid the cost of the additional assessments.) The states participating in SAAL were Kentucky, Maryland, Massachusetts, Missouri, New York, and Oklahoma. The NAAL and SAAL samples were combined (composited) to improve the precision of statistics for the national and SAAL samples. A prison component was added to improve the representation of the target population. The prison component of NAAL included a sample of adult inmates in federal and state prisons, which is representative of the 1,380,000 adults in prisons in 2003. Together, the household and prison samples are representative of the 222,400,000 adults in American households and prisons.

Section 7.2 provides a detailed summary of the four stages of sampling for the household samples. Section 7.3 describes the two-stage sample design for the prison sample. Section 7.4 discusses the assignment of booklet types within the household and prison samples.

---

[1] All people not living in housing units are classified as living in group quarters. There are two general categories: (1) institutionalized population (such as nursing homes and schools, hospitals, and wards for the mentally retarded) and (2) noninstitutionalized population (such as religious group quarters or emergency and transitional shelters).

## 7.2 HOUSEHOLD SAMPLES

The household samples (the national sample and the six SAAL samples) were selected on the basis of a four-stage, stratified area sample: (1) primary sampling units (PSUs) consisting of counties or groups of contiguous counties; (2) secondary sampling units (referred to as segments) consisting of area blocks; (3) housing units containing households; and (4) eligible persons within households. Person-level data were collected through a screener, a background questionnaire, the NAAL literacy assessment, and the oral module.

Section 7.2.1 presents the key features of the household samples and summarizes sample sizes at each sampling stage. Section 7.2.2 discusses the target population and the information sources used to create the frames for each stage of sampling. The selection process for each of the four stages is described in section 7.2.3.

### 7.2.1 Key Design Features and Sizes of the Household Samples

A single area sample was selected for the national NAAL sample, and six additional SAAL samples were selected for Kentucky, Maryland, Massachusetts, Missouri, New York, and Oklahoma. For each sample, the usual procedures for area sampling were followed: a stratified probability proportionate to size design was used for the first two stages and systematic random samples were drawn in the last two stages.

A key feature of the national NAAL sample was the oversampling of Black and Hispanic adults, which was accomplished by oversampling segments with high concentrations of these groups. The SAAL samples did not include any oversampling of minority groups.

Although integrating the NAAL and SAAL samples at the design stage would have been more effective statistically, the states agreed to participate after the NAAL sample design and selection process had been finalized. Therefore, the approach used in the 1992 NALS was followed: selecting the SAAL samples independently of the NAAL sample and combining the samples at the estimation phase by using composite estimation.

The first stage of sampling was the selection of PSUs, which consisted of counties or groups of counties. PSUs were formed within state boundaries, which gave an improved sample for state-level estimation. One PSU was selected per stratum by using probabilities proportionate to their population within households, except in Maryland and Massachusetts where samples of segments were selected as the first-stage units (refer to section 7.2.3.1.1 for more discussion). One hundred PSUs were selected for the national sample, and 54 PSUs were selected in Kentucky, Missouri, New York, and Oklahoma. In Maryland and Massachusetts there were too few PSUs to sample from; therefore segments were selected in the first stage of sampling. After selecting the segments, 20 area clusters (quasi-PSUs) were created for Maryland and Massachusetts by grouping the selected segments into 20 geographically clustered areas to facilitate a cost-efficient approach to data collection. The true first-stage sample size is much larger because a total of 323 first-stage units (i.e., segments) were selected in Maryland and Massachusetts. However, to not mislead readers into thinking the sample is much larger than it really is, this report uses the PSU count of 20 in Maryland and Massachusetts to describe the first-stage sampling activities. Fourteen PSUs were selected for both the national NAAL and the SAAL samples; hence, the sample included a combined total of 160 unique PSUs.

In the second stage of sampling, segments (census blocks or groups of blocks) within the PSUs were selected with a probability proportionate to size; the measure of size (MOS) for a segment was a function of the number of year-round housing units within the segment. In the national sample, the Black and Hispanic populations were sampled at a higher rate than the remainder of the population to increase their sample size, whereas the state samples used no oversampling. Oversampling in the national sample was accomplished by oversampling the high-minority segments in which Black and Hispanic adults accounted for 25 percent or more of the population. There were 1,959 segments selected for the national sample and 861 segments selected across the SAAL samples, with a total of 2,818 unique segments selected across the national and six SAAL samples. (Two segments were selected for both the NAAL and SAAL samples.)

In the third stage of sampling, housing units were selected with equal probability within each segment, except for nonminority households within high-minority segments in the national component. These national sample households were subsampled after screening so that the sampling rates for nonminority persons would be about the same in the high-minority segments as in other segments. The overall sample size of housing units took into account expected losses owing to vacant housing units, units that were not housing units, and expected response rates.[2]

[2] The expected response rates took into account those experienced in the 1992 NALS and recent trends in household studies.

The fourth stage of selection involved listing the age-eligible household members (aged 16 and older) for each selected household. Subsequently, one person was selected at random within households with three or fewer eligible persons, and two persons were selected if the household had four or more eligible persons. The listing and selection of persons within households were performed with the computer-assisted personal interviewing (CAPI) system.

Table 7-1 contains the sample sizes of PSUs, segments, housing units, and persons and the number of persons completing the background questionnaire for the national NAAL household sample.

**Table 7-1. NAAL sample sizes of PSUs, segments, housing units, and persons and numbers of completed background questionnaires, by region, Metropolitan Statistical Area (MSA) status, and segment status: 2003**

| Characteristic | Primary sampling units (percent) | Segments | Housing units | Persons | Completed background questionnaires[1] |
|---|---|---|---|---|---|
| Total | 100 | 1,959 | 24,450 | 16,409 | 12,753 |
| Region[2] | | | | | |
| Northeast | 18 | 373 | 4,510 | 2,864 | 2,228 |
| Midwest | 23 | 404 | 4,833 | 3,354 | 2,688 |
| South | 38 | 758 | 10,295 | 6,405 | 4,943 |
| West | 21 | 424 | 5,812 | 3,786 | 2,894 |
| MSA status | | | | | |
| Non-MSA | 22 | 357 | 4,282 | 2,862 | 2,295 |
| MSA | 78 | 1,602 | 21,168 | 13,547 | 10,458 |
| Segment status[3] | | | | | |
| Low minority | † | 1,091 | 11,648 | 8,441 | 6,394 |
| High minority | † | 868 | 13,802 | 7,968 | 6,359 |

† Not applicable.
[1] Completed background questionnaires included cases that were not complete due to language problems and mental disabilities. These cases were considered a “success” in data collection since race/ethnicity, age, and gender were collected, as well as good information (language problem or mental disability) as to their English literacy skills.
[2] Northeast Region = ME, NH, VT, MA, RI, CT, NY, NJ, PA; Midwest Region = OH, IN, IL, MI, WI, MN, IA, MO, ND, SD, NE, KS; South Region = DE, MD, DC, VA, WV, NC, SC, GA, FL, KY, TN, AL, MS, AR, LA, OK, TX; West Region = MT, ID, WY, CO, NM, AZ, UT, NV, WA, OR , CA, AK, HA.
[3] Segments classified as low minority have less than 25 percent Black and Hispanic adults. Segments classified as high minority have 25 percent or more Black and Hispanic adults. The cutpoint of 25 percent of the population being Black and Hispanic was derived analytically for the 1992 NALS. An analysis for the 1992 sample design showed that the 25 percent concentration provided the minority sample sizes at acceptable design effect levels. The same cutpoint was used for the 2003 NAAL study.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

The SAAL design called for a sample of 12 PSUs in each state. The first stage of sampling for Maryland and Massachusetts was the selection of segments because these states had small numbers of PSUs (i.e., Maryland and Massachusetts had three stages of sampling instead of four; refer to section 7.2.3.1.1 for a discussion of the selection of segments as the first-stage units in Maryland and Massachusetts). Eighteen PSUs instead of 12 were selected for Kentucky to improve the precision of small-area estimation as requested by this state. Twelve PSUs were selected for Missouri, New York, and Oklahoma.

The NAAL and SAAL samples were integrated through a compositing procedure to achieve the maximum efficiency for producing both national and state estimates for the participating states. Table 7-2 summarizes the sample sizes for the combined NAAL-SAAL sample, as well as the numbers of completed background questionnaires. More information about sample yields can be found in the tables in section 8.8.3.

**Table 7-2. NAAL-SAAL combined sample sizes of PSUs, segments, housing units, and persons and numbers of completed background questionnaires, by region, Metropolitan Statistical Area (MSA) status, and state: 2003**

| Characteristic | Primary sampling units[1] | Segments | Housing units | Persons | Completed background questionnaires[2] |
|---|---|---|---|---|---|
| Total | 174 | 2,820 | 35,365 | 23,732 | 18,541 |
| Total unduplicated | 160 | 2,818 | 35,365 | 23,732 | 18,541 |
| | | | | | |
| Region[3] | | | | | |
| Northeast | 37 | 636 | 7,518 | 4,936 | 3,765 |
| Midwest | 35 | 538 | 6,332 | 4,500 | 3,612 |
| South | 81 | 1,222 | 15,703 | 10,510 | 8,270 |
| West | 21 | 424 | 5,812 | 3,786 | 2,894 |
| | | | | | |
| MSA status | | | | | |
| Non-MSA | 59 | 584 | 6,948 | 4,835 | 3,897 |
| MSA | 115 | 2,236 | 28,417 | 18,897 | 14,644 |
| | | | | | |
| State | | | | | |
| Kentucky | 20 | 229 | 2,696 | 1,945 | 1,545 |
| Maryland | 14 | 150 | 1,727 | 1,276 | 1,016 |
| Massachusetts | 9 | 173 | 2,002 | 1,432 | 1,074 |
| Missouri | 13 | 149 | 1,635 | 1,247 | 1,009 |
| New York | 20 | 310 | 3,680 | 2,251 | 1,730 |
| Oklahoma | 14 | 164 | 1,924 | 1,533 | 1,287 |
| Rest of U.S. | 84 | 1,645 | 22,701 | 14,048 | 10,880 |

[1] Because segments were the first stage of sampling for Maryland and Massachusetts, PSU counts for these states represent groups of segments formed for field management purposes.
[2] Includes respondents who did not complete the background questionnaire because of language problems or mental disabilities.
[3] Northeast Region = ME, NH, VT, MA, RI, CT, NY, NJ, PA; Midwest Region = OH, IN, IL, MI, WI, MN, IA, MO, ND, SD, NE, KS; South Region = DE, MD, DC, VA, WV, NC, SC, GA, FL, KY, TN, AL, MS, AR, LA, OK, TX; West Region = MT, ID, WY, CO, NM, AZ, UT, NV, WA, OR , CA, AK, HA.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 7.2.2 Sampling Frames for the Household Study

The target population for the national and state household studies consisted of adults 16 years of age or older who resided in housing units at the time of interview. Area sampling methodology was used to facilitate the selection of a representative sample from the target population. Area sampling requires the formation of frames at each stage of sampling. The next subsections discuss the creation of sampling frames for each stage of selection.

#### 7.2.2.1 Sampling Frame for Primary Sampling Units

For the initial stage of sampling, a PSU frame was created by using the census 2000 public law (PL)-94 county–level files. The PSUs were formed by combining adjacent counties, respecting their population sizes and taking into consideration the travel distance for interviewers. The PSUs were formed as a single county or a group of contiguous counties, depending on the population size and the end-to-end distance within a PSU.

One set of PSUs was created for both the NAAL and the SAAL household samples. The objective of the PSU formation process was to minimize travel distance within a PSU (where the maximum distance was 100 miles), subject to a minimum population size (i.e., a minimum MOS) in a PSU of 15,000. The census 2000 PL-94 county-level data were used to obtain county-level population sizes. The MOS variable was an estimate of the population counts within households derived from the total population counts from the 2000 census.

The PSUs were formed within MSA boundaries. They also were formed within state boundaries, with the exception of two PSUs that were not selected in either the NAAL sample or the SAAL sample. A total of 1,884 PSUs were formed. Table 7-3 presents the distribution of counts of PSUs by region and MSA status, and table 7-4 presents some characteristics of the PSUs by MSA status. The PSU frame for SAAL states was created by subsetting the PSUs from the NAAL PSU frame. For states where PSUs were selected as the first stage, the numbers of PSUs formed before sampling were 73 (Kentucky), 65 (Missouri), 39 (New York), and 45 (Oklahoma).

**Table 7-3. Distribution of formed primary sampling units, by Metropolitan Statistical Area (MSA) and region status: 2003**

| Region[1] | Total | Non-MSA | MSA |
|---|---|---|---|
| Total | 1,884 | 1,497 | 387 |
| Northeast | 143 | 88 | 55 |
| Midwest | 610 | 510 | 100 |
| South | 853 | 695 | 158 |
| West | 278 | 204 | 74 |

[1] Northeast Region = ME, NH, VT, MA, RI, CT, NY, NJ, PA; Midwest Region = OH, IN, IL, MI, WI, MN, IA, MO, ND, SD, NE, KS; South Region = DE, MD, DC, VA, WV, NC, SC, GA, FL, KY, TN, AL, MS, AR, LA, OK, TX; West Region = MT, ID, WY, CO, NM, AZ, UT, NV, WA, OR , CA, AK, HA.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 7-4. Distributions of primary sampling unit characteristics, by Metropolitan Statistical Area (MSA) status, mean, median, and percentiles: 2003**

| Variable | Category | Mean | Median | First quartile | Third quartile | 95th percentile |
|---|---|---|---|---|---|---|
| Population size | Non-MSA | 36,344 | 29,441 | 20,479 | 42,608 | 83,340 |
| | MSA | 577,987 | 255,403 | 134,309 | 627,018 | 2,125,203 |
| | Total | 147,605 | 35,784 | 22,354 | 73,997 | 650,700 |
| Distance (miles)[1] | Non-MSA | 60.3 | 46.8 | 37.3 | 65.1 | 129.8 |
| | MSA | 70.6 | 62.3 | 47.4 | 79.8 | 115.9 |
| | Total | 62.4 | 49.3 | 38.3 | 69.3 | 127.7 |
| Area (square miles) | Non-MSA | 1,878 | 823 | 555 | 1,422 | 6,082 |
| | MSA | 1,840 | 1,364 | 805 | 2,097 | 4,204 |
| | Total | 1,870 | 900 | 572 | 1,716 | 5,660 |
| Number of counties | Non-MSA | 1.5 | 1.0 | 1.0 | 2.0 | 3.0 |
| | MSA | 2.2 | 2.0 | 1.0 | 3.0 | 5.0 |
| | Total | 1.7 | 1.0 | 1.0 | 2.0 | 4.0 |

[1] The distance was computed as the end-to-end distance.. It is beneficial to minimize this distance in order to minimize travel distance for interviewers.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 7.2.2.2 Sampling Frame for Segments

For the second stage of sampling, a frame of segments was created within the selected PSUs by using the census 2000 summary file (SF) block data. A segment consists of a census block[3] (as defined by Census 2000) or a combination of two or more nearby blocks. Within each PSU, the block data from the 2000 census SF1 files were sorted by tract, block group, and block number before creating the segments.

Blocks with no housing units and no population were also included in the formation process in order to include all housing units constructed after the 2000 census. A single block was used as a segment when the number of housing units in the block exceeded 60. Neighboring blocks were combined within a tract to reach either the required minimum of 60 housing units per segment or the end of the tract (segments did not cross tract boundaries).

[3] Blocks are very fine partitions of the United States, formed by using visible semipermanent features such as roads, railroad tracks, mountain ridges, bodies of water, and power lines. The only invisible boundaries used are county, state, and national boundaries. Minor civil division boundaries and property lines are ignored. A block group is a small group of contiguous blocks. A tract is a collection of contiguous block groups all within the same county.

A total of 413,523 segments were formed for the NAAL PSUs, including NAAL PSUs that overlapped with the SAAL PSUs. Another 39,769 segments were formed within PSUs selected for the SAAL only. Therefore, a total of 453,292 segments were formed for the combined NAAL-SAAL PSU sample.

#### 7.2.2.3 Sampling Frame for Housing Units

After segments were selected, the data collection contractor's listers visited each sampled segment to create a sampling frame of housing units for the third stage of sampling. Interviewers constructed a list of all housing units within the segment boundaries, using tract and segment maps created by home office staff. A small number of segments were subdivided, with one part, or "chunk," selected at random for listing. Chunking reduced the burden of listing large sampled segments (generally more than 300 housing units) by dividing the segment into chunks. A chunk was selected with probability proportionate to the estimated number of housing units within the chunk, and listing was conducted within the selected chunk. Of the 2,818 segments selected, 334 were sampled chunks of segments. Table 7-5 shows the distribution of the segments by size.

**Table 7-5. Percentile distributions of NAAL-SAAL sampled segments: 2003**

| Percentile | Actual number of housing units listed[1] |
|---|---|
| 5th | 59 |
| 25th | 71 |
| 50th | 98 |
| 75th | 154 |
| 95th | 238 |
| Mean | 118 |

[1] Counts reflect number of housing units listed, whether from a chunk or a full segment.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

#### 7.2.2.4 Sampling List of Persons

At the fourth stage of sampling, households were screened to determine whether they included any eligible persons, defined as a person 16 years or older who resided in the household; the definition of eligibility included persons who resided in college dormitories but excluded adults in military barracks, halfway homes, and group quarters. A complete list of household members was obtained by interviewers and entered in the CAPI program as part of the screener interview conducted in each sampled household. Before sample selection, the CAPI system determined the eligibility of each person listed.

### 7.2.3 Sample Selection for the Household Study

The NAAL sample and the SAAL sample were selected independently on the basis of a four-stage, stratified cluster sample involving the selection of PSUs, segments, housing units, and eligible persons within households in the selected housing units. The selection of PSUs is discussed in section 7.2.3.1. The selection of segments within PSUs and of housing units within sampled segments is discussed in sections 7.2.3.2 and 7.2.3.3, respectively. The sampling of persons is discussed in section 7.2.3.4. Section 7.2.3.5 describes an approach to improve the coverage of the housing unit sampling frame, and section 7.2.3.6 discusses the probabilities of selection. Differences between the NAAL and SAAL sampling procedures are pointed out within each section. Selection results are summarized for the NAAL sample, the SAAL sample, and the combined NAAL-SAAL sample.

#### 7.2.3.1 Selection of Primary Sampling Units

The PSUs were selected as a stratified probability-proportionate-to-MOS sample, where the MOS was equal to the household population. The PSUs were then stratified on the basis of available variables from the census 2000 PL-94 files. The stratification is described in section 7.2.3.1.1, and the selection process is discussed in section 7.2.3.1.2.

##### 7.2.3.1.1 Stratification of Primary Sampling Units

The NAAL and SAAL samples involved the selection of PSUs, with one PSU selected per stratum. For this selection, the PSUs on the frame were stratified into homogeneous strata. The PSUs with the largest MOS were selected with certainty. Each of the certainty PSUs was treated as a single stratum, and the remaining PSUs were stratified into the appropriate number of noncertainty strata.

The certainty PSUs were identified before the application of the stratification algorithm. The certainty PSUs were the largest PSUs (in terms of MOS) and were selected with probability equal to 1. The certainty cutoff was determined from probability proportionate to size sampling, with the total population in households as the MOS.

On the basis of the analytical cutoff, PSUs were selected with certainty independently for the NAAL and the SAAL samples. The remaining PSUs for each respective sample, excluding the certainty

PSUs, were stratified into noncertainty strata. The NAAL sample stratification process initially formed major strata defined by census division and MSA status (except in the New England and Mid-Atlantic census divisions, where non-MSAs were combined into one major stratum).

Each major stratum was further stratified into the allocated number of substrata. The main objective of the substratification process was to keep the substratum (i.e., ultimate stratum) sizes as equal as possible, both to reduce the variation in workload and to control the variances of the estimates. Table 7-6 presents the variables used for NAAL substratification within each major stratum. The variables used in the substratification process were identified through a regression analysis. The dependent variable for the stepwise regression analysis was the percentage of the population that were high school graduates 25 years and older. Limited by the time that the stratification occurred (just after Census 2000), the independent variables were census division, MSA status, per capita income, percentage of Non-Hispanic black population, percentage of Hispanic population, percentage of non-minority population, and the PSU population size.

The sample designs for Maryland and Massachusetts required one fewer sampling stage than the designs for the other SAAL states and the NAAL sample. Because the numbers of PSUs formed within Maryland and Massachusetts were small, a sample of segments was selected across each state. Within each of these two states, segments were sampled with probability proportionate to size. Table 7-7 presents the variables used for the SAAL substratification process. The variables used in the substratification process were identified through a stepwise regression analysis, using the same variables as described in the preceding paragraph for the NAAL stratification.

## Table 7-6. Variables used in NAAL noncertainty primary sampling unit stratification: 2003

| Census division | Metropolitan Statistical Area (MSA) status | Stratification variables |
|---|---|---|
| New England Division (1) and Middle Atlantic Division (2) | Non-MSA | Per capita income |
| East North Central Division (3) and West North Central Division(4) | Non-MSA | Per capita income, percentage nonminority |
| South Atlantic Division (5) East South Central Division (6) and West South Central Division (7) | Non-MSA | Per capita income, percentage non-Hispanic Black |
| Mountain Division (8) and Pacific Division (9) | Non-MSA | Per capita income |
| New England Division (1) and Middle Atlantic Division (2) | MSA | Per capita income, percentage Hispanic |
| East North Central Division (3) | MSA | Per capita income, percentage non-Hispanic Black |
| West North Central Division (4) | MSA | Per capita income |
| South Atlantic Division (5) | MSA | Per capita income, percentage non-Hispanic Black, percentage Hispanic |
| East South Central Division (6) | MSA | Per capita income, percentage non-Hispanic Black |
| West South Central Division (7) | MSA | Per capita income, percentage non-Hispanic Black, percentage Hispanic |
| Mountain Division (8) and Pacific Division (9) | MSA | Per capita income, percentage nonminority |

NOTE: New England Division (1) = ME, NH, VT, MA, RI, and CT; Middle Atlantic Division (2) = NY, NJ, and PA; East North Central Division (3) = OH, IN, IL, MI , and WI; West North Central Division (4) = MN, IA, MO, ND, SD, NE, and KS; South Atlantic Division (5) = DE, MD, DC, VA, WV, NC, SC, GA, and FL; East South Central Division (6) = KY, TN, AL, and MS; West South Central Division (7) = AR, LA, OK, and TX; Mountain Division (8) = MT, ID, WY, CO, NM, AZ, UT, and NV; and Pacific Division (9) = WA, OR, CA, AK, and HI.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 7-7. Variables used in SAAL noncertainty primary sampling unit stratification: 2003**

| State | Metropolitan Statistical Area (MSA) status | Stratification variables |
|---|---|---|
| New York | Non-MSA | Per capita income |
| | MSA | Per capita income, percent Hispanic |
| Missouri | Non-MSA | Per capita income, percent nonminority |
| | MSA | Per capita income |
| Kentucky and Oklahoma | Non-MSA | Per capita income, percent non-Hispanic Black |
| | MSA | Per capita income, percent non-Hispanic Black, percentage Hispanic |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

#### 7.2.3.1.2 PSU Selection

One PSU was selected independently in each stratum for the NAAL and SAAL samples. A PSU in a certainty stratum had a selection probability of 1, and the PSUs in noncertainty strata were selected with probability proportional to measure of size (MOS) (i.e., the total population within households). The formula for the PSU selection probability, involving the PSU MOS, is provided in table 7-8.

## Table 7-8. NAAL-SAAL selection probabilities: 2003

| Sampling unit | Conditional probability[1] | Overall probability |
|---|---|---|
| PSUs | † | Let $x_{hi}$ = total population in households for PSU $i$ within stratum $h$. <br> $P_{hi} = \frac{x_{hi}}{\sum_{i \in h} x_{hi}}$ |
| Segments | Let $HU_{Nij}$ = number of nonminority housing units (HUs) expected in segment $j$ of PSU $i$. <br> Let $HU_{Mij}$ = number of minority housing units expected in segment $j$ of PSU $i$. <br> Let $MOS_{1ij} = HU_{Nij} + 3HU_{Mij}$, if high-minority NAAL segment.[1] <br> Let $MOS_{2ij} = HU_{Nij} + HU_{Mij}$, if low-minority NAAL segment.[1] <br> Let $MOS_{3ij} = HU_{Nij} + HU_{Mij}$, if SAAL segment. <br> For high-minority NAAL segments, the conditional probability is <br> $CP_{ij} = \frac{n\left[MOS_{1ij}/P_{hi}\right]}{\sum_{ij}\left[MOS_{1ij}/P_{hi}\right] + \sum_{ij}\left[MOS_{2ij}/P_{hi}\right]} = \frac{MOS_{1ij}/P_{hi}}{I_{NAAL}}$, <br> and for low-minority NAAL segments, the conditional probability is <br> $CP_{ij} = \frac{n\left[MOS_{2ij}/P_{hi}\right]}{\sum_{ij}\left[MOS_{1ij}/P_{hi}\right] + \sum_{ij}\left[MOS_{2ij}/P_{hi}\right]} = \frac{MOS_{2ij}/P_{hi}}{I_{NAAL}}$, <br> where $n$ = total size of the target segment sample for NAAL. <br> For SAAL segments, the conditional probability is <br> $CP_{ij} = \frac{n\left[MOS_{3ij}/P_{hi}\right]}{\sum_{ij}\left[MOS_{3ij}/P_{hi}\right]} = \frac{MOS_{3ij}/P_{hi}}{I_{SAAL}}$, <br> where $n$ = total size of the target segment sample for SAAL. | For high-minority NAAL segments: <br> $P_{ij} = P_{hi} \times \frac{MOS_{1ij}/P_{hi}}{I_{NAAL}} = \frac{MOS_{1ij}}{I_{NAAL}}$. <br> For low-minority NAAL segments: <br> $P_{ij} = P_{hi} \times \frac{MOS_{2ij}/P_{hi}}{I_{NAAL}} = \frac{MOS_{2ij}}{I_{NAAL}}$. <br> For SAAL segments: <br> $P_{ij} = P_{hi} \times \frac{MOS_{3ij}/P_{hi}}{I_{SAAL}} = \frac{MOS_{3ij}}{I_{SAAL}}$. |

See notes at end of table.

**Table 7-8. NAAL-SAAL selection probabilities: 2003—Continued**

| Sampling unit | Conditional probability | Overall probability |
|---|---|---|
| Chunking segments and NAAL/SAAL overlap segments | Let $HU_{ij(chunk)}$ = number of housing units expected in the selected chunk in segment *j* of PSU *i*.<br><br>Let $HU_{ij}$ = number of housing units expected in segment *j* of PSU *i*.<br><br>$CP_{ij(chunk)} = \begin{cases} \dfrac{HU_{ij(chunk)}}{HU_{ij}}, & \text{if chunking is implemented.} \\ 1, & \text{otherwise.} \end{cases}$ | For high-minority NAAL segments:<br>$P_{ij(chunk)} = \dfrac{MOS_{1ij}}{I_{NAAL}} \times \dfrac{HU_{ij(chunk)}}{HU_{ij}}.$<br><br>For low-minority NAAL segments:<br>$P_{ij(chunk)} = \dfrac{MOS_{2ij}}{I_{NAAL}} \times \dfrac{HU_{ij(chunk)}}{HU_{ij}}.$<br><br>For SAAL segments:<br>$P_{ij(chunk)} = \dfrac{MOS_{3ij}}{I_{SAAL}} \times \dfrac{HU_{ij(chunk)}}{HU_{ij}}.$ |
| Housing units | Let $\bar{d}$ = average number (integer) of housing units to be sampled within NAAL segments. Let $\bar{d}_{state}$ = average number (integer) of housing units to be sampled within SAAL segments. Then the within-segment sampling rate (i.e., conditional selection probability) is<br><br>$CP_{ijk} = \dfrac{3\bar{d} / CP_{ij(chunk)}}{HU_{ij}}$ for high-minority NAAL segments,<br><br>$CP_{ijk} = \dfrac{\bar{d} / CP_{ij(chunk)}}{HU_{ij}}$ for low-minority NAAL segments, and<br><br>$CP_{ijk} = \dfrac{\bar{d}_{state} / CP_{ij(chunk)}}{HU_{ij}}$ for SAAL segments within a state. | For housing units in high-minority NAAL segments:<br>$P_{ijk} = \dfrac{MOS_{1ij}}{I_{NAAL}} \times \dfrac{HU_{ij(chunk)}}{HU_{ij}} \times \dfrac{3\bar{d} / CP_{ij(chunk)}}{MOS_{1ij}}$<br>$= \dfrac{3\bar{d}}{I_{NAAL}}.$<br><br>For housing units in low-minority NAAL segments:<br>$P_{ijk} = \dfrac{MOS_{2ij}}{I_{NAAL}} \times \dfrac{HU_{ij(chunk)}}{HU_{ij}} \times \dfrac{\bar{d} / CP_{ij(chunk)}}{MOS_{2ij}}$<br>$= \dfrac{\bar{d}}{I_{NAAL}}.$ |

See notes at end of table.

**Table 7-8. NAAL-SAAL selection probabilities: 2003—Continued**

| Sampling unit | Conditional probability | Overall probability |
|---|---|---|
| Housing units—Continued | | For SAAL housing units: $P_{ijk} = \frac{MOS_{3ij}}{I_{SAAL}} \times \frac{HU_{ij(chunk)}}{HU_{ij}} \times \frac{\bar{d}_{state} / CP_{ij(chunk)}}{MOS_{3ij}} = \frac{\bar{d}_{state}}{I_{SAAL}}.$ |
| Missed housing units | If a large number of housing units were found during either the missed structure or hidden housing unit procedure, a sample of *mhu* missed housing units was selected from the *MHU* missed housing units identified. This occurred for the missed structure process only. $CP_{ijk(mhu)} = \begin{cases} \frac{mhu_{ijk}}{MHU_{ijk}} & \text{for missed HU sample or} \\ 1, & \text{otherwise.} \end{cases}$ | For NAAL housing units: $P_{ijk(mhu)} = \frac{\bar{d}}{I_{NAAL}} \times \frac{mhu_{ijk}}{MHU_{ijk}}.$ For SAAL housing units: $P_{ijk(mhu)} = \frac{\bar{d}_{state}}{I_{SAAL}} \times \frac{mhu_{ijk}}{MHU_{ijk}}.$ |
| Persons | Let $M$ = number of eligible persons within a household. If $M \le 3$, then $m = 1$. If $M > 3$, then $m = 2$. $CP_{ijkl} = \frac{m_{ijk}}{M_{ijk}}$ | For NAAL: $P_{ijkl} = \frac{\bar{d}}{I_{NAAL}} \times \frac{mhu_{ijk}}{MHU_{ijk}} \times \frac{\mathrm{m}_{\mathrm{ijk}}}{\mathrm{M}_{\mathrm{ijk}}} \times 0.95,$ where the factor 0.95 is applied because the reserve sample was not released. For SAAL: $P_{ijkl} = \frac{\bar{d}}{I_{SAAL}} \times \frac{mhu_{ijk}}{MHU_{ijk}} \times \frac{\mathrm{m}_{\mathrm{ijk}}}{\mathrm{M}_{\mathrm{ijk}}} \times 0.95.$ |

† Not applicable.

[1] High-minority segments are segments with 25 percent or more Blacks and Hispanics; otherwise segments are classified as low-minority segments.

NOTE: h = stratum; i = PSU; j = segment; k= HU; l = person; Phi is the overall probability of selection for PSU i within stratum h; Pij is the overall probability of selection for segment j within PSU I; INAAL is the sampling interval for the selection of the NAAL segments; ISAAL is the sampling interval for the selection of the SAAL segments.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

For the NAAL sample, 100 PSUs were selected. One PSU was selected from each of 100 (16 certainty and 84 non-certainty) strata. Table 7-1 provides the sample sizes of PSUs by census region and MSA status. Table 7-9 shows the number of PSUs selected in each sample by certainty status.

**Table 7-9. Distribution of sampled PSUs, by certainty status: 2003**

| Sample | Total[1] | PSU certainty status: Noncertainty | PSU certainty status: Certainty |
|---|---|---|---|
| NAAL | 100 | 84 | 16 |
| | | | |
| SAAL | | | |
| Kentucky | 18 | 15 | 3 |
| Maryland | 13 | † | † |
| Massachusetts | 7 | † | † |
| Missouri | 12 | 8 | 4 |
| New York | 12 | 4 | 8 |
| Oklahoma | 12 | 9 | 3 |

† Not applicable.
[1] Because segments were the first stage of sampling for Maryland and Massachusetts, PSU counts for these states represent groups of segments formed for field management purposes.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

The NAAL and SAAL PSU samples were selected independently. As a result, 14 PSUs were selected for both the NAAL and the SAAL (many of them being certainty PSUs). Table 7-10 presents the number of overlapping PSUs by SAAL state.

**Table 7-10. Number of overlapping PSUs in the NAAL and SAAL samples: 2003**

| State | Number of overlapping PSUs[1] |
|---|---|
| Total | 14 |
| | |
| Kentucky | 1 |
| Maryland | 1 |
| Massachusetts | 2 |
| Missouri | 1 |
| New York | 7 |
| Oklahoma | 2 |

[1] Because segments were the first stage of sampling for Maryland and Massachusetts, PSU counts for these states represent groups of segments formed for field management purposes.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 7.2.3.2 Selection of Segments

To increase the number of Black and Hispanic adults in the NAAL sample, segments with moderate to high concentrations of Black and Hispanic adults were given a higher selection probability. Segments in which Blacks or Hispanics accounted for 25 percent or more of the population were oversampled at a rate up to three times that of the remainder of the segments. The housing unit counts served as the MOS for the low-minority segments (segments in which Blacks and Hispanics accounted for less than 25 percent of the population). In high-minority segments, the MOS was the number of White, non-Hispanic households plus three times the number of Black and Hispanic households. Table 7-8 shows the MOS and the probabilities of selection for the second-stage sample. A minimum MOS of 5 was assigned to each segment with fewer than five housing units. Ordering the frame of segments within each PSU by the proportion of Blacks and Hispanics in the segment provided an implicit stratification by minority group for the second-stage sample. A systematic probability proportionate to size sample of segments was selected from the sorted frames. A total of 1,959 segments were selected for the national NAAL sample.

For the SAAL states, a total of 861 segments were selected with probability proportional to MOS, with no oversampling of minority groups. The MOS and probability of selection are defined in table 7-8. The segments were selected independently within each state. For overlapping PSUs, the segments were selected independently for the NAAL and the SAAL. Two segments were selected for both the national NAAL sample and a SAAL sample. Therefore, a total of 2,818 unique segments were selected in the combined sample.

### 7.2.3.3 Selection of Housing Units

The third stage of sampling involved selecting housing units from the frame of addresses in each segment, prepared after the listing operations were complete. After addresses had been selected, the interviewers contacted and screened households to determine whether they included any eligible respondent. Then, in low-minority segments, any household with at least one eligible person was included in the sample. In high-minority segments, which were oversampled, all minority households with at least one eligible person were retained in the sample, but only one-third of the nonminority households (with at least one eligible person) were included in the sample. This subsampling was done so that across all sampled segments, the resulting probabilities for nonminority households were equal under this scheme, which improves the precision of estimates for nonminorities since the variation in sampling probabilities was removed. In the SAAL samples, all households with at least one eligible person were retained in the sample (the SAAL samples did not include any oversampling of minority groups).

The national NAAL sampling approach ensured that the following conditions were met:

- At least one person was selected in each sampled household containing eligible persons 16 years or older, except that in nonminority households in high-minority segments, one-third of households were retained in the sample.
- In households of equal size
  - the probabilities of selection for persons in nonminority households were equal; and
  - the probabilities of selection for persons in minority households in high-minority segments were three times the selection probability of persons in all other selected households.

The housing unit probabilities of selection are presented in table 7-8. A reserve sample of about 5 percent the size of the main sample was selected randomly and set aside to be used in case of a shortfall in the sample. The reserve sample was not released. Housing unit sample sizes are provided in table 7-2.

The following quality control checks are examples of steps taken to ensure the high quality of the samples selected within the sampled segments:

- For the purpose of monitoring the listing operation, a range was generated for the expected number of housing units in the segment (+ 10 percent of the 2000 census count). Whenever the number of listed housing units fell outside the range, the lister called the data collection contractor and provided the reasons for the discrepancy.
- The within-segment sampling rates were applied to the count of listed housing units in the segment. This number was compared with the housing units subsampled from the field listing. This approach provided a check on the subsampling operations.
- From the listings, the address and identification number were keyed and verified. The keyed listings were checked by home office staff against the listing sheets. Any necessary corrections were made before household folders and assignment logs were produced.
- As the survey got under way, regular quality checks on the age, gender, and race distributions of the sample persons were made by comparing actual with expected distributions for the PSU.

The sample assignments, which specify the households and within-household subsampling, were checked by applying the algorithm used to generate them to the segment estimates of persons and households for minority and nonminority sample cases.

#### 7.2.3.4 Selection of Persons

After selection, the addresses of the sampled housing units were loaded into the CAPI system. The within-household sampling was conducted by randomly selecting one adult from households with three or fewer eligible persons or two adults in households with four or more eligible persons. The random selection algorithm was programmed into the CAPI system, and the selected person(s) was displayed on the screen. The random selection of respondents was accomplished by assigning random numbers to each eligible person in the household and selecting the person(s) with the smallest (or the two smallest) random number(s). The selection of two adults in households with four or more eligible persons prevented a substantial increase in variances owing to high sampling weights, which would have resulted if the survey had selected only one person in households with large numbers of eligible adults. Sizes of the person samples are provided in table 7-2.

Most residents of college dormitories were expected to be available for interview at their family homes because the data collection period included the spring and summer breaks. However, if it was not possible to reach students at their family homes, arrangements were made to interview as many as feasible in their dormitories.

#### 7.2.3.5 Procedures for Selecting Missed Structures and Hidden Housing Units

The missed structure and hidden housing unit procedures were developed to correct for any undercoverage that occurred during the listing operation. Procedures were implemented during data collection to handle any housing units identified through the hidden housing unit and missed structure procedures. The hidden housing unit procedure looked for housing units within a structure not included during the listing operation. If five or more hidden housing units were found, the statistician and the field director determined whether a sample of the hidden housing units had to be selected. Any sampling reduced the amount of interview work and clustering within the segments; however, it also increased the sampling error because of the unequal probabilities of selection resulting from the subsampling procedure.

For the missed structure procedure, interviewers looked for entire structures missed during the listing operation within a subsample of segments. The subsample of segments designated for the quality check was selected at a rate such that the inclusion of all units found retained the self-weighting feature of the sample stratum. If more than five missed structures were found in the segment, the statistician and the field director determined whether a sample of the missed structures had to be selected. As with the missed housing unit procedure, any sampling reduced the amount of interview work and clustering within the segments; however, it also increased the sampling error because of the resulting unequal probabilities of selection.

In the national NAAL sample, 288 housing units were added through the missed structure procedure, and 153 housing units were added through the hidden housing unit procedure. Table 7-11 provides the number of housing units added by the missed structure and hidden housing unit procedures for the NAAL sample and each of the SAAL states. In total, 1.8 percent of the NAAL/SAAL combined sample consisted of housing units added through the missed structure and hidden housing unit procedures.

**Table 7-11. Number of housing units added by the missed structure and hidden housing unit procedures: 2003**

| Sample | Total | Housing units added through missed structure procedure | Housing units added through hidden housing unit procedure |
|---|---|---|---|
| Total combined NAAL/SAAL | 662 | 455 | 207 |
| NAAL | 441 | 288 | 153 |
| SAAL | 221 | 167 | 54 |
| Kentucky | 40 | 35 | 5 |
| Maryland | 19 | 10 | 9 |
| Massachusetts | 43 | 23 | 20 |
| Missouri | 8 | 7 | 1 |
| New York | 18 | 0 | 18 |
| Oklahoma | 93 | 92 | 1 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 7.2.3.6 Overall Probabilities of Selection

Table 7-8 provides the selection probabilities at each stage of sampling. The conditional probability in column 2 shows the selection probability at a particular stage of sampling, and the overall probability in column 3 shows the multiplicative overall selection probability across all previous stages. Therefore, the overall probability in the last row in table 7-8 is the overall probability of selecting a person into the NAAL or SAAL samples.

Hence, an explicit expression for the overall probability of selecting a person into the national NAAL sample is given by the product of the conditional probabilities from each selection stage (given that the reserve sample was not released):

$$P_{ijkl} = P_{hi} CP_{ij} CP_{ij(chunk)} CP_{ijk} CP_{ijk(mdu)} CP_{ijkl},$$

which, for NAAL, reduces to

$$P_{ijkl} = \frac{\bar{d}}{I_{NAAL}} \frac{mhu_{ijk}}{MHU_{ijk}} \frac{m_{ijk}}{M_{ijk}} \times 0.95$$

and, for SAAL, reduces to

$$P_{ijkl} = \frac{\bar{d}}{I_{SAAL}} \frac{mhu_{ijk}}{MHU_{ijk}} \frac{m_{ijk}}{M_{ijk}} \times 0.95$$

Where,

$mhu_{ijk}$= number of selected missed housing units associated with housing unit $k$, within segment $j$ of PSU $i$,

$MHU_{ijk}$= number of missed housing units found that were associated with housing unit $k$, within segment $j$ of PSU $i$,

$m_{ijk}$ = number of eligible persons selected within housing unit $k$, within segment $j$ of PSU $i$,

$M_{ijk}$ = number of eligible persons within housing unit $k$, within segment $j$ of PSU $i$,

$I_{NAAL}$ is the sampling interval for the selection of the NAAL segments, and

$I_{SAAL}$ is the sampling interval for the selection of the SAAL segments.

## 7.3 PRISON SAMPLE

This section describes the design and selection of the state and federal inmate sample for the 2003 NAAL Prison Study. A two-stage sample was used to select inmates. At the first stage, 114 prisons were selected from the frame, with probabilities proportionate to an MOS. At the second sampling stage, an average of about 12 inmates were selected from the participating sampled facilities.

Section 7.2.1 describes the sampling frames for the study. Procedures used to select the sample of prisons and to select inmates within those facilities are described in sections 7.2.2 and 7.2.3, respectively.

### 7.3.1 Sampling Frames for the Prison Study

The sampling frames for the Prison Study are discussed below, including an overview of the frames and data sources used to create the sampling frames.

#### 7.3.1.1 Overview of Prison Study Sampling Frames

The target population consisted of inmates 16 years and older from state, federal, and private prisons in the United States. The sampling frame was created primarily from two data sources: the Bureau of Justice Statistics 2000 Census of State and Federal Adult Correctional Facilities (referred to in the following text as the Prison Census) and the 2003 Directory of Correctional Facilities of the American Correctional Association (ACA).

The facility universe for the NAAL Prison Study was consistent with the Prison Census. As defined for the Prison Census, the 2003 NAAL target population included the following types of state and federal adult correctional facilities: prisons; prison farms; reception, diagnostic, and classification centers; road camps; forestry and conservation camps; youthful offender facilities (except in California); vocational training facilities; drug and alcohol treatment facilities; and state-operated local detention facilities in Alaska, Connecticut, Delaware, Hawaii, Rhode Island, and Vermont. Facilities were included in the NAAL Prison Study if they were

- staffed with federal, state, local, or private employees;
- designed to house primarily state or federal prisoners;
- physically, functionally, and administratively separate from other facilities; and
- in operation between September 2003 and March 2004.

The Prison Study sample also included private facilities housing prisoners under exclusive contract to state governments and the Federal Bureau of Prisons.

Specifically excluded from the NAAL Prison Study were

- privately operated facilities that were not exclusively for state or federal inmates;
- military facilities;
- Immigration and Naturalization Service facilities;
- Bureau of Indian Affairs facilities;
- facilities operated and administered by local governments, including those housing state prisoners;
- facilities operated by the U.S. Marshals Service, including the Office of the Detention Trustee;
- hospital wings and wards reserved for state prisoners; and
- facilities housing only juvenile offenders.

Even though they contain inmates up to age 21, juvenile facilities were excluded from the NAAL for two reasons: (1) to remain consistent with the facilities listed in the Prison Census and (2) to promote cost efficiency because it would not have been cost effective to visit these facilities to sample the small number of inmates 16 years of age and older.

Inmate sampling frames were created by interviewers at the time they visited the prisons. The frame consisted of all inmates occupying a bed the night before inmate sampling was conducted.

#### 7.3.1.2 Data Sources Used to Create Sampling Frames for the Prison Study

The Bureau of Justice Statistics 2000 Prison Census included more than 1,600 facilities meeting the criteria provided in the previous section. The Prison Census data included facility addresses, capacity, inmate population, and security level, all of which were important information for sampling and data collection.

The 2003 ACA directory contained an updated list of more than 6,000 adult and juvenile state correctional departments, institutions, programs, and probation and parole/aftercare services. The directory also included updated inmate population figures, security level, and gender of the inmates, which were all helpful for sample design purposes.

The Prison Census list of facilities was compared with the ACA directory list to arrive at a sampling frame of prisons eligible for the study. After comparing the ACA and Prison Census information, project statisticians needed clarification for cases with unknown eligibility status. The data collection contractor called each state's department of corrections and the Federal Bureau of Correctional Facilities to verify that the facilities were eligible for the study or to retrieve missing sampling information. Additionally, the number of cases in question was greatly reduced by obtaining information from various corrections-related websites.

Before sample selection, much work was done to separate work camps, annexes, satellites, and boot camps from their main facility. The sources used for this separation were the ACA directory, telephone calls, and websites. Table 7-12 shows the results of the frame creation operation.

The facilities were selected in late 2003, and inmates were selected and assessed in early 2004. The selection procedures are detailed in sections 7.2.2 and 7.2.3.

**Table 7-12. Summary of data sources used to create the prison sampling frame: 2003**

| Source[1] | Count | Percent |
|---|---|---|
| Total | 1,837 | 100.0 |
| | | |
| Census | 1,559 | 84.9 |
| American Correctional Association | 92 | 5.0 |
| Website | 1 | 0.1 |
| | | |
| Separated from main facility on basis of … | | |
| Telephone call | 7 | 0.4 |
| Website | 51 | 2.8 |
| American Correctional Association Directory | 127 | 6.9 |

[1] 2000 Census of State and Federal Adult Correctional Facilities (Census) and 2003 American Correctional Association Directory (ACA).
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 7.3.2 Selection of Prison Sample

The first-stage sampling units (or PSUs) were state or federal adult correctional facilities. The list of prisons was sorted before sample selection, which implicitly stratified the facilities (as described below), resulting in lower sampling variation than would be achieved with a simple random sample of facilities. The prisons were systematically selected from the sampling frame with probabilities proportional to the number of inmates in the facility.

The probability of facility *i* being selected is given by

$$P_i = \frac{a\hat{x}_i}{\sum_{i=1}^{A} \hat{x}_i},$$

where *A* indicates the number of prisons on the sampling frame, *a* indicates the number of prisons selected for the sample, and $\hat{x}_i$ indicates the estimated number of inmates in facility *i* as it appeared on the sampling frame.

To determine the best sort order for the sampling frame, the data collection contractor conducted regressions of three literacy measures (prose, document, and quantitative) from the 1992 NALS, using census region and security level as the independent variables (the only variables on the current sampling frame that were also on the 1992 NALS public use file). The results showed that census region was a significant variable in explaining variation in literacy among inmates. Therefore, the facilities on the frame were ordered by census region first, followed by security level (supermaximum/maximum, medium, minimum, or other), type (federal, state, or private), and the number of inmates in the facility.

The frame was sorted in a serpentine fashion, with census region sorted first, in ascending order. Within the first level of census region, security level was sorted in ascending order. Within the second level of census region, security level was sorted in descending order. The same pattern continued for all of the sort variables. This process resulted in a list of prisons in which like units were adjacent to each other more often than in the traditional sort order.

A sample of 114 prisons was selected, allowing for prison nonresponse or ineligibility (e.g., closed). Table 7-13 provides the distribution of the prison sample by cooperation status and census region.

**Table 7-13. Distribution of selected prisons, by cooperation status and census region: 2003**

| Prison-level characteristic | Total prisons sampled | Cooperation status | | |
|---|---|---|---|---|
| | | Participants | Refusals | Ineligibles |
| Total | 114 | 107 | 3 | 4 |
| | | | | |
| Census region[1] | | | | |
| Northeast | 17 | 16 | 0 | 1 |
| Midwest | 22 | 19 | 3 | 0 |
| South | 50 | 48 | 0 | 2 |
| West | 25 | 24 | 0 | 1 |

[1] Northeast Region = ME, NH, VT, MA, RI, CT, NY, NJ, PA; Midwest Region = OH, IN, IL, MI, WI, MN, IA, MO, ND, SD, NE, KS; South Region = DE, MD, DC, VA, WV, NC, SC, GA, FL, KY, TN, AL, MS, AR, LA, OK, TX; West Region = MT, ID, WY, CO, NM, AZ, UT, NV, WA, OR , CA, AK, HA.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 7.3.3 Selection of Inmates Within Facilities

The second-stage units consisted of inmates selected within a sampled prison. Inmates were selected with a probability inversely proportional to the prison's population size so that the product of the first- and second-stage selection probabilities would be constant. In practice, the number of sampled inmates varied within prisons because of differences between the anticipated and actual sizes of the inmate populations and also because of constraints on the sample size per prison. The sample design was intended to provide a constant overall probability of selection across all inmates.

The conditional probability of inmate $j$ being selected in prison $i$ is given by

$$CP_{ij} = \frac{b_i}{\hat{x}'_i}$$

Where

$b_i$ =the expected number of inmates to be selected in prison $i$ and

$\hat{x}'_i$ = the updated inmate population of prison $i$, obtained through a telephone call to the facility after its selection into the sample.

The expected number of inmates to be selected in prison $i$ was calculated as

$$b_i = \frac{R_i}{k} b ,$$

where

$b$ =12.12, the average inmate sample size inflated for anticipated nonresponse;

$$R_i = \frac{\hat{x}'_i}{\hat{x}_i} \text{ ; and}$$

$$k = \frac{a'}{\sum_{i=1}^{a'} \frac{1}{R_i}}$$

and where

$a'$ =107, the number of participating prisons.

The expression for *k* is the harmonic mean of the $R_i$ values. Note that *k* is equal to 1 if $\hat{x}'_i = \hat{x}_i$ for all sampled prisons (i.e., the number of inmates on the frame is equal to the number of inmates in the prison as determined through telephone contact). If all $\hat{x}'_i > \hat{x}_i$, then *k* will be greater than 1, and $b_i$ will tend to be an average of size *b*.

Substituting the expression of $b_i$ in the formula for the conditional probability, $CP_{ij}$ gives

$$CP_{ij} = \frac{R_i b}{k\hat{x}'_i} = \frac{\hat{x}'_i}{\hat{x}_i} \frac{\sum_{i=1}^{a'} \frac{1}{R_i}}{a'} \frac{b}{\hat{x}'_i} = \frac{b \sum_{i=1}^{a'} \frac{1}{R_i}}{a'\hat{x}_i}.$$

The overall selection probability of an inmate is thus

$$P_{ij} = P_i CP_{ij} = \frac{a\hat{x}_i}{\sum_{i=1}^{A} \hat{x}_i} \frac{b \sum_{i=1}^{a'} \frac{1}{R_i}}{a'\hat{x}_i} = \frac{ab \sum_{i=1}^{a'} \frac{1}{R_i}}{a' \sum_{i=1}^{A} \hat{x}_i}.$$

Note that $P_{ij}$ is constant across all inmates.

An upper bound of 16 inmates was set to constrain the size of the inmate sample per prison. This upper bound was dictated by the practical limitations of interviewing a large number of inmates in each prison. If a prison's expected sample size exceeded the upper bound, it was truncated to the upper bound, and the sample sizes for the other prisons were inflated to yield the total expected inmate sample size. This iterative process was continued until there were no prisons with an expected inmate sample size greater than the upper bound. In addition, a lower bound of 9 inmates was set to justify the cost of traveling to prisons with a small number of interviewers.

Inmates in each prison were selected from a list of inmates occupying a bed the previous night. The interviewers received forms and instructions to follow when sampling inmates from the lists.

The interviewers had a laptop computer preprogrammed with a sampling algorithm. The statisticians assigned both the random number and the sampling interval to the prison before the fieldwork began; these values were preloaded into the sampling algorithm. The facility name, location, security level, type (federal, state, private), and gender composition (male only, female only, mixed) were also loaded on the laptop. The interviewers were required to verify all sampling information because it was also used in the sample weighting process.

The algorithm required that the interviewer enter the number of inmates on the list. After the number of inmates had been entered, the sampling algorithm compared the value with a preloaded acceptable range. In general, the acceptable ranges were within 10 percent of the expected inmate population. If the number of inmates fell outside the acceptable range, a message appeared on the laptop instructing the interviewer to contact the home office to receive a new sampling rate. After the interviewer entered the appropriate sampling rate, the laptop displayed the number of inmates to be sampled and the selected line numbers. The interviewer circled these line numbers on the list, and those inmates were selected.

Tables 7-14 through 7-16 show the background questionnaire and exercise sample counts for the inmate sample by prison and inmate characteristics. Weighted response rates are provided in chapter 11. Another component of the NAAL assessment was the Adult Literacy Supplemental Assessment (ALSA). A person took an ALSA assessment if he or she did not pass the core assessment. There were 29 inmates who took the ALSA assessment.

**Table 7-14. Background questionnaire sample counts by cooperation status and prison characteristics: 2003**

| Prison-level characteristic | Actual total inmates sampled | Complete[1] | Nonresponse |
|---|---|---|---|
| Total | 1,298 | 1,173 | 125 |
| Prison type | | | |
| Federal | 146 | 136 | 10 |
| State/private | 1,152 | 1037 | 115 |
| Security level | | | |
| Supermaximum/ maximum | 311 | 263 | 48 |
| Medium | 583 | 536 | 47 |
| Minimum/other | 404 | 374 | 30 |
| Gender composition of prison | | | |
| Male | 1,206 | 1,087 | 119 |
| Female only /mixed gender composition | 92 | 86 | 6 |
| Census region[2] | | | |
| Northeast | 183 | 159 | 24 |
| Midwest | 237 | 219 | 18 |
| South | 589 | 556 | 33 |
| West | 289 | 239 | 50 |

[1] Completed background questionnaires included cases that were not complete due to language problems and mental disabilities. These cases were considered a "success" in data collection since race/ethnicity, age, and gender were collected, as well as good information (language problem or mental disability) as to their English literacy skills.
[2] Northeast Region = ME, NH, VT, MA, RI, CT, NY, NJ, PA; Midwest Region = OH, IN, IL, MI, WI, MN, IA, MO, ND, SD, NE, KS; South Region = DE, MD, DC, VA, WV, NC, SC, GA, FL, KY, TN, AL, MS, AR, LA, OK, TX; West Region = MT, ID, WY, CO, NM, AZ, UT, NV, WA, OR , CA, AK, HA.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 7-15. Exercise sample counts by cooperation status and prison characteristics: 2003**

| Prison-level characteristic | Total exercises attempted | Complete[1] | Nonresponse |
|---|---|---|---|
| Total | 1,161 | 1,147 | 14 |
| Prison type | | | |
| Federal | 136 | 133 | 3 |
| State/private | 1,025 | 1,014 | 11 |
| Security level | | | |
| Supermaximum/maximum | 259 | 255 | 4 |
| Medium | 531 | 528 | 3 |
| Minimum/other | 371 | 364 | 7 |
| Gender composition of prison | | | |
| Male | 1,076 | 1,062 | 14 |
| Female only/mixed gender composition | 85 | 85 | 0 |
| Census region[2] | | | |
| Northeast | 154 | 152 | 2 |
| Midwest | 218 | 216 | 2 |
| South | 551 | 545 | 6 |
| West | 238 | 234 | 4 |

[1] Includes exercises coded as complete, reading/writing barrier, language problem, mental disability, or physical disability, as well as those coded as partial completes for the following reasons: reading/writing barrier, language problem, mental disability, or physical disability.
[2] Northeast Region = ME, NH, VT, MA, RI, CT, NY, NJ, PA; Midwest Region = OH, IN, IL, MI, WI, MN, IA, MO, ND, SD, NE, KS; South Region = DE, MD, DC, VA, WV, NC, SC, GA, FL, KY, TN, AL, MS, AR, LA, OK, TX; West Region = MT, ID, WY, CO, NM, AZ, UT, NV, WA, OR , CA, AK, HA.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 7-16. Exercise sample counts by cooperation status and background questionnaire variables in the Prison Study, by inmate characteristic: 2003**

| Inmate characteristic | Total exercises attempted | Complete[1] | Nonresponse |
|---|---|---|---|
| Total | 1,161 | 1,147 | 14 |
| Age | | | |
| 16–29 | 388 | 386 | 2 |
| 30–49 | 659 | 647 | 12 |
| 50+ | 114 | 114 | 0 |
| Gender | | | |
| Male | 1,086 | 1,072 | 14 |
| Female | 75 | 75 | 0 |
| Race | | | |
| Hispanic | 223 | 222 | 1 |
| Non-Hispanic Black only | 491 | 482 | 9 |
| Other | 447 | 443 | 4 |
| Education | | | |
| Less than high school | 646 | 638 | 8 |
| High school | 347 | 342 | 5 |
| More than high school | 168 | 167 | 1 |

[1] Includes exercises coded as complete, reading/writing barrier, language problem, mental disability, or physical disability, as well as those coded as partial completes for the following reasons: reading/writing barrier, language problem, mental disability, or physical disability.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

## 7.4 ASSIGNMENT OF BOOKLETS TO RESPONDENTS

The NAAL assessment used 26 types of booklet to measure prose, document, and quantitative scales. The booklet types were assigned randomly to the housing units selected for the NAAL and SAAL samples. Before data collection, the selected housing units were sorted by PSU, segment, and their geographic sequence within their segment. Booklet types 1 through 26 were assigned to each housing unit by picking a random number *l* between 1 and 26 and assigning it to the first housing unit record. Booklet types for the remaining housing units were assigned sequentially *l*+1, *l*+2, …, 26, 1, 2, .., 26, 1, 2, …. The sequential numbers became the value of the booklet type assigned.

Each interviewer carried a spare bundle of booklets. The booklet types within each spare bundle were sorted with a random start between 1 and 26. If two persons were selected in a household, the top booklet from the spare bundle was administered to the second sampled person. Under the assumption of random nonresponse, the booklet type assignment process was expected to result in an equal distribution of booklet types across respondents.

A similar process was developed to randomly assign booklet types to inmates in the Prison Study. Each prison was assigned a random set of booklet types, and the interviewers took the bundle of booklets and distributed them systematically to each sampled inmate. The prison sample booklet assignment process was expected to result in an equal distribution of booklet types across inmates.

*This page is intentionally left blank.*

# CHAPTER 8

# MAIN HOUSEHOLD STUDY DATA COLLECTION AND QUALITY CONTROL

Michelle Amsbary, Martha Berlin, Thomas Krenzke, and Leyla Mohadjer, Westat

## 8.1 INTRODUCTION

To gather information on adults' literacy skills for the National Assessment of Adult Literacy (NAAL), trained staff interviewed a nationally representative sample of about 12,500 adults, aged 16 and older, residing in private households and college dormitories across the United States. Study participants were randomly selected to represent the adult population in the country as a whole. Black and Hispanic adults were sampled at a higher rate than the remainder of the population to ensure reliable estimates of the literacy proficiencies of these groups.

To give states an opportunity to explore the skill levels of their populations, they were invited to participate in a concurrent assessment. Six states elected to participate in the State Assessment of Adult Literacy (SAAL). Approximately 1,000 adults were interviewed in each of the following states: Maryland, Massachusetts, Missouri, New York, and Oklahoma. A larger sample of 1,500 adults participated in Kentucky, because of a request from the state officials that a larger sample be drawn. To allow comparisons of the state and national results, identical instruments were administered to the state and national samples, and the data were gathered at the same time. Unlike the national sample, the SAAL involved no oversampling of Black and Hispanic adults in high-minority areas.

The main household study was conducted from May 2003 through February 2004. Approximately 400 trained interviewers, some of whom were bilingual in English and Spanish, visited households to select and interview adults. Each study participant was asked to spend approximately 90 minutes responding to a series of diverse literacy tasks and answering questions about his or her demographic characteristics, educational background, reading practices, and other areas related to literacy. Also incorporated into the study protocol were a new component to assess oral reading fluency (the Fluency Addition to the NAAL, or FAN) and an alternative assessment to gather as much information as possible about adults with limited English literacy skills (the Adult Literacy Supplemental Assessment, or ALSA).

After completing an interview, the interviewers edited materials for legibility and completeness. The interviewers sent their completed work directly to the data collection contractor's home office, where the hard-copy assessment booklets and electronic questionnaire data were receipted and processed. Data gathered through computer-assisted personal interviewing (CAPI) were edited by trained staff. Assessment booklets were prepared and sent for scoring, and oral module data files were processed and scored. Final edited and scored data were prepared for analysis.

## 8.2 LISTING

The implementation of an area probability design such as the one used in the NAAL requires the development of a list of housing units in each second-stage sampling unit, or area segment. This section describes the procedures used to carry out the address listing operation. For the national sample, the NAAL design involved 100 primary sampling units (PSUs) and 1,959 area segments. The design for the state samples involved 74 PSUs (14 of which overlapped with the NAAL PSUs) and 861 area segments (2 of which overlapped with the NAAL segments). Hence, the total household sample was distributed across 160 unique PSUs and 2,818 unique area segments.

During a 12-week period in the summer of 2002, a total of 332,821 housing units were listed in the 2,818 area segments. A field organization of nearly 200 people was assembled to carry out the listing operation.

### 8.2.1 Staff Organization for Listing

The staff for the NAAL listing operation included 10 regional supervisors and 167 listers. The supervisors reported to one of two field managers, who reported directly to the field director. On average, each supervisor recruited and supervised 17 listers in about 280 segments in his or her region. Listers were recruited in May and June 2002. Of the 167 interviewers hired as listers, 123 had or were currently working for the data collection contractor, 30 had experience with other social research organizations, and 14 were new to social research.

### 8.2.2 Listing Materials

A segment folder was prepared for each sampled segment. The folder contained (1) a tract map; (2) a segment map; (3) listing sheets on which the lister recorded each address; (4) a segment profile form; and (5) a form containing general comments and any special instructions.

The tract and segment maps in each folder defined and described the sample segments, permitting the listers to identify the exact boundaries of the sampled areas. The size of the segments varied substantially, depending on the urban or rural character of the area. In dense residential areas, segments may have consisted of one or more blocks. In rural areas, segments may have covered many miles. The 11” × 17” tract map provided an overall picture of the location of the segment within a larger geographic area and within the county. This map gave the listers a geographic context to help them locate the segment. In some segments, the computer program used to generate maps was unable to produce clear and complete maps for the selected geographic areas. In that situation, an enlarged tract map, with clearer boundaries and street names, was also included.

The 11” × 17” segment map was a more detailed picture, showing all streets and other features of the area to be listed. Occasionally, a section of the segment was too dense (i.e., included too many streets), or there was not enough space to print street names and other descriptive information on the map. An enlarged segment map was provided in this circumstance. Listers recorded two things on the segment map: (1) arrows that indicated their route of travel while listing, and (2) the listing sheet line numbers that corresponded to the first and last housing units on each street or boundary.

The U.S. Census Bureau’s TIGER System file was used to produce the segment maps. The TIGER file digitized all intersections of geographic boundaries used in the 2000 census. This information was used to generate maps of selected blocks, combinations of blocks, or other geographic units.

The listing sheets were used by the lister to record the complete address, including street name, house number, and, if appropriate, apartment number, of every housing unit encountered. In rural areas, where house numbers were not always available, listers described the housing unit’s location in relation to other landmarks.

The Segment Profile Form was used by the lister to collect basic information about the demographic makeup of the segment. The lister also noted the status of new construction or demolition, as well as an estimate of the proportion of the segment used for seasonal dwelling.

The Special Instructions and General Comments Form was used to communicate information to the lister and for the lister to note any special circumstances encountered in the segment.

### 8.2.3 Training Listers

Of the 167 interviewers recruited for the listing operation, 39 had listing experience within the past 3 years. These listers were trained through home study only, using a listing manual, a listing video, and a home study guide. This guide was a reference manual with practice exercises interspersed throughout and a final examination. Trainees were instructed to read the manual and complete the exercises and the final examination from the guide. The completed exercises and examination were then mailed to the appropriate supervisor for review. The remainder of the text was used as a reference manual during the listing operation. The experienced listers who completed the home study session were not required to attend the in-person training.

The 10 supervisors and two field managers attended a 1-day in-person training on the listing procedures. The 128 interviewers who had not conducted listing in the past 3 years attended a 2-day, in-person lister training session that included lectures, an audiovisual training presentation, and field practice. Before attending training, these trainees also completed the home study package described above, and brought the completed exercises and final examination with them for the training staff to review and evaluate. The training program covered the fundamental concepts and basic procedures of listing, problematic aspects of listing, special procedures for working in rural areas, and administrative procedures. Listing procedures unique to NAAL were also presented, including instructions for listing group quarters and other structures that did not qualify as housing units, such as military barracks, hospitals, and transient hotels or motel rooms.

A segment near the training site was selected for field practice listing as part of the 2-day training program. The training staff prepared practice segment folders containing the tract and segment maps, as well as listing sheets. Before the training session, the training staff listed all the addresses in the practice segment, noting any problem areas. During training, trainees were taken to the practice segments and required to list the segment individually. Training staff then reviewed the listings with the trainees, checking the recording and discussing any problems.

### 8.2.4 Listing Operation

The listing operation began immediately after training and was completed by mid-September 2002. During the initial stages of listing, the lister located the assigned segment using the maps from the segment folder and, when necessary, a local map to verify boundaries. Before beginning to record addresses, the lister "cruised" the segment to verify boundaries, make an approximate count of the housing units in the segment, and correct the segment and tract maps, if necessary.

To keep the listing cost-efficient, very large area segments—those containing 300 or more housing units according to the 2000 census count—were subdivided into smaller, more manageable areas, or chunks, according to instructions provided by statistical staff. One chunk was then selected, with probability proportionate to size, as the area to be completely listed. (Chapter 7 provides more detail on chunking procedures.) A total of 334 segments had housing unit counts in excess of 300 and were subdivided in this fashion, using a software designed to manage the listing and chunking effort.

If no major problems were encountered during the cruising stage, the lister began the actual address listing operation, starting in the northwest corner of the segment. The starting point and the direction of travel were indicated on the segment map. As the lister traveled through the segment, following the specified listing route, he or she recorded the address of each housing unit on the listing sheet. If no house or apartment number was evident for a housing unit, the lister recorded a detailed description of the unit and its location.

Because the NAAL design was based on the 2000 census data and the listing operation was carried out approximately 2 years after the field operations for the census, relatively few structural changes had occurred in the segments. Hence, in most segments, the difference between the expected and actual numbers of housing units was not great. For the most part, segment boundaries were also still intact and could be easily located.

Field managers, field supervisors, and home office staff monitored the listing effort by using automated progress reports distributed twice a week. Statisticians also monitored the listing operation for yield compared with census figures.

Completed segment listings were returned to the data collection contractor, where they were reviewed for completeness, accuracy, legibility, and adherence to procedures. A segment tracking application was used to track lister assignments and the status of the segment folders. Segment listings were batched for data entry, and address information was then coded, keyed, and entered into the survey control file.

Of the 2,818 segments selected for listing, two segments yielded no housing units. In the remaining 2,816 segments, nearly 332,820 housing units were listed. Of these, approximately 35,500 housing units were selected to form the national sample (approximately 25,500 housing units) and the state samples (approximately 10,000 housing units across the six participating states).

### 8.2.5 Quality Control Procedures

As described in section 8.2.5.1, quality control checks applied to the listing operation included a thorough review of each lister's initial assignment. Additionally, procedures were implemented to identify and sample housing units and structures missed during the listing operation. These quality control checks were conducted during the data collection phase of the study and are described in section 8.2.5.2.

### 8.2.5.1 Quality Control of Listing Sheets

Each lister was required to mail his or her first two completed segment listings to the supervisor for review before working on additional segments. The supervisor reviewed the listings for completeness, accuracy, legibility, and adherence to procedures and provided immediate feedback to the lister. On the basis of this review of the lister's first assignments, the supervisor decided on the type and number of segments to assign to that lister. Trained home office staff conducted a further review of each completed listing.

### 8.2.5.2 Quality Control of the Listing Operation

As a check on the completeness of the address-listing operation, NAAL interviewers performed two procedures to detect and measure omissions in listing. The hidden housing unit and missed structure procedures were performed during data collection to correct any undercoverage during the listing operation. As the names imply, the procedures separated the detection of missed housing units into two parts—identifying hidden housing units within multiunit structures and detecting completely missed structures or units constructed since the listings were prepared. Section 7.1.3.5 provides the rationale for the hidden housing unit and missed structure procedures, describes how segments and structures were selected for these procedures, and provides the exact number of housing units added through the two procedures. Each of the two procedures is described below.

#### 8.2.5.2.1 Hidden Housing Unit Procedure

The hidden housing unit procedure was implemented in a sample of the selected housing units. It was designed to detect individual units within listed structures that were not visible to the lister. Such units might be in multiunit structures, such as apartment buildings or duplexes, or they might be separate dwelling quarters within what appeared to be a single-family structure, such as a self-contained in-law apartment in the basement.

The assignment label on the front of the household folder indicated whether the housing unit had been selected for the hidden housing unit procedure. For housing units that required the hidden housing unit procedure, the CAPI screener displayed the text to be read to the respondent to determine whether there were any other living quarters at the address, such as a basement or attic apartment. In multiunit structures, the interviewer also compared the numbers on mailboxes and doorbells against the listing sheet and looked around the outside of the structure for additional units or entrances, being particularly careful to look for basement, unnumbered, or out-of-the-way apartments that might be hidden and easy to miss.

If no hidden housing units were discovered, the interviewer simply recorded that the procedure had been carried out. If four or fewer additional units were discovered in any housing unit, they were automatically added to the sample, and the interviewer began efforts to conduct interviews in those households. If five or more hidden housing units were discovered in any unit, a subsampling procedure was used to control the number of additional units added to the sample. When this situation occurred, the interviewer called the supervisor or home office for subsampling instructions.

#### 8.2.5.2.2 Missed Structure Procedure

The missed structure procedure was conducted in a sample of segments. A message on the segment folder instructed the interviewer to perform the procedure before conducting any interviews in that segment. Using the tract and segment maps and the listing sheets, the interviewer recanvassed the entire segment to look for single-family houses or multiunit structures that had been omitted from the listing sheets.

If no missed structures were discovered, the interviewer simply checked a box on the missed structure form to verify that the procedure had been performed. If missed structures were found, the interviewer listed all the newly discovered housing units on the missed structure listing sheet. If five or fewer missed structures were discovered in a segment, they were all automatically added to the sample and the interviewer began efforts to interview in the households. If more than five missed structures were discovered in a segment, interviewers were instructed to call their supervisors to determine whether all or a subsample of units should be added to the sample.

## 8.3 DATA COLLECTION INSTRUMENTS AND INTERVIEWER MATERIALS

The development and content of the cognitive exercise items, the ALSA, and the oral module—the primary data collection instruments used in the NAAL—are described in detail in chapter 2. What follows is a brief overview of the materials used for the assessments, as well as a discussion of the other materials used during data collection. This includes automated instruments and systems such as the screener, the background questionnaire, the assessment interviewer guide, and the interviewer and supervisor management systems, as well as hard-copy materials such as advance materials, the oral module booklet, noninterview report forms, interviewer manuals, and various field aids. Many of these materials and instruments were previously developed for the 2001 NAAL field test and revised as appropriate for the main data collection effort.

### 8.3.1 Assessment Materials

The administration of the core and main assessment, as well as the ALSA, involved the use of many hard-copy materials, in addition to the automated instruments described in the subsequent sections. The development of these materials are discussed in detail in chapter 2, but briefly summarized below.

The data collection effort used 26 unique assessment booklets. Each booklet began with 7 core assessment items, followed by three blocks of main assessment items. Assessment items were arranged into 13 unique blocks and then spiraled to reduce the possibility of question order effects, resulting in 26 booklets.

As in the field test, the administration of the assessment required the use of several stimulus materials, including an almanac, a colon cancer pamphlet, a "Medicare and You" brochure, a newspaper, and a calculator.

The ALSA assessment was produced in a separate booklet and required the use of 9 stimulus materials, including a Coca-Cola can, a box of pancake mix, a yard sale flyer, and a TV guide.

### 8.3.2 CAPI Data Collection Instruments and System Features

The following sections provide a brief description of the individual automated instruments, the screener, background questionnaire, and interviewer guide, as well as the Interviewer and Supervisor Management Systems.

#### 8.3.2.1 Screener

The CAPI screener was used to collect household information and to select one or more members of the household for participation in the background questionnaire and the literacy assessment. The screener began with a household enumeration, in which the interviewer recorded the first name of all household members starting with the person (or one of the persons) who owned or rented the home, designated the reference person. The interviewer then entered each household member's relationship to the reference person, gender, and age. The interviewer asked about race and ethnicity only for household members aged 16 and older. The screener could be administered to any household member aged 16 or older.

To facilitate the validation of screeners and the subsequent followup of the case, respondents were asked to provide a telephone number where he or she could be reached. Finally, if the housing unit had been selected for the hidden housing unit procedure, the interviewer followed the procedures provided at the end of the screener.

The different selection criteria used for the national and state samples were programmed into the screener instrument. The sampling procedure used to select the appropriate background questionnaire and exercise respondent(s) in each household was implemented by the CAPI system; the interviewer had no discretion about whom to include in the sample. The CAPI screener selected one respondent in households having one to three eligible members and two respondents in households having four or more eligible members.

A Spanish version of the screener was administered by a bilingual interviewer in households where the members spoke only Spanish. Additionally, if the household members did not speak either English or Spanish, interviewers were permitted to use a translator, such as a household member under 16 years of age, a neighbor or friend, or a paid assistant.

#### 8.3.2.2 Background Questionnaire

The background questionnaire was an approximately 25-minute instrument that included questions on a variety of topics, including general and language background; educational background and experience; political and social participation; labor force participation; literacy practices; job training and skills; demographic information; family literacy; household income and welfare participation; health questions; and additional demographics.

Demographic information collected during the screener interview, such as age and gender, was directly imported into the background questionnaire. The CAPI program controlled the background questionnaire instrument flow by using answers to prior questions to determine which questions should be asked and which should be skipped for each respondent. Hard and soft edits were also programmed directly into the CAPI program and inconsistencies were reconciled with the respondent during the interview.

#### 8.3.2.3 Interviewer Guide

Each of the 26 versions of the assessment booklet had a corresponding interviewer guide, which contained instructions for facilitating the interview and guiding the respondent through the assessment booklet and cued the interviewer when to read an instruction or hand stimulus materials to the respondent. The interviewer guide also specified the exact amount and type of assistance the interviewer could provide to the respondent during the assessment. The interviewer guide was programmed as a CAPI application. The CAPI system automatically presented the correct version of the interviewer guide, depending on which assessment booklet had been assigned to the case.

The interviewer guide also contained a module for scoring the respondent's responses to the seven core assessment items. On the basis of the responses to these items, as well as the language in which the core items were administered, the CAPI system implemented an algorithm to determine whether the respondent should proceed with the main assessment or be directed to ALSA. (See section 13.3 for more detail on the core scoring procedure.)

#### 8.3.2.4 Interviewer Management System

An interviewer's assigned cases and other study activities were managed with an integrated software system called the Interviewer Management System. This system was developed for the 2001 field test and revised for the main household study. The Interviewer Management System had various features that were accessed by using one of two laptop modes of operation, stand-alone mode or online mode. The online mode features required access to the data collection contractor's centralized database.

The Interviewer Management System stand-alone mode provided the following capabilities:

- case browse, allowing the interviewer to review assignments;
- hidden housing unit processing;
- status review for the case and its individual tasks;
- the ability to launch and conduct all CAPI instruments;
- creation of zip disks of oral module recordings; and
- entry of status codes and other information on an Electronic Record of Calls (EROC).

The Interviewer Management System online mode provided the following capabilities:

- data transmission;
- time and expense reporting;
- shipment of case materials; and
- e-mail.

The five primary functions of the Interviewer Management System were (1) Browse Case, (2) Browse Person, (3) Data Transmission, (4) E-mail, and (5) Ship Disk. Each function is briefly discussed below.

***Browse Case***

The Browse Case window displayed the following information about each of the interviewer's assigned cases:

- case identification (ID) number;
- street address, city, ZIP code, and state;
- overall interview status and status date; and
- assigned assessment booklet type for sample person #1.

The Browse Case window also contained the Activity Log. The Activity Log displayed a history of previous EROC entries. The items included in this display were the following:

- the contact date;
- the contact status or result;
- the contact type (in-person, telephone, or mail);
- who was contacted;
- comments about the contact; and
- the appointment date, if applicable.

The EROC allowed the interviewer to enter interim status codes only. Finalized statuses were entered by the supervisor, using the Supervisor Management System application (see section 8.3.2.5), or by the CAPI system automatically upon completion of each interview task.

The Browse Case window also provided the functionality for updating addresses and completing the hidden housing unit procedure. If an address correction was needed, as determined during the study introduction that preceded the screener, the interviewer entered revised address information into the Interviewer Management System. This action provided information for updating the home office files to be used in conducting validation interviews or sending refusal conversion letters. The hidden housing unit button was available only for the cases that had been selected for the procedure. Interviewers entered the number of housing units to be added, as well as the street address for these cases. The Interviewer Management System then added the new cases to the interviewer's Browse Case assignment.

***Browse Person***

When a sample person was identified during screening, sample person-level data were added to the system. For completed screeners in which one or two sample persons were selected, Browse Person was used to view sample-person-level information, such as sample person ID number, name, age, gender, telephone number, sample-person-level status, status date, and assigned assessment booklet type. Browse Person functionality was also used to complete the sample person-level tasks, such as the background questionnaire, the assessment, and the oral module. The Interviewer Management System controlled the order of administration of the instruments by requiring that the background questionnaire be finalized before the interviewer guide associated with the assessment became available.

***Data Transmission***

The Data Transmit button was used by the interviewers to enter their time and expense data, transmit case data, and record shipments of cases to the data collection contractor. The data transmission and shipment functions are discussed in more detail below.

Data transmission involved sending electronic information from the interviewer's laptop to the home office and sending new study data to the interviewer's laptop. Items that were sent from the home office to the interviewer's laptop included new or transferred assignments from the supervisor and case status code updates entered by the supervisor. Items that were sent from the interviewer's laptop to the home office included EROC entries, interview data for completed or partially worked cases, and data concerning finalized cases that were ready to be shipped to the home office.

As part of the preparation for shipping finalized case materials, interviewers indicated in the Interviewer Management System which cases would be included in each shipment, as well as the tracking number used for the mailing.

***E-mail***

Outlook Express was used to send and receive e-mail. Interviewers were in frequent e-mail communication with their supervisors and other interviewers.

***Ship Disk***

The Ship Disk process transferred the voice recordings collected during the oral module from the laptop to a zip disk. Interviewers were given a zip drive to assist in this process.

#### 8.3.2.5 Supervisor Management System

The Supervisor Management System, a component of the Data Management System, was designed for the 2001 field test and revised as necessary for the main study. Supervisors used the Supervisor Management System to manage the case work within their region. The numerous functions of the Supervisor Management System are described below:

- case review, assignment, reassignment, and unassignment to interviewers;
- case search using the following criteria: PSU, case ID, interviewer, status, and status date;
- assignment of the final status of cases, at both the household and sample person levels;
- review of time and expense data that were recorded by interviewers;
- report production (see section 8.7.2. for a detailed discussion of the reports); and
- tracking of the distribution and use of the assessment and ALSA booklets.

### 8.3.3 Interviewer Materials

The administration of the NAAL interview required numerous hard-copy interviewer materials, including several advance materials, the household folder, handcards, noninterview report forms, interviewer manuals, and the oral module booklet. These materials are discussed in the following sections.

#### 8.3.3.1 Advance Materials

During preparation for the 2001 field test, considerable effort was made to develop introductory materials that would convince respondents of the study's legitimacy and importance. These materials were revised slightly for the main study.

**Advance letter.** Before the interviewer's first contact with the household, the home office mailed an advance letter with a brochure. The letter introduced the study, identified the sponsor, stated the study's purpose, and asked for cooperation. It also provided the toll-free study hotline number, as well as the name and telephone number of the National Center for Education Statistics (NCES) Project Officer.

**Brochure.** The informative and attractive brochure included with the letter explained the study in detail and emphasized the importance of participation.

**U.S. Department of Education letter of introduction.** This letter was signed by the NCES Project Officer and verified that the interviewer was an authorized representative of the U.S. Department of Education. The letter was used whenever respondents needed further convincing that the study and interviewer were legitimate.

**Community authorization letter.** This general letter was intended to be displayed to apartment managers, postal employees, police departments, or other professional people whom interviewers might encounter in the community. It provided assurances that the interviewer was not selling or soliciting but was a trained professional working on a government-sponsored education study.

**Sorry-I- missed-you card.** This card was left when the interviewer visited a household and no one was home. Interviewers often personalized the card with a brief message or left their name and telephone number on the card; this provided some familiarity and recognition when the interviewer returned to the household.

**Refusal conversion letters.** Several versions of refusal conversion letters were developed and sent to households that refused to participate in the study. See section 8.8 for a list of these letters.

Spanish-language versions of the advance letter, brochure, U.S. Department of Education letter of introduction, and sorry-I-missed-you-card were produced and distributed as well.

#### 8.3.3.2 Household Folder

One household call record folder was produced for each sampled housing unit in the study. The household folder helped the interviewers keep track of the status of all cases in their assignment.

A label on the cover specified the case ID, the address of the housing unit, the listing sheet line number associated with the housing unit, the barcode number of the assessment booklet preassigned to the case, and the case control code. It also indicated whether the housing unit had been selected for the hidden procedure.

The front of the folder also contained a study introduction, for convenient access at the doorstep, and included an address verification question. The Spanish translation of the study introduction was located on the inside front cover. Step-by-step instructions for performing the hidden housing unit procedure were also included on the inside front cover, as well as space for recording any necessary notes.

The Record of Actions was located on the back cover of the household folder. Interviewers used the Record of Actions grid to record the status and outcome of every contact attempt with the household. For each contact, the interviewer recorded the following information:

- date, day of week, and time;
- contact type (in-person, telephone, or mail);
- result code of appropriate instrument, for both sample person #1 and sample person #2; and
- any applicable comments.

The interviewer stored the case-specific assessment booklet in the household folder, as well as any noninterview report forms, oral module interviewer guides, ALSA booklets, or assessment booklets used with a second sample person, if required for that case. This entire package was returned to the home office when the case was completed. If a case was transferred to another interviewer, the household folder was transferred as well.

#### 8.3.3.3 Handcards

A bound handcard booklet was designed to facilitate the flow of the screener and background questionnaire interviews and improve the efficiency of the sample person’s reporting because the response options would not have to be read aloud for every question. Two handcards listing race and ethnicity response options were used for the screener. The background questionnaire used 10 handcards listing the sets of response categories used most frequently throughout the interview. The handcards were translated into Spanish, with the translation bound to the flip side of the handcard booklet. Handcard instructions, indicating which handcard was to be used for each item, appeared at the top of the applicable CAPI screen.

#### 8.3.3.4 Noninterview Report Forms

When a household member did not complete the screener, or when a sampled respondent did not complete the background questionnaire, the assessment booklet or ALSA, or the oral module, the interviewer was required to complete a noninterview report form. One version of the form was used to document nonresponse at the screener level; another version was used for nonresponse resulting at the background questionnaire, assessment, ALSA, or oral module level. The information collected on these forms served two important purposes: (1) field supervisors reviewed the forms to determine the case’s potential for nonresponse conversion and (2) the data collected on the form were processed for nonresponse analysis.

The screener noninterview report form was completed if the sampled address was determined to be vacant or not a housing unit, or if the interviewer was unable to complete a screener at that address. In the latter case, the interviewer provided information about attempts to contact the household and the reason for noncompletion. If a household member refused to participate, the interviewer described the reasons, in the respondent’s own words, as well as the perceived strength of the refusal. The interviewer also provided any other information that might help another interviewer contact the household, complete the screener, or both, such as what language was spoken in the household and whether a refusal letter would be helpful.

The sample person noninterview report form collected similar information for the background questionnaire, the ALSA, the assessment, and the oral module, including the specific reasons the respondent did not complete the particular instrument(s) and additional information, such as the type of disability, the non-English language spoken, details about reading and writing difficulty, or circumstances surrounding the refusal.

#### 8.3.3.5 Interviewer Manual

The study-specific interviewer manual included an introduction to the study and an overview of interviewer responsibilities. The text covered field materials and procedures for locating sampled households; contacting respondents; and administering the screener, the background questionnaire, the assessment booklet, the ALSA, and the oral module. Question-by-question specifications for the screener, the background questionnaire, the interviewer guide, and noninterview report forms were included as well. The interviewer manual also contained information on maintaining quality control procedures, using the Interviewer Management System features, keeping records, completing the Time and Expense Report, and reporting to the supervisor. A detailed table of contents and section markers helped the interviewer locate specific information in the manual. Interviewers received the manual at training.

#### 8.3.3.6 Oral Module Booklet

The oral module booklet contained the tasks that the respondent was required to read aloud during the oral module component of the interview. As described in section 5.3.4.4, respondents wore a headset with a microphone that recorded samples of their reading and answers to questions directly onto the interviewer's computer. The oral module booklet contained eight unique reading passages, spiraled to create 16 different sets consisting of two passages each. The CAPI oral module instrument instructed the interviewer which set of passages to administer to each respondent.

The reading passages were followed by one list of digits, one list of letters, one list of English words (subdivided into three sections that got progressively harder), and one list of English pseudowords (subdivided into three sections that got progressively harder). These were presented in the same order for all respondents. (See section X.X for a detailed description of the tasks included in the Oral Module Booklet.)

## 8.4 FIELD STAFF TRAINING

The following sections describe the training of the field supervisors and interviewers. The training plan adopted included one large session, with the interviewers divided into 20 small groups according to supervisory region. The field supervisor served as the assistant trainer in each room and was paired with a lead trainer. NCES and various contractor and subcontractor staff members monitored the sessions.

A challenge of the training plan was to prepare field staff to both conduct traditional interviews and administer literacy assessments. Interviewers were trained to take an active role in conducting the screener and the background questionnaire and to be prepared to answer any questions the respondent might raise. In the role of assessment administrator, however, the interviewers had to remain very much in the background, observing and facilitating but intervening only at certain well-defined points and refraining from offering help in completing the literacy tasks even if it was requested.

### 8.4.1 Approach to Training

The basic approach to interviewer training was to maximize trainees' involvement and participation in the training, to provide ample opportunity for supervisory staff to observe and evaluate trainee performance, and to provide trainees with detailed reference documents.

Each training room had a lead trainer and an assistant trainer responsible for approximately 15 to 18 trainees. The training staff was composed of members of the data collection contractor's staff. Lead trainers consisted of home office staff, the four NAAL field managers, other employees with training experience, and supervisors from other data collection projects. The 20 regional supervisors served as the assistant trainers.

### 8.4.2 Train-the-Trainer Trainings

Approximately 50 lead and assistant trainers were trained on the NAAL training program during a training session in late March 2003. The 4-day training was a simulation of the interviewer training program (described in detail in section 8.4.5), although the pace was accelerated because of the experience level of the group. This simulation of the interviewer training program not only prepared the supervisors for their subsequent responsibilities but also provided a dress rehearsal for training staff and an opportunity to evaluate and refine the training materials. After completing the training, lead and assistant trainers were prepared to lead small groups of interviewers through scripted, interactive reviews of the NAAL data collection instruments.

Following the lead and assistant trainer training, a separate training session was conducted in April 2003 for the staff who would be supporting each training room's software and hardware needs, referred to as runners and data display operators. Given the reliance on CAPI instruments throughout the interviewer training, it was necessary to train more than 50 people on how to navigate and resolve problems with the NAAL CAPI applications. An abbreviated simulation of the interviewer training was conducted over a 2-day session, focusing exclusively on the training sessions that involved CAPI applications and the use of the computer.

### 8.4.3 Supervisor Training

A 1½-day training was held for the field managers and supervisors immediately following the train-the-trainers training session. Supervisor training was conducted by the NAAL field director, with support from NAAL systems staff on the technical aspects of the Supervisor Management System. The training covered management techniques and duties specific to the NAAL data collection. The supervisor manual was distributed at the training.

After completing their training, field managers and supervisors returned home to prepare to train interviewers and assume supervisory responsibility immediately after interviewer training.

### 8.4.4 General Interviewer Techniques

Novice interviewers received 4 hours of in-person training on General Interviewer Techniques (GIT) prior to project-specific training. The in-person GIT training program included an audiovisual presentation, interactive participation, written exercises, and a question-and-answer period. The training introduced the interviewers to survey research; provided examples of survey questions, recording conventions, and interviewing terminology; and taught them basic listening and probing skills for obtaining accurate data. The interviewers completed exercises on applying probing techniques and answering respondent questions. The importance of data quality was also reviewed. This training was in addition to the GIT home study and exercises, required of all interviewers.

CAPI Train, a self-administered tutorial that introduced the procedures for conducting a CAPI interview, was completed by most interviewers, as part of their home study package before training. The tutorial instructed trainees on types of questions, function keys, and special commands. The training also included practice in logging on to the computer and using the keyboard, particularly the function keys used to manage the flow of the instruments.

For the handful of interviewers who were hired late in the recruitment process and did not receive their laptops prior to training, a separate 2-hour CAPI Train session was held after GIT training. Interviewers completed the self-guided tutorial at their own pace. The session was supervised by NAAL systems staff.

### 8.4.5 Project-Specific Training

NAAL project-specific training for the 342 interviewers consisted of a 6-day, in-person training program, preceded by home study. The training was conducted in late April 2003. Twenty groups of interviewers were trained concurrently. Supervisors (assistant trainers) were assigned to the training room with the interviewers from their region. Holding numerous simultaneous training sessions at one site allowed the NAAL field director, data collection contractor home office staff, and NCES staff to observe all training sessions while maintaining a manageable number of interviewers in each training room.

Because of interviewer attrition, 59 interviewers from across the country were recruited and trained at a session held in August 2003. Two additional small attrition trainings for a total of 10 interviewers were held in fall 2003. These training programs, led by NAAL project staff, were identical to the program used at the initial interviewer training session.

#### 8.4.5.1 Interviewer Training Materials

The training materials were carefully scripted to cover every concept that the interviewers needed to know, and the scripts were organized into training guides. The elaborate preparation of training materials accomplished two purposes. First, it achieved standardization, which is particularly important when a large staff of interviewers is being trained in separate sessions. Second, it allowed all trainers to study the training guides, rehearse their roles, and be completely prepared for training. This was particularly important in a training effort that required a large training staff. The scripted materials eliminated the necessity for the trainer to improvise. This preparation allowed the NAAL training sessions to move smoothly and on schedule, which gave the interviewers the confidence that they were being trained by knowledgeable people.

#### 8.4.5.2 Interviewer Training Techniques

The general approach to training centered on five basic training techniques that have been extensively used and refined by survey operations professionals over the past 30 years. The following paragraphs briefly describe the five techniques and how they were used for training on NAAL.

**Home study.** About 2 weeks before training, interviewers received a home study package and their laptop computer. The home study included an overview of NAAL, instructions on how to set up and test the laptop computers in the interviewers' homes, an e-mail tutorial, the CAPI Train tutorial and exercises, and the GIT home study guide and exercises. Completed exercises were collected and reviewed by training staff at the in-person training.

**Demonstration.** The first session at training was a videotaped demonstration of the entire NAAL interview. The demonstration interview introduced the trainees to the NAAL instruments and gave them an idea of their role and responsibilities on the project. The trainees were able to see the overall flow of the interview before receiving instruction on each instrument.

**Interactive lecture.** This technique provided trainees with detailed instructions for administering the questionnaires. The lead trainer used a scripted lecture to present the basic concepts of the instrument to the entire group of trainees. Trainees took turns playing the role of interviewer and asking the questions, while the lead trainer provided responses from the script. The lead trainer's script included instructions to interrupt the script at appropriate times to review certain sections of the interviewer manual, point out some of the less obvious features of the instrument, or explain certain terms. All trainees were required to follow along on their computers and enter the responses provided by the trainer. Several runners assigned to the room ensured that trainees entered the correct responses and were all on the correct item. A response was entered into a laptop by a trained data display operator, and then projected on a screen in front of the group. Trainees were instructed to check their entry against the entry on the screen. Interactive lectures were used for the initial presentations of the screener, background questionnaire, interviewer guide, ALSA, and oral module. The scripts used for the interactive lectures presented increasingly complex scenarios so that trainees became familiar with the various types of cases they would encounter.

**Practice exercises.** Written exercises reinforced and tested trainees' comprehension of certain concepts. They were particularly well-suited for evaluating the trainees' comprehension of some of the more complicated instrument issues, such as navigating the screener enumeration grid, scoring the core exercise items, and collecting industry and occupation information in the background questionnaire.

**Dyad role playing.** Role playing provided additional practice and gave trainees a feeling for the overall flow of the interview. Trainees were arranged in pairs (dyads), as designated by the training team. One member of each pair was given a scripted copy of the interview instruments, complete with data entry instructions, and played the role of the respondent while the other trainee conducted the interview (played the role of the interviewer). With the next script, the members of the pair reversed roles. Two role playing scripts were used. The scripts began with the screener and ended with the administration of the oral module.

Additionally, paid respondents recruited by a local focus group facility were brought in toward the end of the training session. Interviewing these respondents gave the trainees the opportunity to conduct a real nonscripted interview before working their first case assignment. Training staff observed these practice interviews and provided feedback after the interview.

#### 8.4.5.3 In-Person Training Program for Interviewers

Most of the 6-day interviewer training was devoted to teaching procedures for administering the data collection instruments—screener, background questionnaire, assessment booklet and interviewer guide, ALSA, and oral module. In addition, instruction was provided on gaining respondent cooperation, locating households, using the Interviewer Management System, assigning status codes, and completing administrative forms. Table 8-1 presents an overview of the training program.

Training interviewers to administer the exercises presented a particular challenge. The role of the assessment administrator was different in important ways from that of an interviewer, requiring interviewers to switch roles in the middle of the interview. During the administration of the screener and background questionnaire, a dynamic interaction took place between the interviewer and the respondent. Although the interviewer needed to remain neutral and avoid leading the respondent, he or she provided reassurance and encouragement. The administration of the assessment exercises, in contrast, required the interviewer to take a much more passive role, observing the respondent's performance without intervening (except as directed in the interviewer guide) and studiously avoiding any temptation to provide assistance, even when help was requested.

**Table 8-1. Overview of the NAAL study-specific interviewer training session: 2003**

| Day | Topic | Presentation mode |
|---|---|---|
| 1 | Introduction and overview of the study | Plenary session |
| | Demonstration of the NAAL interview | Plenary session |
| | Advance materials | Learning community |
| | Procedures for gaining respondent cooperation | Learning community |
| | Locating households | Learning community |
| | Use of Interviewer Management System | Learning community |
| 2 | Screener | Interactive lectures |
| | Screener enumeration grid practice | Individual exercise |
| | Background questionnaire | Interactive lectures |
| 3 | Background questionnaire | Interactive lectures and individual exercises |
| | Assessment booklet and interviewer guide | Interactive lectures |
| | Core scoring procedures | Interactive lectures and individual exercises |
| | Administrative procedures | Interactive lectures |
| 4 | Oral module | Interactive lectures |
| | Screener, background questionnaire, core and main assessment, and oral module | Dyad role playing |
| 5 | Day in an interviewer's life | Interactive lectures |
| | Gaining respondent cooperation | Round table discussion |
| | Live respondent practice | One-on-one |
| | Meeting with supervisors | Learning community |
| 6 | ALSA | Video, interactive lectures, dyad role playing |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

The issues involved in making the switch from interviewer to assessment administrator were discussed in a lecture conducted by a lead trainer. In addition to instructing interviewers on the mechanics of administering the assessment, the trainer spent time discussing problems that might occur in the administration of the assessment, such as those arising for respondents with limited English-reading abilities, respondents with physical or mental conditions that might affect their performance on the assessment, and the special needs of the elderly population. The trainees then practiced administering the assessment, received instruction on core scoring procedures, and reviewed recordkeeping procedures specific to the assessment booklets.

In general, interviewers were discouraged from making assumptions about a respondent's capacity or ability to complete a study instrument; they were instructed to allow the respondent to attempt the questionnaire and the assessments as long as the respondent was willing. In addition, the training materials emphasized the fact that disabilities that might prevent the respondent from completing one instrument might not prevent him or her from completing another instrument. For example, although a blind respondent would be unable to complete the assessment booklet, he or she would not be prevented from responding to the background questionnaire. Conversely, although a respondent who had insufficient English skills might be prevented from responding to the background questionnaire, he or she might be able to attempt the assessment booklet. Although the CAPI system required that the background questionnaire be completed before the assessment was administered, a technique was developed for situations such as these, to allow interviewers to complete the background questionnaire and attempt the assessment.

During training, evening meetings were held with all lead and assistant trainers, other data collection contractor home office staff, and field management staff to discuss any problems that had arisen during the day. Minor modifications to the training program or schedule were discussed as a group. Any necessary changes or clarifications to the materials were made and distributed to the training staff.

Trainees with potential performance problems were identified and remedial measures were discussed at these nightly meetings. Such trainees were closely observed and were paired during dyad role plays with a staff member who could assist them during the mock interview. They were also required to attend evening practice sessions at which they could focus on the element of the interview causing the most problems, such as navigating the screener enumeration grid or using the Interviewer Management System. One-on-one conversations were held with these trainees about their progress. At the end of training, eight trainees who were unable to master the procedures and techniques required for the job were released from the study.

The trainee group included 60 Spanish-speaking interviewers. Following the study-specific training, these interviewers were assembled into three training communities, which were led by a Spanish-speaking field manager, supervisor, and home office staff member. During this 6-hour training session, interviewers worked with the Spanish translation of the screener, the background questionnaire, the core assessment items, the oral module, the ALSA, and advance materials. This gave the supervisors the opportunity to assess the Spanish-speaking abilities of the bilingual interviewers.

## 8.5 CONDUCT OF THE FIELD WORK

The NAAL field period began in May 2003, immediately following the completion of the interviewer training session, and lasted for approximately 40 weeks, until the beginning of February 2004. The national and state samples were worked simultaneously.

The following sections describe the field operations, including the general approach, the schedule and production, and the reporting systems used to manage the field effort.

### 8.5.1 Field Organization

The national and state study components were carried out by a large field organization, headed by the NAAL field director, who reported directly to the data collection contractor project director and was supported by 4 field managers and 20 field supervisors across the United States. The supervisors oversaw an interviewing staff of approximately 350 interviewers. This section presents a general description of the field organization and the responsibilities of the staff at each level.

#### 8.5.1.1 Recruiting Field Staff

Field staff were recruited and hired directly, not through interviewing services. Interviewers were hired from the areas in which the interviewing assignments were located. The primary source of potential field staff was the data collection contractor's computerized field personnel file, containing information on more than 12,000 persons who have worked on their field studies in the previous 4 to 5 years. The file excludes those who received an unsatisfactory review during a previous assignment. This system can produce lists, by geographic area, of available field personnel who meet the qualifications for a project. The system contains demographic information on languages spoken, special field skills, and time and geographic availability. Project evaluations are also in the system, including productivity, accuracy, cooperation, dependability, and length of service for each project. The four field managers and 20 regional supervisors were hired in late 2002 and early 2003.

When recruiting interviewers, supervisors and field managers assessed both the basic skills and personal traits of applicants. An interviewer had to have basic reading and computational skills and be able to follow instructions. Computer skills were desirable as well. Desirable personality traits included receptivity to others' ideas, open-mindedness, and motivation. Additionally, a respondent's willingness to grant an interview often depends on his or her initial perception of the interviewer. Candidates who would appear "neutral" to the target population were favored.

Of the 342 interviewers recruited, 334 successfully completed training and began work on the study. The field organization included approximately 60 interviewers who were bilingual in Spanish and English.

#### 8.5.1.2 Field Staff Attrition and Replacement

The level of interviewer attrition was greater than that experienced for the 1992 NALS, but was consistent with that found on recent field studies of similar size and length of field period. The leading causes of attrition were poor performance, such as inadequate interviewer production rates or a high number of hours per completed case; personal or family illness; and acceptance of a full-time position elsewhere. A total of 86 interviewers were released throughout the field period, and another 110 resigned, for a total attrition rate of 47 percent. Different approaches were used to deal with attrition problems depending on when and where they occurred. In some cases, new interviewers were hired and trained. In other cases, other interviewers working in the PSU were able to complete the remaining work, or interviewers from other areas traveled to the PSUs where the attrition occurred.

To compensate for attrition and slow production in a few areas, an additional 63 interviewers were recruited in July 2003; of these, 59 completed training. Two small-scale attrition trainings were conducted in September and November 2003 as well.

The characteristics of field workers, including all recruited interviewers, supervisors, field managers, and the field director, are shown in table 8-2. Nearly half (224, or 46 percent) had worked previously for the data collection contractor, and a similar number (216, or 45 percent) had worked as interviewers for other field organizations. The field staff was primarily middle aged, with most (322, or 67 percent) between the ages of 30 and 59, only a small number (13, or 3 percent) under 30, and slightly fewer than one-third of them (148, or 31 percent) aged 60 or older. Like most field staffs, the majority (380, or 79 percent) were female.

**Table 8-2. Characteristics of field workers on NAAL main study, by gender, age, experience, and assignments: 2003**

| Characteristics | Total persons | Female | Male |
|---|---|---|---|
| Total persons | 483 | 380 | 103 |
| | Percent | Percent | Percent |
| Age (years) | | | |
| Total | 100.0 | 100.0 | 100.0 |
| 20–29 | 2.7 | 2.6 | 2.9 |
| 30–39 | 11.0 | 11.1 | 10.7 |
| 40–49 | 22.3 | 22.1 | 23.3 |
| 50–59 | 34.5 | 29.1 | 33.3 |
| 60–69 | 22.4 | 25.2 | 23.0 |
| 70–79 | 6.8 | 7.8 | 7.0 |
| 80+ | 0.5 | 1.0 | 0.6 |
| Survey research experience with the data collection contractor (years since hired) | | | |
| Total | 100.0 | 100.0 | 100.0 |
| ≤ 12 months | 45.0 | 51.4 | 46.4 |
| 1–4 years | 34.5 | 40.0 | 34.6 |
| 5–9 years | 11.1 | 4.9 | 9.7 |
| 10–14 years | 6.1 | 7.8 | 6.4 |
| 15–19 years | 2.9 | 1.0 | 2.5 |
| 20–24 years | 0.5 | 0 | 0.4 |
| Field assignments worked | | | |
| Total | 100.0 | 100.0 | 100.0 |
| 1–4 | 80.5 | 83.5 | 81.1 |
| 5–9 | 13.7 | 11.7 | 13.3 |
| 10–14 | 2.6 | 3.9 | 2.9 |
| 15–19 | 1.6 | 0 | 1.2 |
| 20–24 | 0.5 | 1.0 | 0.6 |
| 25–29 | 0.3 | 0 | 0.2 |
| 30+ | 0.8 | 0 | 0.6 |

See notes at end of table.

**Table 8-2. Characteristics of field workers on NAAL main study, by gender, age, experience, and assignments: 2003—Continued**

| Characteristic | Total percent | Female (percent) | Male (percent) |
|---|---|---|---|
| Survey research experience: other companies | | | |
| Total | 100.0 | 100.0 | 100.0 |
| None | 41.1 | 58.3 | 44,7 |
| 1–4 years | 39.7 | 31.1 | 37.9 |
| 5–9 years | 8.4 | 5.8 | 7.9 |
| 10–14 years | 4.2 | 1.9 | 3.7 |
| 15–19 years | 3.4 | 1.0 | 2.9 |
| 20–24 years | 1.6 | 0 | 1.2 |
| 25+ years | 1.2 | 1.9 | 1.7 |
| Highest level of education | | | |
| Total | 100.0 | 100.0 | 100.0 |
| Unknown/no data | 0.3 | 0 | 0.2 |
| Some high school | 0 | 1.9 | 0.4 |
| High school | 15.3 | 3.9 | 12.8 |
| Vocational certificate | 7.9 | 10.7 | 8.1 |
| Some college | 40.5 | 31.1 | 36.8 |
| Bachelor’s degree | 21.0 | 20.4 | 20.0 |
| Graduate work | 3.4 | 9.7 | 14.5 |
| Graduate degree | 11.6 | 22.3 | 13.2 |

NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 8.5.2 Field Management

The management of the NAAL data collection effort included the project director, the field director, field managers, and regional supervisors. The following sections describe the reporting structure of this organization, as well as the procedures and tools used to assist in the reporting.

#### 8.5.2.1 Reporting Structure

The data collection contractor home office staff who oversaw the NAAL field organization included the project director and several supporting staff members. The field director and regional supervisors were located in the field. The field director coordinated all activities related to field operations and kept in close touch with the four field managers to address issues of production, cost, response rates, shipment of closed-out work, and other issues.

For purposes of field operations, the 160 NAAL PSUs were divided into 20 regions, each headed by a supervisor who typically lived in the region. Each field manager had responsibility for five regions and five regional supervisors. The field supervisor's primary responsibility was overseeing the work of an average of 15 to 18 interviewers in his or her region.

An important part of the supervisor's job was determining the optimal flow of work to each interviewer. On the basis of the weekly conference call, the supervisor decided when the interviewer was ready for an additional assignment. Supervisors tried to maintain a balance between somewhat competing goals—keeping interviewers supplied with enough work to stay productive and not allowing cases to languish by giving an interviewer more work than he or she could close out in 2 or 3 weeks.

#### 8.5.2.2 Reporting Procedures and Tools

The smooth progress of field work depended on the ongoing monitoring of the interviewers' work and regular communication among all members of the NAAL project staff and NCES. The following sections describe the major mechanisms and procedures used for reporting during the NAAL field period.

Interviewers were required to contact their supervisors by telephone at a regularly scheduled time once a week to discuss all aspects of their work (response rates, production and cost performance, and quality control results). Each outstanding case in the interviewer's assignment was reviewed and discussed. The supervisor and interviewer discussed any problems reflected in the Supervisor Management System-generated data collection reports, such as low response rates or a high number of hours per complete case. (See section 8.7.2 for a discussion of these Supervisor Management System reports.)

At least once a week, each supervisor had a telephone conference with his or her field manager to discuss progress in the region. Discussion centered on the week's Supervisor Management System reports as well as on current progress as reported to the supervisor during the interviewers' weekly calls. The weekly conferences between field managers and supervisors were used to discuss problems in the region, the prospects and plans for completing the remaining work, and what help, if any, the supervisor needed in order to complete all work in the region by the end of the field period. The results of quality control procedures were also discussed. If the quality control reports indicated problems with the quality of an interviewer's work, appropriate steps to correct the problem were discussed.

Additionally, the field director conducted a formal weekly telephone call with the field managers. The discussion began with a review of the Supervisor Management System reports. These reports were used during the weekly calls to identify any problems in an area, discuss plans for completing the remaining work, and determine whether any help was needed to meet production goals.

Once a week, a home office staff meeting was held with the project director, the field director, and other home office staff members. The results for each region were reviewed, and any studywide problems were reviewed (e.g., interviewer attrition, Interviewer Management System software or hardware issues, or distribution of supplies and materials). Strategies for solving problems were discussed and passed on to the field managers and other staff for implementation.

Twice a month, a summary of data collection progress was sent to the NCES Project Officer. Finally, once a month, NCES and the NAAL contractors attended a meeting at which field progress and problems uncovered during the review of work were discussed. Any important changes in the field work strategy were discussed before implementation.

## 8.6 DATA COLLECTION OPERATIONS

The following sections describe the general approach to the NAAL data collection operation, the specific schedule and plan for production, and the procedure used to effectively administer the interview.

### 8.6.1 General Approach

The NAAL field effort used an approach that has been effective for many previous surveys involving large, complex, in-person data collection operations, including the 1992 National Adult Literacy Survey. Under this approach, the field effort occurs in three overlapping stages:

- Initial phase. Each area segment[1] is assigned by the regional supervisor to an interviewer, who follows certain rules in making a prescribed number of calls to every sampled housing unit in the segment.

- Reassignment phase. Cases that do not result in completed interviews during the initial phase are reviewed by the regional supervisor, and a subset is selected for reassignment to another interviewer in the same or a nearby PSU.

- Special nonresponse conversion phase. The home office assembles a special traveling team of the most experienced or productive interviewers to perform a nonresponse conversion effort, under the supervision of a subset of the field supervisors.

[1] Census blocks or groups of blocks within PSUs.

The assignments in the initial phase are controlled by the regional field supervisor. In NAAL, the supervisors had two local interviewers available in most PSUs. Each area segment was assigned to one of the interviewers on the basis of the racial/ethnic composition of the area and the proximity of the segment to the interviewer's home.

During the initial phase, interviewers were instructed to make up to four in-person calls to the household to complete a screener and up to four additional in-person contacts—after completing the screener—to administer the background questionnaire, the assessment, and the oral module, which had to be completed during the same visit. After the prescribed number of in-person attempts to complete the instrument without contacting a respondent, the interviewer consulted with the supervisor to determine further attempt and contact strategies.

To maximize the chances of finding respondents at home, most contacts were made during prime interviewing hours (3 p.m. to 9 p.m. on weekdays and 10 a.m. to 9 p.m. on Saturdays and Sundays). Contacts at each housing unit were to be staggered on different days of the week and at different times of the day. All calls to complete the screener had to be made in person. If the screener was completed and the background questionnaire, the assessment, and the oral module could not be completed in the same visit, the interviewer was permitted to use the telephone to set an appointment to administer these instruments in person. The initial phase was considered complete when the interviewer reported a definitive outcome for the case or when the full complement of calls had been made.

Interviewers mailed completed cases to the home office twice a week. Copies of the entire segment folder and its associated materials were returned when the interviewer completed all possible cases in the segment. At the time the assignment was made, the interviewer was given 2 to 3 weeks (depending on the size of the segment) to complete the initial effort for all cases in the segment. Most productive interviewers were able to handle up to five segments simultaneously during the initial phase. Less productive interviewers were given only two or three segments at a time.

### 8.6.2 Schedule and Production

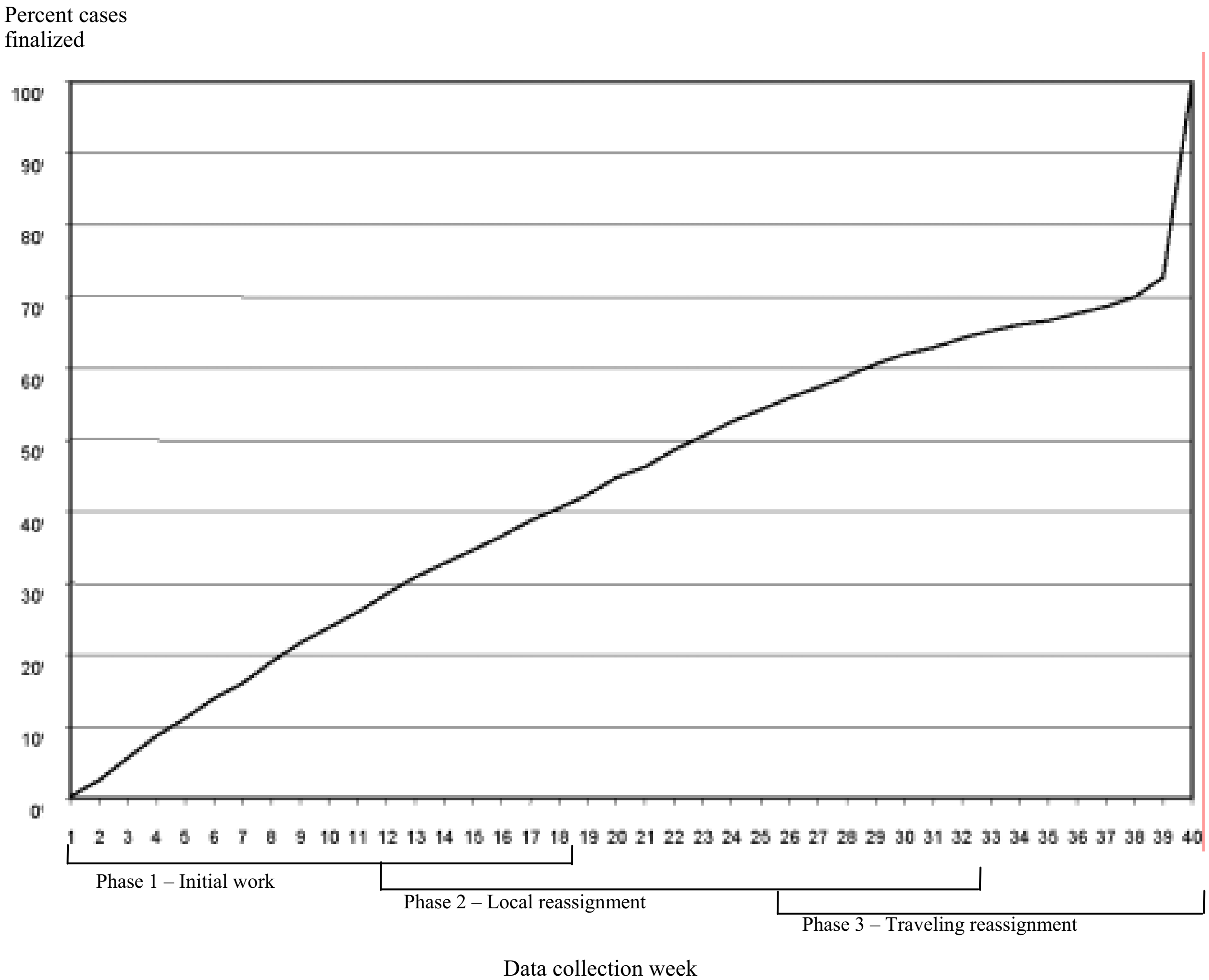
The original plan for the NAAL field effort envisioned a 26-week (6-month) field period, in which 4 months would be used to complete the initial complement of calls to all assigned households and 2 months would be reserved at the end for intensive nonresponse conversion by the traveling team of interviewers. The second phase was planned to overlap with the first, beginning at about month 3.

To allow more time to increase response rates, a decision was made late in the field period to extend the field period to approximately 40 weeks. Figure 8-1 shows the cumulative percentage of cases that were closed out, by week of the field period, for all household sample cases. The three-phase approach to data collection, as described in section 8.6.1, is shown here as well. (The spike in closed-out cases at the end of the field period occurred because nonresponse cases were not finalized until the end of the study.)

The pace of the field effort was influenced by several factors unique to NAAL. First, the design called for respondents to complete the background questionnaire, the assessment, and the oral module in the same visit; thus, it was necessary for a respondent to have a period of more than 1 hour during which he or she was reasonably unlikely to be interrupted. This requirement reduced the likelihood that respondents would be available on the interviewer's first visit and necessitated additional callbacks to ensure the completion of the case. In addition, because of the very large number of cases, across-the-board decisions to improve productivity took considerable time to implement.

Further, several natural disasters, such as hurricanes, fires, and snowstorms, made interviewing in selected areas impossible for periods of time. Finally, the climate of fear and suspicion fueled by the September 11, 2001 terrorist attacks created an environment in which people were less likely to participate.

**Figure 8-1. Percentage of closed-out cases, by week of field period and phases of data collection: 2003**




SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 8.6.3 Data Collection Procedures

The interviewer's first task was to conduct an in-person screening interview with a respondent at each of the sampled housing units that was occupied. Interviewers were instructed to administer the screener to a household member aged 16 or older, in order to identify eligible sample persons in the household.

Interviewers were required to complete the background questionnaire, the assessment (either the main assessment or the ALSA), and the oral module with the selected respondent in the same visit. The interviewer informed the respondent of the amount of time needed to complete all instruments and attempted to transition directly into the interview after completing the screener. In situations where this was not possible, the interviewer set an appointment time to complete the interview and confirmed the appointment by telephone at a later date.

If a respondent demonstrated reluctance to participate, through either numerous broken appointments or a voiced refusal, the interviewer completed a noninterview report form and discussed further strategies with the supervisor.

A toll-free hotline was established for respondents to call with any additional questions (see section 3.3.6.3). The telephone number for this hotline was included on numerous study materials, such as the advance letter, brochure, letter of introduction, and community authorization letter. A total of 785 calls were received by the hotline over the course of the study. Most of the calls were requests for study verification, appointment scheduling, or refusals.

The average administration time for the NAAL instruments is displayed in table 8-3.

**Table 8-3. Average administration time, by instrument: 2003**

| Instrument | Average administration time (minutes) |
|---|---|
| Total | 91.8 |
| Main assessment sample persons | |
| Screener | 4.5 |
| Background questionnaire | 27.4 |
| Core and main assessment items | 45.7 |
| Oral module | 14.2 |
| Total | 77.1 |
| ALSA sample persons | |
| Screener | 4.5 |
| Background questionnaire | 27.4 |
| Core and supplemental study | 32.2 |
| Oral module | 13.0 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

## 8.7 QUALITY CONTROL MEASURES AND FEEDBACK TO STAFF

In addition to the CAPI range and logic checks and home office review of completed cases, the quality control measures described in the following sections were implemented to ensure high-quality work in the data collection phase. The procedures were designed to assess the quality and completeness of data as collected and to provide timely feedback to the supervisors, the home office, and the interviewers.

### 8.7.1 Validation

A 10 percent subsample of completed and noninterview cases were randomly selected for validation. Because falsification activity that goes undetected for a long time is very costly to correct, it is desirable to perform validation as soon after interviewing as possible. The validation efforts conducted by the field supervisors were supplemented by experienced interviewers from the data collection contractor's Telephone Research Center to ensure timely validation efforts.

Validation was performed by telephone when possible. On all cases selected for validation for which telephone numbers were not available, in-person validation was performed by a different interviewer who worked in the same or a nearby PSU.

For completed cases, the validation interview verified that contact had been made and confirmed the respondent's address at the time of the contact. Then three questions from the background questionnaire were asked again, to verify that the responses were consistent with those provided earlier. The respondent was also asked how long the interviewer spent with the respondent on the day of the interview, whether the interviewer assisted the respondent with the assessment, whether the interviewer conducted the oral module, and how much the respondent had been paid for participating in the study. For incomplete interviews, the validation interviewer verified contact (if the interviewer's original report of the case indicated that contact had occurred), confirmed the respondent's address and the number of household members at the time of contact, and, if possible, tried to schedule an appointment for an interview.

As soon as validation for a case was completed, a validation result code for the case was entered into the Data Management System. The Data Management System was used to monitor the progress of the validation effort and to ensure that at least 10 percent of each interviewer's work was being validated. At the end of data collection, 5,458 cases had been validated either by telephone or in person, for an overall validation rate of just over 15.4 percent of finalized screeners. Some of the validation workload in excess of the original 10 percent requirement resulted from situations in which additional verification was required to dismiss or confirm suspected falsification of interviews.

Of the nearly 400 field interviewers who worked on NAAL, 17 were discovered to have submitted some fraudulent work. Each of these interviewers was dismissed as soon as falsification was confirmed, and all of their completed work was validated in person by another interviewer. Through in-person and telephone validation, it was determined that 263 of the 684 cases completed by these interviewers (38 percent) had been falsified. The data associated with these cases was deleted and the cases were reassigned to the field for completion.

### 8.7.2 Computer-Generated Reports

The Supervisor Management System was used to manage and monitor the progress of the field work and provided critical management information to field and home office staff. One of the Supervisor Management System components was a reports mechanism. Reports were updated in real time as an interviewer or supervisor entered information or transmitted data.

The reports allowed all levels of management to monitor, on a daily basis, the progression of completion rates, response rates, and distribution of cases in pending codes by region, interviewer, PSU, and segment for each of the study instruments. Managers were also able to view the daily EROC information entered by interviewers. Costs were entered and monitored on a weekly basis as well.

The reports available to the management staff are discussed below.

The Data Collection Reports and the Interviewer Reports allowed the supervisors to view the overall status of production and response rates for their region and the nation.

The Interviewer Conference Report gave the supervisor a detailed view of pending cases for every interviewer, including the status for each component of the case. The Interviewer Cost Summary Report listed each interviewer’s weekly and cumulative hours, expenses, and number of completed background questionnaires. Using these two reports as a starting point, supervisors discussed each pending case as well as overall completion rates, response rates, and costs with each interviewer during the weekly conference call. Plans and strategies for handling each pending case and overall performance were discussed during these calls, and workloads were adjusted as necessary. The quality of an interviewer’s work was also discussed, based on the results of tape edits, interviewer observations, and validations.

The Missing Booklet Report listed finalized cases for which booklets had not been receipted in the field room.

The Interviewer Transmission Report documented each interviewer's last transmission of new data. Supervisors used this report to determine whether an interviewer was having difficulty in transmitting because of a problem in understanding the transmission process or a problem with the laptop. Supervisors monitored this report very closely because it affected the accuracy of all other reports.

The Assignment History of Nonfinalized Cases Report allowed supervisors to view and track the history and progress of reassigned cases. The supervisors could view, in one place, each interviewer's attempts to complete the instruments for cases that had been reassigned for efficiency and nonresponse conversion.

The Unassigned Cases Report documented any unassigned cases and also allowed the supervisor to check for entry errors in assigning or transferring cases.

The Validation Report documented all cases preselected for validation, including the completion or finalization date of each case to indicate that it was available for validation. Supervisors used this report to document the results of their validation efforts and to select any additional or substitute cases for validation. A summary report by regional and national totals was also available.

Additional reports were run each week to allow the project statisticians to monitor the sample yield for various populations in the national and state samples. Selected variables from the sample selection file, which carried census race characteristics by segment and by PSU, were merged with production data from the Supervisor Management System to allow a comparison of projected and actual results.

Reports including interview timing and scheduling data were also produced for each interviewer, from the CAPI database of finalized cases. Field managers used the reports to look for anomalies within the data that might identify interviewer falsification of interviews. Examples of these anomalies are very short instrument administration times, a short amount of time between interviews conducted at two different households, and interviews conducted very early in the morning or late in the evening.

### 8.7.3 Observation

Two methods were used to observe NAAL interviews: (1) tape recording of interviews for review by supervisors and (2) in-person observations by home office staff and field managers and supervisors.

#### 8.7.3.1 Tape-Recorded Interviews

Supervisors relied on the review of tape-recorded interviews to “observe” each interviewer. Each interviewer was required to tape record two complete interviews, that is, the entire screener, background questionnaire, assessment, and oral module (to the extent that the respondent was able to complete the instruments). So that cases worked early in the data collection period could be reviewed and any errors in administration corrected, interviewers were asked to tape their 3rd and 20th interviews; if the respondent did not give consent, the next interview was taped instead. After listening to the tape, the supervisor completed a tape edit form and gave the interviewer feedback on the quality of the interviewing techniques and any mistakes or areas for improvement. A total of 569 tape-recorded interviews were evaluated.

#### 8.7.3.2 In-Person Observations

Interviewer observations were performed by the home office field managers; other members of the contractors’ and NCES staffs; and supervisors whose field offices were in sampled PSUs. Interviewer observations were performed for two main purposes. One purpose was to give home office staff an opportunity to observe respondents’ reactions to the study and also to observe how well the field procedures worked. Supervisors identified their strongest interviewers for this type of observation. The second purpose was to observe interviewers whose performance was of some concern, either because of their evaluation during training or because they were assigned to a particularly difficult area. Interviewers new to interviewing were observed as well.

Interviewers were typically observed locating sampled housing units, making screener contacts, setting appointments, and completing at least one background questionnaire, assessment, and oral module. During an interview, the observer listened but did not participate in any way. After the interview, when the observer and interviewer had left the respondent’s home, the observer used an interviewer observation form to evaluate the quality of the interviewer’s work. Interviewers were evaluated on the following points: ability to gain access to the household, organization of material and equipment, knowledge of the study, administration of the instruments, and general interviewing techniques. A total of 31 field observations were completed.

### 8.7.4 CAPI Help Desk

A CAPI help desk was established and operated by staff specially trained in the NAAL instrumentation. If interviewers or supervisors experienced technical problems with the CAPI system, they could call the toll-free help desk number and receive assistance in resolving the problem. The help desk received 1,575 calls during the data collection period.

The largest number of calls to the help desk related to data cleanup (cases where data required editing or cleaning by the home office); transmission, time and expense reporting, and shipment (e.g., connecting the laptop to the home office through the telephone lines, sending and receiving cases, and accessing or entering shipment data); laptop hardware and equipment (e.g., mouse, power cord, and carrying case); and the Interviewer Management System or e-mail (problems with user names or passwords, launching CAPI instrumentation, or receiving e-mail messages).

## 8.8 ACHIEVING HIGH RESPONSE RATES

Response rates on household studies are influenced by three broad categories of factors:

- the ability of the interviewers to obtain cooperation;
- the effectiveness of "callback" procedures; and
- the efforts made by interviewers and supervisors to convert initial nonresponse cases to completed interviews.

These factors are described in the following subsections.

### 8.8.1 Interviewers' Ability to Obtain Cooperation

An important factor in maximizing response rates is the ability of the interviewers to encourage respondents to participate. Two sessions during the interviewer training program involved round table discussions and exercises focused specifically on techniques for handling reluctant respondents, answering questions, and avoiding refusals. Before working with actual households, the interviewers progressed through several stages of practice during training. During the first stage, they conducted role-playing exercises with one another until they felt comfortable and could demonstrate an adequate level of skill in gaining cooperation. After reaching this point, the interviewers conducted a practice interview with a paid volunteer respondent. Training staff monitored the interviewers closely during this segment of the training.

To assist the interviewers in gaining respondent cooperation, all sampled households received a cover letter and brochure approximately 7 to 10 days before the interviewer attempted to complete the screener.

### 8.8.2 Callback Procedures

Developing an effective strategy for visiting housing units is a fundamental of good interviewing. Interviewers were trained in the following rules to build response rates:

- make trips at different times of day (morning, afternoon, or evening), taking into account that late afternoons and evenings would be the most productive hours in most cases;
- make trips on different days of the week; and
- make trips on weekends (Saturday or Sunday).

If they were unable to complete a screener during the first four visits to a housing unit, the interviewers were instructed to complete a noninterview report form and call the supervisor, who would either authorize more visits or assign the case to another interviewer.

### 8.8.3 Efforts to Convert Nonresponse

Each type of nonresponse required a different strategy for conversion. The conversion strategies are summarized in the following sections.

#### 8.8.3.1 Refusals

Refusals are the most difficult type of nonresponse case to convert. When a respondent refused or broke off an interview, the interviewer completed the noninterview report form to capture information about the reason for the refusal. Using this information, the interviewer and supervisor could decide to send the respondent one of several refusal conversion letters or transfer the case to a different interviewer. Seven refusal conversion letters were developed for the main data collection:

- general refusal letter;
- no selling letter;
- too busy letter (in English and Spanish);
- sample person selected letter (in English and Spanish);
- elder letter (written in large print);
- locked building letter; and
- a simple, short letter designed for those who were not expected to take the time to read a more detailed letter.

#### 8.8.3.2 Not at Home

Interviewers were supplied with sorry-I-missed-you cards that could be left at the door when nobody was home. Interviewers were instructed to return to the home at different times of the day and on different days of the week to attempt to contact household members.

#### 8.8.3.3 Language Problem

Whenever possible, interviewers fluent in Spanish were sent to Spanish-speaking households. When a bilingual interviewer was not available, interviewers were instructed to locate a bilingual household member or neighbor or to hire a translator from a local community center to assist in conducting the screener component of the interview.

In households where neither English nor Spanish was spoken, interviewers tried to arrange for a household member, child, or neighbor (aged 16 or older) to assist in the translation, for the screener only. The background questionnaire, assessment, and oral module were not administered to respondents who could not speak English or Spanish.

#### 8.8.3.4 Illness

Whenever a respondent was too ill to participate, interviewers filled out a noninterview report form and discussed the situation with their supervisor.

#### 8.8.3.5 Vacant Housing Unit

If a housing unit was vacant during the interviewer's first visit, the case was closed out as vacant. If, however, the interviewer made initial contact with a household but returned to find that the housing unit was vacant or that a new family had moved in, the interviewer attempted to interview the household members who had lived in the housing unit at the time of the original contact.

### 8.8.4 Strategies To Increase Response Rates

Many strategies were employed throughout the NAAL field period to increase response rates. They included collecting information about the sampled households from neighbors, administering a hard-copy version of the screener, sending FedEx mailings, and implementing interviewer incentive programs. These strategies are discussed below.

#### 8.8.4.1 Neighbor Information

After unsuccessful attempts to contact the sampled address, interviewers were allowed to use neighbor information as a mechanism to identify ineligible housing units in high-minority segments only. Interviewers were required to collect the following information from two neighbors of the selected household: whether the housing unit was occupied, the best time and day to contact the household, and whether any household members were Hispanic or Black. The names, addresses, and telephone numbers of the neighbors were collected as well. If neither neighbor report indicated that there were minority members in the sampled household, the supervisor finalized the case as ineligible. In all other circumstances, the interviewer was required to continue efforts to interview the sampled household. This procedure was implemented in November 2003 and used for the remainder of the field period.

#### 8.8.4.2 Hard-Copy Screener

To increase response rates, a hard-copy version of the CAPI screener was developed and implemented in segments where a hard-copy screener was deemed preferable to administering a CAPI instrument at the doorstep. This approach was used in households where the presence of the laptop computer was expected to be intimidating or imposing to potential respondents. Once the data had been collected on the hard-copy form, NAAL interviewers transferred the information into the CAPI screener to determine whether any household members were eligible for the study.

The hard-copy screener was used by the NAAL interviewers, as well as by approximately 55 experienced interviewers on short-term loan from other studies conducted by the data collection contractor. The non-NAAL interviewers were teamed with a NAAL interviewer who followed up with the selected respondent to complete the remainder of the interview, if an eligible respondent was identified. Nearly 1,000 screeners were completed by these non-NAAL interviewers.

#### 8.8.4.3 FedEx Mailings

In December 2003 and January 2004, more than 6,000 FedEx mailings were sent to households where interviewers had made numerous visits but were unable to find anyone at home and where attempts at refusal conversion had been unsuccessful. The FedEx mailings contained the NAAL advance letter, the study brochure, and a brief letter, when appropriate. These materials helped introduce the interviewer to households that had not received or read the previous mailings.

#### 8.8.4.4 Interviewer Incentives

Several interviewer incentive plans aimed at increasing production were put into place throughout the last several months of the data collection period. These plans encouraged interviewers to complete as many interviews as possible during periods when they were less likely to work, such as harsh weather conditions and the winter holiday season. These additional incentive plans were found to spur production. All interviews completed as part of an incentive plan were validated.

### 8.8.2 Reasons for Nonresponse

There are two major reasons for nonresponse: (1) literacy-related nonresponse,[2] that is, sample persons who did not respond because of language problems, reading or writing barriers, or mental disabilities; and (2) other nonresponse, that is, sample persons who did not respond for other reasons (e.g., refusals or unlocatable cases). Table 8-4 shows the sample counts and percentages by reason for nonresponse for the national and state samples for the screener, the background questionnaire, the assessment, and the oral module. Among these instruments, literacy-related nonresponse was most prevalent for the background questionnaire (1.5 percent of sample persons).

Of the 35,365 housing units sampled in the national and state components, 13.2 percent were unoccupied; that is, they were either vacant (8.6 percent) or were not a housing unit (4.6 percent). A screener was completed at 25,123 of the 30,694 occupied housing units in the sample, for an overall unweighted screener response rate[3] of 81.8 percent. The largest category of screener nonresponse was "refusals," which occurred in 3,207, or 9.1 percent, of the occupied households. The second largest category was maximum callbacks, which occurred in 1,625, or 4.6 percent, of the occupied households. There were 160 cases of literacy-related nonresponse (0.5 percent), all classified as language problems.

The screening effort identified 22,270 households that included at least one eligible person. Among the 23,732 sample persons, 18,186 (76.6 percent) completed the background questionnaire. Refusals, which accounted for the largest number of nonrespondents to the background questionnaire, occurred in 3,032 cases (including 153 cases where someone refused for the sample person), or 12.7 percent of all sample persons. The next category was maximum callbacks, which accounted for 1,401

---

[2] Of the 18,541 cases that were weighted in the household sample, 439 did not complete the assessment due to mental disability, language problem or reading/writing barrier. These literacy-related cases are counted as respondents for response rate and nonresponse bias analysis purposes, since some useful data was collected from the screener and background questionnaire (for some), and some knowledge of the English literacy skills was also obtained. After the weighting of the data and the computation of response rates, a decision was made to exclude disabilities and cases with reading/writing barriers from the target population. However, these cases are in scope for this report.

[3] Weighted response rates are discussed in chapter 11.

cases (5.9 percent). A total of 355 sample persons (1.5 percent) were considered literacy-related nonrespondents because of a language problem (211 cases, or 0.9 percent) or mental disability (144 cases, or 0.6 percent).

Table 8-4 also shows the reasons for nonresponse to the assessment. Of the 23,732 sample persons, 17,172 (72.4 percent) completed the assessment, and an additional 548 (2.3 percent) partially completed it. The reasons for partial completion were both literacy related (1.0 percent) and nonliteracy related (1.3 percent). Further, 5,540 sample persons did not attempt the assessment because they did not complete the background questionnaire. Of the 6,012 sample persons who did not attempt the assessment, 290, or 1.2 percent of all sample persons, were classified as refusals.

Only 51 sample persons (0.2 percent) were classified as having literacy-related reasons for not attempting the assessment, including language problems (27 cases, or 0.1 percent), reading/writing barriers (14 cases, or .1 percent), and mental disabilities (10 cases, or 0 percent).

The reasons for nonresponse to the oral module are also provided in table 8-4. Of the 23,732 sample persons, 17,057 (71.9 percent) completed the oral module, and an additional 23 (0.1 percent) partially completed it. The reasons for partial completion were both literacy related (9 cases) and nonliteracy related (14 cases). Further, 6,199 sample persons did not attempt the oral module because they did not complete the assessment. Of 6,652 sample persons who did not attempt the oral module, 191, or 0.8 percent of all sample persons, were classified as refusals. Only 145 sample persons (0.6 percent) were classified as having literacy-related reasons for not attempting the oral module, including language problems (70 cases, or 0.3 percent), reading/writing barriers (50 cases, or 0.2 percent), and mental disabilities (25 cases, or 0.1 percent). For response rates associated with the household, by major instruments and by key variable, refer to chapter 11.

**Table 8-4. Sample counts by instrument and reasons for nonresponse: 2003**

| Instrument and reasons for nonresponse | National household total | State household total | Combined household total |
|---|---|---|---|
| Screener | | | |
| Total housing units | 25,450 | 9,915 | 35,365 |
| | Percent | Percent | Percent |
| All housing units (percent) | 100.0 | 100.0 | 100.0 |
| Complete, 1 sample person | 56.2 | 65.6 | 58.8 |
| Complete, 2 sample persons | 4.1 | 4.1 | 4.1 |
| Complete, subsampled out using neighbor information | 0.6 | 0 | 0.4 |
| Complete, subsampled out | 10.6 | 0 | 7.6 |
| Illness | 0.5 | 0.8 | 0.6 |
| Language problem | 0.4 | 0.5 | 0.5 |
| Maximum callbacks | 4.9 | 3.8 | 4.6 |
| Not a housing unit | 4.6 | 4.6 | 4.6 |
| Other | 0.9 | 0.8 | 0.8 |
| Refused | 8.2 | 11.2 | 9.1 |
| Unavailable for field period | 0.2 | 0.2 | 0.2 |
| Vacant | 8.8 | 8.3 | 8.6 |
| Background questionnaire | | | |
| Total sample persons | 16,409 | 7,323 | 23,732 |
| | Percent | Percent | Percent |
| All sample persons (percent) | 100.0 | 100.0 | 100.0 |
| Complete | 76.1 | 77.8 | 76.6 |
| Language problem | 1.0 | 0.6 | 0.9 |
| Maximum callbacks | 6.5 | 4.5 | 5.9 |
| Mental disability | 0.6 | 0.6 | 0.6 |
| Other | 0.8 | 0.7 | 0.8 |
| Physical disability | 1.4 | 1.3 | 1.4 |
| Sample person refused | 11.8 | 12.8 | 12.1 |
| Someone refused for sample person | 0.7 | 0.6 | 0.6 |
| Unavailable for field period | 1.1 | 1.0 | 1.0 |

See notes at end of table.

**Table 8-4. Sample counts by instrument and reasons for nonresponse: 2003—Continued**

| Instrument and reasons for nonresponse | National household total | State household total | Combined household total |
|---|---|---|---|
| Assessment | | | |
| Total sample persons | 16,409 | 7,323 | 23,323 |
| | Percent | Percent | Percent |
| All sample persons (percent) | 100.0 | 100.0 | 100.0 |
| Reading/writing barrier | 0.1 | 0.1 | 0.1 |
| Complete | 71.6 | 74.2 | 72.4 |
| Language problem | 0.1 | 0.1 | 0.1 |
| Mental disability | 0.1 | 0 | 0 |
| Not applicable, did not complete background questionnaire | 23.8 | 22.2 | 23.3 |
| Other | 0.1 | 0.1 | 0.1 |
| Partial complete, reading/writing barrier | 0.6 | 0.2 | 0.5 |
| Physical disability | 0.4 | 0.3 | 0.4 |
| Partial complete, refused, core incomplete | 0.2 | 0.1 | 0.2 |
| Partial complete, language problem | 0.5 | 0.2 | 0.4 |
| Partial complete, mental disability | 0.1 | 0.1 | 0.1 |
| Partial complete, physical disability | 0.3 | 0.8 | 0.5 |
| Partial complete, refused, core complete | 0.7 | 0.4 | 0.6 |
| Sample person refused | 1.2 | 1.0 | 1.2 |
| Someone refused for sample person | 0 | 0.1 | 0.0 |
| Unavailable for field period | 0 | 0.1 | 0.1 |
| Oral module | | | |
| Total sample persons | 16,409 | 7,323 | 23,732 |
| | Percent | Percent | Percent |
| All sample persons (percent) | 100.0 | 100.0 | 100.0 |
| Reading/writing barrier | 0.2 | 0.2 | 0.2 |
| Complete | 71.0 | 73.8 | 71.9 |
| Language problem | 0.4 | 0.1 | 0.3 |
| Mental disability | 0.1 | 0.1 | 0.1 |
| Not applicable, did not complete assessment | 26.8 | 24.5 | 26.1 |
| Other | 0.1 | 0.2 | 0.1 |
| Partial complete, reading/writing barrier | 0 | 0 | 0 |
| Partial complete | 0 | 0 | 0 |
| Physical disability | 0.3 | 0.5 | 0.4 |
| Partial complete, language problem | 0 | 0 | 0 |
| Partial complete, mental disability | 0 | 0 | 0 |
| Partial complete, physical disability | 0 | 0 | 0 |
| Sample person refused | 0.9 | 0.6 | 0.8 |
| Unavailable for field period | 0 | 0 | 0 |

NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 8.8.3 Demographic Profiles of Respondents

This section presents demographic profiles of the respondents at each stage. Table 8-5 presents the distributions of respondents with respect to selected demographic and other characteristics, for the screener, background questionnaire, assessment, and oral module. The counts and percentages presented in the table are unweighted and were computed from the combined NAAL-SAAL household sample.

Table 8-5 shows some simple distributions among the samples of respondents at the various stages of data collection. For a comprehensive analysis of the differences between respondents and nonrespondents, a complete nonresponse bias analysis is provided in chapter 11. The nonresponse bias analysis provides an evaluation of the potential for bias owing to nonresponse to the screener and the background questionnaire.

Table 8-5 also provides the sample counts of screener respondents. The respondent counts reflect the number of housing units that completed the screener or were ineligible because of subsampling procedures. Minority status was defined as high if the sample housing unit was from a segment in which at least 25 percent of the population was Black or Hispanic and was defined as low otherwise. The majority of respondent housing units came from the South and were from metropolitan statistical areas (MSAs).

The background questionnaire respondent counts reflected both the number of sample persons who completed the background questionnaire and the number who did not complete the background questionnaire for literacy-related reasons. About 36 percent of the sample were minority sample persons (i.e., Hispanic or non-Hispanic Black adults), which was a reflection of the oversampling procedures. The distributions of respondents for the assessment and the oral module were very similar to the distribution for background questionnaire respondents because the level of nonresponse to the assessment and oral module was low among background questionnaire respondents. It should be noted that assessment and oral module respondents were defined as those who (1) completed the assessment; (2) those who only partially completed the assessment (because of a language problem, mental disability, physical disability, or reading/writing barrier); and (3) those who did not attempt the assessment (or oral module) because of a language problem, mental disability, physical disability, or reading/writing barrier.

**Table 8-5. Demographic profile of respondents in the combined household sample, by instrument: 2003**

| Characteristic | Screener total | Background questionnaire total | Assessment total | Oral module total |
|---|---|---|---|---|
| Total sample | 25,123 | 18,541 | 17,668 | 17,300 |
| | Percent | Percent | Percent | Percent |
| Total percent | 100.0 | 100.0 | 100.0 | 100.0 |
| Region | | | | |
| Northeast | 19.6 | 20.3 | 20.2 | 20.2 |
| Midwest | 18.4 | 19.5 | 19.4 | 19.4 |
| South | 44.3 | 44.6 | 45.3 | 45.6 |
| West | 17.7 | 15.6 | 15.0 | 14.9 |
| MSA[1] status | | | | |
| MSA | 80.6 | 79.0 | 78.7 | 78.8 |
| Non-MSA | 19.4 | 21.0 | 21.3 | 21.2 |
| Minority status[2] | | | | |
| High | 45.7 | 40.4 | 40.3 | 40.2 |
| Low | 54.3 | 59.6 | 59.7 | 59.8 |
| Gender | | | | |
| Male | † | 43.3 | 42.9 | 42.8 |
| Female | † | 56.7 | 57.1 | 57.2 |
| Age | | | | |
| 16–29 | † | 25.4 | 25.8 | 26.0 |
| 30–49 | † | 39.2 | 39.3 | 39.4 |
| 50–69 | † | 24.6 | 24.6 | 24.6 |
| 70+ | † | 10.8 | 10.3 | 10.1 |
| Race/ethnicity | | | | |
| Hispanic | † | 17.2 | 17.1 | 17.0 |
| Black | † | 18.9 | 19.2 | 19.2 |
| Other | † | 63.9 | 63.7 | 63.8 |

† Not applicable.
[1] Metropolitan Statistical Area.
[2] If a sampled segment (group of census blocks) contained 25 percent or more Black and Hispanics, then the segment was classified as high minority. Otherwise the segment was classified as low minority.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

## 8.9 DATA PREPARATION AND PROCESSING

During the data collection period, interviewers returned materials to the home office in two formats: hard copy and electronic. The hard-copy materials returned included assessment booklets, ALSA questionnaires, household folders, noninterview report forms, and oral module interviewer guides. The electronic data consisted of screener, background questionnaire, assessment interviewer guide, oral module, and interviewer observation data, as well as status code updates, and data indicating which hard-copy materials had been shipped by the interviewer to the home office.

The Data Management System was used to support data processing activities, among other functions. The Data Management System consisted of three main components: the Supervisor Management System (see section 8.3.2.2), receipt control functions, and the reporting mechanism. This section of the report focuses on the receipt control functions.

During the field period, interviewers typically transmitted interview data to the home office daily. During transmission, the updated status codes and questionnaire data for all completed and in-process cases were combined and sent to a server through a secure dial-up connection. Transmitted files were backed up and held on the server. Approximately every 5 minutes, an automatic process determined whether new transmissions had been received from the field interviewers. This process was an ongoing operation during the field period.

The process that handled data transmissions performed two functions: It updated the Data Management System database with case status information, and it moved the completed interview data to processing directories on the project server. Once a day, the newly transmitted Blaise data files created during the interview were decrypted and concatenated to create a project-level Blaise database for each instrument.

The study database was updated almost immediately after the transmitted data were received. Backup processes were in place to ensure that transmitted data were received successfully. During the conversion of the Blaise data to the study database, two other products were created. The first was a Blaise interview browse area that allowed project staff to locate data on individual cases. This interview browse function was used primarily in resolving issues reported to the hotline. The second product was a SAS data set that was used to complete ad hoc reporting requests, such as task timing reports.

### 8.9.1 Overview of Data Preparation and Processing Activities

Several data preparation and processing systems and procedures were developed to support the extraction of screener, background questionnaire, and interviewer guide data from the Blaise system and into SAS data files. Cases transmitted by interviewers were appended to a cumulative database, where they were then prepared for a review of interviewer comments and coding of other-specify and open-ended responses. Once these steps were completed, the data were converted into the study database. Frequencies were run daily on the study database and reviewed for outliers and inconsistencies. After the frequency review task, the data were loaded into the Blaise Editing System, batched, tracked, and cleaned.

### 8.9.2 Receipt, Batching, and Editing of Cases

Once data had been moved into the Blaise Editing System, manageable batches of 100 cases were created. The Blaise Editing System tracking file, which contained all the records within a batch, was then developed. A sequential batch number was used to link the tracking file records to the batch file records. After batching, a utility in the Blaise Editing System assigned batches to a data editor.

The Blaise Editing System management report identified the cases within each batch that required data cleaning. During the creation of the batches, any automated edits that had been triggered and suppressed by the interviewer during the interview were reinstated and the interview status was reset. Additional postcollection editing, including edits considered too time-consuming or complex to be executed in an interview situation, was conducted as well. The cases were then verified, coded, adjudicated, and prepared for delivery.

### 8.9.3 Processing of Hard-Copy Assessment Materials

Interviewers were instructed to return completed interview materials (including assessment and ALSA booklets, household folders, oral module interviewer guides, and noninterview report forms) to the home office twice a week. All shipments were tracked electronically with the Interviewer Management System and the Data Management System. Before the interviewer mailed the items, the electronic shipping record section of the Interviewer Management System was used to indicate each item included in the shipment. After marking each case and item being sent, the interviewer entered the shipping tracking number into the Interviewer Management System. The home office was then able to see which case materials were in transit.

The status and location of all assessment and ALSA booklets were constantly monitored. The Data Management System automatically changed the status associated with each booklet to reflect when it was sent to an interviewer, when it was marked as shipped by the interviewer, and when it was received at the home office. Upon arrival, completed interview materials were receipted in the Data Management System, and items received were compared against the list of items expected. Any discrepancies were brought to the attention of the systems management team and, where needed, the field director and supervisors.

Once an assessment booklet had been receipted, it was filed by case ID and, where appropriate, sample person ID within case. ALSA booklets, household folders, and oral module interviewer guides were filed separately by case ID. Noninterview report forms were maintained in the associated household folder.

As part of the standard NAAL quality control procedures, all seven core items in the completed assessment booklets were rescored by trained home office staff. The validation results were entered into a specially designed core scoring program. The program compared the interviewer's scoring with that of the home office staff, enabling supervisors to give interviewers feedback on their performance.

Early in the data collection period, the core items were rescored for 100 percent of the receipted assessment booklets. When interviewers were determined to be proficient at scoring the core items, no further core validation was conducted for those interviewers. Home office staff continued to rescore 100 percent of the core items for the remaining interviewers.

The in-house validation of the core items continued through the middle of January 2004, as the end of data collection approached. Home office staff rescored a total of 13,608 core assessments. (See section 13.3.2.2 for a further discussion of core validation.)

### 8.9.4 Processing of Oral Module Data and Zip Disks

Interviewers were required to mail a zip disk of oral module recordings to the home office once each week. Each disk contained all the oral module interviews completed by the interviewer during the week. The files received from the interviewers were copied into a backup directory. These files were maintained until confirmation was received from the oral module scoring contractor that the recordings had been successfully received and processed. Zip disks were sent for scoring on a weekly basis throughout the data collection period.

### 8.9.5 Data Processing for the ALSA

As discussed in section 8.9.3, ALSA booklets were receipted in the home office. Following receipt, the booklets were reviewed and edited, during which trained staff reviewed notes written in the margins and ensured that every questionnaire item had a valid response associated with it and that the skip patterns had been correctly followed by the interviewer. After the booklets had been edited, they were prepared for data entry. All booklets were double-keyed by independent data entry staff at a rate of 100 percent. Any discrepancies between the two entry processes were reconciled. After the data entry procedure was completed, the codebooks and associated frequencies were reviewed and adjudicated.

### 8.9.6 Delivery of Assessment Booklets to the Scoring Contractor

Assessment booklets were delivered to the assessment scoring contractor at four points during the field period. Each shipment included an electronic file listing all the items included in the shipment for verification purposes.

### 8.9.7 Quality Control of Data in the Study Database

As discussed above, a detailed reporting system, integrated with the Data Management System, was used throughout the main study data collection. Data in these reports were reviewed by the field director, field managers, supervisors, and home office staff throughout the field period.

In addition, receipt control reports were used to track and verify that assessment booklets were sent to the home office on a timely basis after completion. The integrated receipt control system also tracked the assessment booklets throughout the field period and documented the location of each booklet.

At the end of the field period, an extensive data validation process was completed to ensure that Blaise interview data existed for each completed case in the Data Management System. Status codes were compared to ensure that each status within the Data Management System accurately reflected the existence of the Blaise interview data. The CAPI instrument data were compared, and all discrepancies were documented, reviewed, and corrected, if feasible to do so, before delivery of the final data to the analysis contractor. In addition, some problem cases were restored from laptop backup disks to ensure the existence of data for all completed cases.

Additionally, the assessment scoring data were reviewed to ensure that a score had been received for each case ID and booklet barcode number sent for scoring. When discrepancies could not be resolved, differences were documented in the final delivery documents.

Finally, a thorough process of reconciling FAN data was conducted to ensure that data were available for all finalized cases with a completed FAN interview. Any cases that were determined to be missing from the scoring database were reprocessed. The associated responses were then generated and the data were redelivered.

*This page is intentionally left blank.*

# CHAPTER 9

# CORRECTIONAL INSTITUTION STUDY DATA COLLECTION AND QUALITY CONTROL

Michelle Amsbary, Terri Annis, and Martha Berlin, Westat

## 9.1 INTRODUCTION

The purpose of the Correctional Institution Study, hereinafter referred to as the Prison Study, was to assess the literacy skills of adult inmates in federal, state, and private correctional facilities, using a sample of approximately 1,100 inmates. The inclusion of the inmate sample in the National Assessment of Adult Literacy (NAAL) helped improve estimates of the literacy levels of the total U.S. population and made it possible to report on the proficiencies of this important segment of society. The Prison Study component was developed in consultation with the Bureau of Justice Statistics of the U.S. Department of Justice and the Federal Bureau of Prisons.

Data collection took place following the completion of the main household study. The field period ran for approximately 4 months, from mid-March through July 2004. A total of 1,173 background questionnaire interviews were completed in 107 facilities. A complete description of the sample design is provided in chapter 7.

To ensure comparability with the NAAL household study, the inmates completed the same literacy tasks as the household population. However, to address issues of particular relevance to the prison population, a revised version of the background questionnaire was developed. New questions included queries about current offenses, criminal history, and participation in prison programs, as well as education and workforce experience. (See section 2.3 for a complete description of questionnaire development and content.)

As in the Prison Study conducted as part of the 1992 National Adult Literacy Survey (NALS), the rules of almost all facilities precluded monetary or in-kind incentives. Instead, a personalized certificate of participation was placed in the inmate's file upon completion of the interview.

Before the main study data collection, a small pretest was conducted at three correctional institutions, one in Maryland and two in Texas. Additionally, the Spanish version of the background questionnaire was tested at a Maryland facility. The pretest evaluated the ease of administration of the study instruments, administration time, within-facility procedures, and inmate reaction to the study. On the basis of the pretest experience, minor changes were made to the background questionnaire to facilitate administration, and administrative procedures were refined to reflect lessons learned.

## 9.2 GAINING COOPERATION

The permission and cooperation of federal, state, and correctional facility officials were required before interviewing in prisons. Representatives from the Bureau of Justice Statistics and from the Office of Vocational and Adult Education within the U.S. Department of Education provided considerable assistance in gaining cooperation from federal and state correctional agency officials. Letters of endorsement were obtained from the Correctional Education Association and the American Correctional Association. All of these organizations and individuals contributed to the success of the negotiation process. Of the 110 eligible facilities originally selected for the study, 107 (97.3 percent) agreed to participate.

The following steps were used to gain cooperation at the sampled facilities:

- The data collection contractor mailed letters to the correctional agencies of all states in which prisons had been selected for the study. A letter was also mailed to officials at the Federal Bureau of Prisons. The letter explained the study and asked for permission to contact selected facilities within the agency's jurisdiction. Letters were followed up with telephone calls to answer questions, secure cooperation, and determine prison contact procedures. For the federal prisons and several state institutions, the study protocol and instruments had to be approved by individual institutional review boards.

- The state or federal official, in most cases, informed the warden at the sampled facility that the facility had been selected and urged the facility to participate. The data collection contractor then contacted the facility. The contractor's prison negotiator provided additional information about the study and described the sample selection process. The warden was asked to approve the study protocol and to designate a prison official to serve as coordinator for the study. The prison negotiator and the designated coordinator then worked out details, such as the interviewing procedures within the facility.

- The interviewers assigned to conduct interviews at a facility contacted the prison coordinator 2 days before the scheduled sampling date to reconfirm negotiated arrangements, including production of the inmate list for sampling, and to resolve any outstanding details.

Facility negotiations included (1) procedures for providing interviewer security within the institution and (2) interviewer clearance procedures required by the facility. Prison coordinators were asked to arrange a secure, private room for each interview. If this arrangement was not possible, interviews were conducted in partitioned or private areas of larger rooms where the inmate would not be interrupted and could be assured of confidentiality. Depending on the security regulations of each facility, respondents were either brought to the interview session by a guard or received a pass to meet with the interviewer unescorted. To minimize misinformation and deter refusals, facilities were requested to "call out" selected inmates without providing an explanation of the study. The interviewer was responsible for introducing the study and gaining inmate cooperation.

The interviewer clearance process varied from state to state and facility to facility. In three states, department of corrections officials conducted interviewer background checks and notified the facilities that the interviewers had been cleared. Elsewhere, each facility had its own requirements. In general, the interviewer's name, address, Social Security number, date of birth, and driver's license number were submitted to the facility and were processed by a recognized clearance agency. Study materials were generally reviewed by prison officials during the negotiation process. In some cases, special permission was required to allow interviewers to bring in the laptop used to administer the interview. Most facilities also required that interviewers obtain a signed informed consent form from inmates before the interview. The form included statements on confidentiality and assurances that participation or nonparticipation would not affect release or parole eligibility.

## 9.3 DATA COLLECTION MATERIALS AND INSTRUMENTS

The materials used in the conduct of the Prison Study are listed below:

- **Facility folder:** This was similar to the segment folder used in the household study (see section 8.2.2). A unique folder was developed for each facility and contained the facility scheduling sheet, facility sampling form, and sample listing sheet (described below).

- **Facility scheduling sheet:** This document contained information concerning the facility, including its location, contact information, logistical information, and security arrangements. Interviewers were required to review this form very thoroughly prior to visiting the facility because it included information that was unique to each facility. In addition to contact information, the facility scheduling sheet included details on where, when, and how inmate sampling would be conducted.

- **Facility sampling form:** This form was used in conjunction with a module in CAPI to select the inmates for participation (see section 9.4).

- **Sample listing sheet:** This sheet was used to schedule interviews and document the results of the interviews for each selected inmate. This form was provided to the facility contact to aid in scheduling the interviews and ensuring that the appropriate inmates were available at the indicated time.

- **Screener:** The screener instrument was much abbreviated from that used in the household component of the study. The Prison Study screener collected simply the sample person's first name and institutional identification (ID) number, as assigned by the facility.

- **Background questionnaire:** While essentially the same instrument as administered in the household component, some of the questions were changed, deleted, or added to adapt the interview and information gathered to the correctional environment.

- The **assessment, supplemental study**, and **oral module** were exactly the same as in the household component.

- **Certificate of participation:** Following the completion of the entire interview, inmates received a certificate of participation for their files. (In some cases, institutions did not permit the certificates; in other institutions, the certifications had to be put directly into the inmates' files.)

- **Prison noninterview report form (NIRF):** If a sampled inmate did not complete any component of the interview, the interviewer documented the reasons on the prison noninterview report form.

- **Participant folder:** Each sample person was assigned a participant folder. All contact results were recorded in the Record of Actions on the participant folder.

The materials discussed above were developed specifically for the Prison Study and the use of the materials are discussed in the subsequent sections.

## 9.4 INTERVIEWER SELECTION AND TRAINING

To conduct the data collection, 43 interviewers were recruited from among the NAAL household study workforce. Criteria for selection included proximity to sampled facilities, experience in interviewing in correctional facilities, availability, and willingness to interview in correctional facilities. The interviewer manual documented procedures specific to interviewing the prison population, including instructions for the facility contact and sampling forms, question-by-question specifications for each instrument, and instructions for reporting information. Training materials focused on the following aspects of the study:

- the background and purpose of the study, including an overview of facility negotiations (see section 9.2 for more details on gaining cooperation of the prisons);

- inmate sampling forms and procedures;

- specifications for administering the background questionnaire;
- procedures for working within correctional facilities, obtaining inmate cooperation, and reporting results of the study; and
- administrative procedures.

Interviewer training was conducted during a 2-day in-person session. Particular emphasis was given to the inmate sampling procedures, including numerous hands-on practice exercises. Interviewers also received information on security procedures and working in the correctional environment. Further details are provided in the following sections.

### 9.4.1 Sampling Procedures

A significant portion of the training program focused on the sampling procedures for randomly selecting inmates in each of the correctional facilities. Much of the training time was devoted to the procedures involved in the creation of a sampling frame, or a list of all inmates who were eligible for the study, and selection of inmates. Although the facilities provided the inmate lists (based on detailed information provided by the study recruiters), interviewers were required to verify that the lists were current and contained all eligible inmates, and in some cases, arrange the lists into a suitable sampling frame.

The inmate lists provided by the facilities differed significantly in structure, format and content, based on the facility's size and record-keeping system. Interviewers were trained on how to accommodate and work with the various list structures through numerous training exercises.

Interviewers were required to follow the steps outlined on the facility sampling form to verify that the list of inmates provided by the facility included the following:

- all inmates who had a bed assigned and slept at the facility the previous night;
- all inmates who had a bed assigned but were temporarily absent (on furlough, in court, in the hospital, or in another facility); and
- all inmates who were admitted the previous day and had a bed assigned as of the previous night.

After adding the name and/or prison ID number of any eligible inmates who did not previously appear on the list, interviewers ensured that the list did not include any of the following:

- any inmates who were released from the facility prior to the previous night;
- any inmates who ordinarily slept elsewhere (hospital, halfway house, work release center, local jail);
- any inmates who were admitted to the facility after the previous night;
- any inmates who were scheduled to be admitted but were not officially at the facility the previous night;
- any inmates who had escaped or were away without leave;
- any inmates who were sleeping, eating, working, visiting, or for any other reason in the facility the previous night but did not ordinarily have a bed assigned there;
- any inmates who were released from the facility prior to the previous night; or
- any inmates who were under the age of 16.

If any of these inmates were found to be on the list, interviewers were trained to delete their names from the list. Once the list was determined to contain all eligible inmates, it was necessary for the interviewers to conduct a final count of all inmates on the list. A few strategies were developed to assist the interviewers in the task of manually counting the inmates on the list, and these strategies were practiced during the training session. Once the final count of inmates was written on the facility sampling form, a CAPI module was utilized to select a sample of inmates for interviewing.

After the interviewer input several pieces of information regarding the facility, including the final count of inmates, the CAPI module indicated how many were sampled and their corresponding line numbers from the lists. The final sampling task for interviewers was to determine which inmate name/ID corresponded with the line number selected by the CAPI system—these were the inmates who would be interviewed. Interviewers practiced these procedures through five in-depth, realistic training exercises. After selecting the sample at the first two facilities, each interviewer was required to contact the home office sampling coordinator by telephone to review the sampling process and ensure that the proper procedures were followed before scheduling interviews with the sampled inmates. The sampling coordinator remained available to discuss questionable or problematic sampling procedures for the remainder of the field period.

### 9.4.2 Inmate Participation

The interviewer training session also covered procedures for scheduling the interviews and obtaining respondent cooperation within a correctional institution. Once the sampling procedure was completed and the sample listing sheet filled out, the sheet was provided to the facility contact who then determined the actual interview date, time, and location for each inmate and added the relevant information to the sheet. Interviews were scheduled a minimum of 2 hours apart.

Arrangements for an appropriate interview setting were made with the facility during the initial negotiation process (see section 9.2). Most interviews took place in administrative offices, attorney/client rooms, or classrooms. Facility contacts were asked not to discuss the NAAL study with the inmates prior to bringing them to the scheduled interview. Interviewers presented inmates with a letter of introduction upon arriving at the interview. The interviewer also discussed the certificate of participation and the informed consent form, as appropriate, before beginning the interview.

### 9.4.3 Background Questionnaire

The general format and structure of the Prison Study background questionnaire was identical to that used in the household component of NAAL. Many of the questions were also identical, but some were deleted and others, which pertained directly to the prison environment, were added. The entire background questionnaire instrument was covered in training, with particular emphasis given to the prison-specific items, such as questions about current offenses, criminal history, and participation in prison programs, as well as education and workforce experience.

### 9.4.4 Administrative Procedures

The final part of training focused on administrative procedures such as dismissing inmates at the end of the interview, presenting the certificate of participation, accounting for all interview materials, special security situations, recordkeeping, the assignment of result codes, and completing the prison noninterview report form. The administrative procedures and recordkeeping forms remained essentially the same as those employed in the household component, but interviewers were trained in the techniques needed to apply them to the Prison Study.

Interviewers were trained in the procedures specific to working within correctional facilities. For instance, correctional institutions had strict dress codes that interviewers had to be aware of prior to their visit. Additionally, many facilities had items they considered contraband, including food containers, tobacco products, chewing gum, and paper clips.

Security procedures were another focus of the training. Interviewers had to be prepared to undergo numerous security procedures at the facility, in addition to the background checks conducted prior to their being hired for the study. Facilities required numerous security procedures, ranging from requests for identification and use of facility visitor ID badges to physical searches such as metal detectors.

## 9.5 DATA COLLECTION

On average, 4 days were required to select the sample of inmates and administer the interviews in each facility. Two interviewers usually entered the facility on Monday morning to sample inmates and submit the list of selected inmates to the prison coordinator for scheduling interview appointments. Interviewing typically began on Tuesday and proceeded at the rate of about four interviews a day. Interviewer assignments were guided by the proximity of the interviewer's home to the facility, by the need for bilingual interviewers for Spanish-speaking inmates, and by interviewer availability.

The response rates achieved on this unique effort were quite favorable. Of the 1,298 inmates selected, 1,173 (91 percent) completed the background questionnaire.[1] The assessment booklet was completed by 1,125 inmates and partially completed by another 22 inmates. These rates are a significant achievement, especially because the interviewers had no control over the availability of selected inmates within the short data collection period at each prison and the ability of inmates to complete the assessment. Although interview appointments were scheduled by the prison coordinator, sessions were sometimes delayed, interrupted, or canceled because of inmate transfers, unscheduled inmate count-downs, facility lockups, or movement of a sampled inmate to solitary confinement, restricted housing, or a unit for the mentally ill. Prisons also changed or canceled appointments to accommodate inmate obligations, such as attorney meetings or court appearances.

---

[1] This number includes 12 inmates who did not complete the background questionnaire for literacy-related reasons, including language problems and mental disabilities.

## 9.6 QUALITY CONTROL

The measures used to ensure the collection of high-quality data included daily communications among interviewers, regional supervisors, and the home office. At the first two assigned facilities, each interviewer telephoned the home office sampling coordinator to review the sampling results immediately after completing inmate selection. Thereafter, interviewers called the coordinator if problems or questions arose during the sampling process. Any problems were referred to the data collection contractor's statistical staff before the sampling and interviews could proceed. Interviewers also reported the sampling results to their regional supervisor at the conclusion of the sampling process. Finally, after each day of interviewing, interviewers were required to contact their supervisor to discuss any special issues or concerns about the facility or inmate interview process.

Special editing specifications were prepared for the Prison Study. Editors were trained on the requirements and documents specific to the prison interviews.

*This page is intentionally left blank.*

# CHAPTER 10

# REDUCING THE RISK OF DATA DISCLOSURE

Thomas Krenzke, Sylvia Dohrmann, and Laura Alvarez-Rojas, Westat

Over the past decade, concerns about the disclosure of information related to individual survey respondents have increased. New laws have been put in place since the Privacy Act of 1974 to further ensure the protection of confidential data. The National Center for Education Statistics (NCES) and data contractors pledge confidentiality to respondents. The Education Sciences Reform Act of 2002 explicitly requires that NCES protect the confidentiality of all those responding to NCES-sponsored surveys so that no individual respondent can be identified. More specifically, NCES Standard 4-2, *Maintaining Confidentiality* (NCES 2002), provides guidelines for limiting the risk of data disclosure for data released by NCES. Data disclosure occurs when an individual respondent has been identified through the use of the survey item responses and other external data sources. This chapter describes the procedures for reducing the risk of data disclosure for the National Assessment of Adult Literacy (NAAL) in accordance with the guidelines specified in NCES Standard 4-2.

Several types of data were collected and derived during the NAAL sampling, data collection, and weighting processes. These variables were reviewed to determine their disclosure risk levels. The confidentiality analysis used a three-step process to reduce disclosure risk: (1) determining the disclosure risk arising from existing external data; (2) coarsening the data (described in section 10.1.2); and (3) swapping the data (described in section 10.3). Westat conducted the risk analysis, coarsening, and data swapping procedures to produce the following files containing data from all components of NAAL:

- Household Study public-use microdata file (PUMF);
- Prison Study PUMF;
- Household Study restricted-use microdata file (RUMF); and
- Prison Study RUMF.

Following the NCES guidelines, the RUMFs contain noncoarsened, swapped data, and the PUMFs contain coarsened and swapped data. Access to the RUMFs is restricted because users need a license to obtain the file. The PUMFs will be readily accessible to the public. The data swapping step for the RUMF was consistent with the one used for the PUMF to ensure consistency between statistics derived from the two datasets. In addition, the following confidentiality procedures were established for disseminating data through the RUMFs and PUMFs:

- identify personal identifiers, geographic information, and contextual variables (variables that can indirectly identify a geographic area);
- evaluate the existence of other publicly available files;
- evaluate the disclosure risk associated with release of the sampling and variance estimation variables; and
- evaluate the disclosure risk associated with release of key variables (i.e., visible variables) through extensive frequency tables.

Sections 10.1 and 10.2 discuss some outcomes of the risk analysis for the household sample and the prison sample, respectively. A general discussion of the data swapping procedures is provided in section 10.3.

## 10.1 HOUSEHOLD SAMPLE

One aspect of the disclosure risk analysis for the PUMF was a review of each background questionnaire variable and groups of background questionnaire variables to determine whether any of the data presented a nonnegligible risk of individual disclosure. Several types of variables were available from the household sample and were analyzed for disclosure risk. These included variables collected through the survey and assessments, as well as variables created during weighting. These variables are summarized below:

- case identifiers;
- disposition codes for the survey instruments: screener, background questionnaire, assessment, and oral module;
- demographics: age, race/ethnicity, and gender;
- variance stratum and variance unit;
- sampling weights: weights from all stages of weighting, including the base weight, nonresponse-adjusted weight, trimmed weight, and final weight;
- weight adjustment factors, including compositing factors;
- number of eligible persons in the household and number selected;
- weighting variables, such as age category, census region, metropolitan statistical area identifier, educational attainment, race/ethnicity, and country of birth, and including imputation flags for educational attainment and country of birth;
- Background questionnaire data responses, including census codes for the respondent's industry and occupation;

- other background questionnaire variables, such as start and end times of the interview; and

- assessment scores.

Careful attention was given to these background questionnaire items and combinations of items. Even a very limited amount of demographic detail—such as income, occupation, age, year of immigration to the United States, foreign language spoken, and country of birth—can increase the chance that an individual can be identified. As discussed in section 10.1.1, personal and geographic identifiers were removed. Section 10.1.2 presents outcomes from the risk analysis in the form of variable suppression and recodes.

### 10.1.1 Personal Identifiers and Geographic Identifiers

Any information that might be used to directly identify respondents and/or sampled locations was suppressed from the PUMF. This information included direct personal identifiers such as names (only first names were collected), addresses, and telephone numbers. Explicit geographic identifiers, such as state or county, were also suppressed, with the exception of census region and state identifiers for the State Assessment of Adult Literacy (SAAL). The SAAL states were Kentucky, Maryland, Massachusetts, Missouri, New York, and Oklahoma. The primary sampling unit (PSU) identification number (ID) was also suppressed. Further, because the case identifiers assigned during sample design, sample selection, and data collection had embedded geographic identifiers, the original case IDs were replaced with a sequential number to mask any pattern (e.g., alphabetic or geographic order).

### 10.1.2 Data Coarsening

In general, data coarsening includes several types of procedures that decrease disclosure risk by reducing the amount of information released. Coarsening approaches include removing direct identifiers, limiting geographic detail, categorizing continuous variables, performing top- and bottom-coding,[1] and recoding values into broader categories. Weight adjustment factors and intermediate weights were removed because they posed a disclosure risk and provided minimal analytical value. Targeted or local suppression was also performed by removing the sensitive item value from the record or suppressing (or deleting) the variable from the file. During the NAAL data-coarsening step, some variables with high disclosure risk were suppressed or recoded; the original data values were retained for low-risk variables. The swapping procedure was used to add uncertainty to the otherwise individual identifying variables.

[1] With top-coding, the largest values of a variable are replaced with an upper limit, reducing the appearance of outlier data. Similarly, bottom-coding replaces the smallest values with a lower limit.

Variables with potential for high risk of disclosure were those known facts about individuals as collected by NAAL. These include income, race, occupation, and personal and geographic identifying variables. Variables with high disclosure risk that could not be recoded further were suppressed from the PUMF. The information contained in these variables was too specific to be released and could not be effectively recoded. Such variables include, for example, text responses to other-specify items. These types of items may have high risk of disclosure.

As described in the remainder of this section, seven major classes of variables were coarsened: age, race/ethnicity, language, education, income, occupation/industry, and all others. The recodes are a result of many cross-tabulations between background questionnaire variables and demographic/geographic variables. Of primary concern were cases that were rare in the population. Data swapping was also implemented to further protect the identity of the individual.

**Age.** The respondent's date of birth, collected on the background questionnaire, was converted to a single year of age. Whenever age was missing, it was reconciled with the screener data (the screener asked for a single year of age for each person in the household). The screener had a backup question that requested the response in age categories. These screener age categories were used to impute a single year of age for respondents still missing an age; the single year of age was released on the RUMF. The single years were categorized as shown in table 10-1 for the PUMF. Other age-related recoded variables are also shown in table 10-1.

**Table 10-1. Source, description, and categories of age-related recoded variables for household sample: 2003**

| Variable name | Source variables | Variable description | Categories of recoded variable |
|---|---|---|---|
| DARRIVE | BQ1029 | Immigration age | U.S. born; 0–18; 19 or more years |
| DLIVEUS | BQ1030 | Years lived in United States | Combined highest age categories: 1–5; 6 or more years |
| DENGAGE | BQ1065 | Age when respondent spoke English | 1–10; 11 or more years; does not speak English |
| DAGE | CALCAGE | Age calculated from date of birth | Included screener data (where background questionnaire item was missing) to recode into categories: 16–18; 19–24; 25–39; 40–49; 50–64; 65 or more years |
| DHSAGE | BQ1032, BQ1205, BQ1210 | Age when respondent graduated from high school or obtained General Educational Development (GED) credential | 16–19; 20 or more years; not applicable |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Race/ethnicity.** Table 10-2 contains the race/ethnicity variables and the recodes. The recode for respondent race/ethnicity used screener data when background questionnaire items were missing.

**Table 10-2. Source, description, and categories of race/ethnicity recoded variables for household sample: 2003**

| Variable name | Source variables | Variable description | Categories of recoded variable |
|---|---|---|---|
| DCBIRTH | BQ1025, BQ1025S | Country where respondent was born | United States; other |
| DMCBIRTH | BQ1875, BQ1875S | Country where respondent's mother was born | United States; other |
| DFCBIRTH | BQ1885, BQ1885S | Country where respondent's father was born | United States; other |
| DRACE | BQ2440, BQ2445_a-e, BQ2450_a-e | Race/ethnicity | White, Black, Hispanic, other (including multiracial) |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Language.** Table 10-3 shows the language variables and the method of coarsening. All items that allow several responses, including other-specify variables (i.e., variables containing text responses not contained in the response categories provided), were coded into five categories:

- English only;
- English and Spanish (with or without other);
- English and other;
- Spanish only or with other; and
- other only.

This coding scheme preserved the most frequent responses (English and Spanish) while capturing multilingual respondents and those who spoke no English. Recoding for the variable indicating other languages the respondent speaks now includes the language he or she speaks best.

**Table 10-3. Source, description, and categories of language recoded variables for household sample: 2003**

| Variable name | Source variables | Variable description | Categories of recoded variable |
|---|---|---|---|
| DHMLANG | BQ1045a-e, BQ1045Sa-e | Language spoken at home | Five categories: English only; English and Spanish (with or without other); English and other; Spanish only or with other; other only |
| D1STLAN | BQ1050a-e, BQ1050Sa-e | Language spoken before school | Five categories: English only; English and Spanish (with or without other); English and other; Spanish only or with othe; other only |
| DLANGRW | BQ1060, BQ1060S | Language respondent first learned to read and write | English; Spanish; other |
| DCLANGS | BQ1090, BQ1090S | Language respondent usually speaks now | English; Spanish; other |
| DOLSOPT | BQ1095a-e, BQ1095Sa-e | Other language often spoken now | Incorporated BQ1090 and recoded into five categories: English only; English and Spanish (with or without other); English and other; Spanish only or with other; other only |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Education.** Table 10-4 shows education variables and the categories of their recodes. The recode for educational attainment consists of nine categories:

- still in high school;
- less than high school/some high school;
- General Educational Development (GED)/high school equivalency;
- high school graduate;
- vocational/trade/business school;
- some college;
- associate's/2-year degree;
- college graduate; and
- graduate studies/degree.

**Table 10-4. Source, description, and categories of education recoded variables for household sample: 2003**

| Variable name | Source variable | Variable description | Categories of recoded variable |
|---|---|---|---|
| DEDBFUS | BQ1040 | Education before coming to the United States | Did not attend school/primary, elementary, secondary or more |
| DEDATTN | BQ1205 | Highest education level attained | Still in high school (HS); less than HS/some HS; General Educational Development/HS equivalency; HS graduate; vocational; some college; associate's degree; B.S. or B.A.; graduate studies/degree |
| DDTYPE | BQ1225 | Diploma type | Regular diploma from school in United States or U.S. government school outside United States; regular diploma from school outside United States; General Educational Development credential or certificate of completion; combined "Did not receive HS diploma" with missing |

See note at end of table.

**Table 10-4. Source, description, and categories of education recoded variables for household sample: 2003—Continued**

| Variable name | Source variable | Variable description | Categories of recoded variable |
|---|---|---|---|
| DRFSSCHC | BQ1235 | Reason respondent stopped school | Financial problems; did not do well in school; did not like school or was bored in school; was expelled from school or asked to leave; wanted to work; wanted to go into the military; personal reasons; other |
| DSGRDHS | BQ1245 | State in which respondent graduated high school | Current; other |
| DSGRDCO | BQ1257 | State in which respondent received college degree | Current; other |
| DMED | BQ1880 | Mother's highest educational attainment | Less than HS/some HS; General Educational Development credential/HS equivalency; HS graduate; vocational; some college; associate's degree; B.S. or B.A.; graduate studies/degree |
| DFED | BQ1890 | Father's highest educational attainment | Less than HS/some HS; General Educational Development credential/HS equivalency; HS graduate; vocational; some college; associate's degree; B.S. or B.A.; graduate studies/degree |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Income.** The household background questionnaire included several income-related questions; these included some weekly wage amounts, as well as income amounts for the entire year. Using the variables in table 10-5, variables were derived for the weekly wage during the previous week, income adequacy, personal income, and household income.

Data on total personal and household income were gathered through a series of questions. First, the respondent was asked one question with 13 categories (14 in the case of household income). If the respondent refused that question, he or she was asked a series of questions designed to categorize income into one of eight categories. The income variable was created in the eight categories to use all of the information possible.

**Table 10-5. Source, description, and categories of income-related recoded variables for household sample: 2003**

| Variable name | Source variables | Variable description | Categories of recoded variable |
|---|---|---|---|
| D_WKLYWAGE | BQ1500,<br>BQ1505,<br>BQ1505S,<br>BQ1515,<br>BQ1520 | Weekly wage (previous week) | Less than $300;<br>$300–$499;<br>$500-$649;<br>$650-$1,149;<br>$1,150-$1,949;<br>greater than or equal to $1,950 |
| DINCOME | BQ1895,<br>BQ1910a-j,<br>BQ1920a-j,<br>BQ2430 | Income adequacy | Below poverty threshold; other |
| DBQ2421 | BQ2421,<br>BQ2422,<br>BQ2423,<br>BQ2424,<br>BQ2425,<br>BQ2426,<br>BQ2427,<br>BQ2428 | Approximate personal income | Less than $5,000;<br>$5,000–$9,999;<br>$10,000–$14,999;<br>$15,000–$19,999;<br>$20,000–$29,999;<br>$30,000–$39,999;<br>$40,000–$59,999;<br>$60,000 or more |
| DBQ2430 | BQ2430,<br>BQ2432,<br>BQ2433,<br>BQ2434,<br>BQ2435,<br>BQ2436,<br>BQ2437,<br>BQ2438 | Approximate household income | Less than $10,000;<br>$10,000–$14,999;<br>$15,000–$19,999;<br>$20,000–$29,999;<br>$30,000–$39,999;<br>$40,000–$59,999;<br>$60,000–$99,999;<br>$100,000 or more |
| DWFTIME | BQ2155 | Length of participation in welfare programs | Less than 2 years; 2 or more years |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Occupation and industry codes.** Each respondent was asked to specify his or her occupation and industry. These responses were then categorized into the standard four-digit values used by the U.S. census. These detailed codes had a high disclosure risk. However, the census also uses some standard combinations of two-digit classifications (such as combining all codes beginning with 01 with all codes beginning with 02). Using these classifications reduced the risk of data disclosure.

Other derived variables are provided in appendix A.

## 10.2 PRISON SAMPLE

As for the household sample, the statistical disclosure control procedures for the prison sample were based on a disclosure risk analysis. After the sources of disclosure risks had been identified, variable suppression, coarsening, and swapping procedures were applied to reduce such risks. The application of these techniques followed NCES guidelines, specifically NCES Standard 4-2, *Maintaining Confidentiality* (NCES 2002).

All variables collected or derived through the sampling, data collection, and weighting process were compiled. The following list shows the variables available and analyzed for data disclosure risk:

- case identifiers;
- disposition codes for the background questionnaire, assessment, and oral module;
- variance stratum and variance unit;
- prison-level variables: census region (Northeast, Midwest, South, West); type of prison (federal, state, private); gender composition of the prison (male only, female only, mixed); and security type (supermaximum, maximum, medium, minimum, administrative, other);
- inmate full sample weight and replicate weights;
- adjustment factors and intermediate weights (prison base weight, prison nonresponse-adjusted weight, inmate base weight); and
- weighting variables, such as imputation flags for weighting purposes and raking dimensions.

The background questionnaire collected the following data responses:

- Section A: General and Language Background;
- Section B: Educational Background and Experiences;
- Section C: Prison Experiences;
- Section D: Prison Work Assignments and Labor Force Participation;
- Section E: Political and Social Participation;
- Section F: Literacy Practices;
- Section G: Demographic Information;
- Section H: Household Income and Welfare Participation;

- Section I: Health Questions; and
- Section J: Additional Demographics (race/ethnicity variables).

After the variables had been compiled, personal identifiers and geographic information were treated as discussed in section 10.2.1. Weight adjustment factors and intermediate weights were removed because they posed a disclosure risk and provided minimal analytical value. The ability to match with administrative data and other external data was assessed, as discussed in section 10.2.2. Data coarsening procedures are discussed in section 10.2.3, and section 10.3 provides a general discussion of data swapping.

### 10.2.1 Personal Identifiers and Geographic Information

Any information that could be used to directly identify persons and/or prisons was suppressed from the file. This information includes the following:

- Direct and proximate identifiers:
  - direct identifiers of prisons such as names, addresses, and telephone numbers and names of inmates.
- Explicit geographic information:
  - Explicit geographic identifiers (except census region).

Further, because the original numerical IDs provide key identifying information, the case IDs were replaced by a sequential number to mask any pattern, such as alphabetic or geographic order.

### 10.2.2 Ability to Match With Administrative Data and Other External Data

The Bureau of Justice Statistics 2000 Census of State and Federal Adult Correctional Facilities (referred to as the Census) is a prison-level data source for which data are publicly available and which presents a potential risk of disclosure. The Census includes more than 1,600 facilities. The Census data include address, capacity, inmate population, and security level, all of which are important characteristics for sampling and data collection. A record linkage analysis was conducted but did not show a risk of disclosure from matching to the Census data. "Record linkage analysis" is a technical term relatively well known among researchers of disclosure risk. Skinner and Elliott (2002) provide a good discussion of record linkage techniques.

### 10.2.3 Data Coarsening

The disclosure analysis for the PUMF included a review of each background questionnaire variable to determine whether any of the proposed data presented a nonnegligible risk of individual disclosure. For the disclosure risk analysis, frequencies were processed. If the number of observations for a category was low, then the variables were generally recoded into broader classifications or were suppressed. Once the initial coarsening process was complete, cross-tabulations between key reporting variables and several identifiable variables (education, race/ethnicity, age, and other background questionnaire items) were processed. The cross-tabulations further determined the risk level of each variable. The analysis of the cross-tabulations identified more recodes, and decisions were made about the inclusion of certain high-risk variables on the file. Concerns about some of the variables with a high risk of disclosure were reduced through variable suppression. For these variables, the risk of disclosure could not be reduced further by recoding the data.

As shown in table 10-6, year of immigration to the United States, the respondent's age, and the language variables are among the variables that were reclassified. The industry and occupation codes proposed for the household sample were used for the prison sample.

**Table 10-6. Source, description, and categories of recoded variables for prison sample public-use microdata file: 2003**

| Variable name | Source variables | Variable description | Categories of recoded variables |
|---|---|---|---|
| DAGEC | CALCAGE | Age calculated from date of birth | Age: 16–24; 25–39; 40–49; 50 or more years |
| DCBIRTH | BQ1015 | Country of birth | United States (50 states and DC); other |
| DARRIVE | BQ1020 | Immigration age | U.S. born; 0–18; 19 or more years |
| DEDBFUS | BQ1050 | Education before coming to United States | Did not attend school/primary, elementary, secondary or more |
| DHMLANGC | BQ1055a-e | Language spoken at home when growing up | English only; English and other (including Spanish); other only |
| D1STLANG | BQ1060a-e | Language spoken before school | English only; English and other (including Spanish); other only |
| DLANGRWC | BQ1070 | Language respondent first learned to read and write | English; other |

See note at end of table.

**Table 10-6. Source, description, and categories of recoded variables for prison sample public-use microdata file: 2003—Continued**

| Variable name | Source variables | Variable description | Categories of recoded variables |
|---|---|---|---|
| DENGAGE | BQ1075 | Age respondent learned to speak English | 1–10; 11 or more years; do not speak English |
| DCLANGSC | BQ1100, BQ1105a-e | Language respondent usually speaks now | English, other; English only; English and other; other only |
| DOLANGSB | BQ1110 | Other language respondent speaks best | Spanish; other |
| DEDATTNC | BQ1205, BQ1208, BQ1215 | Educational attainment | Less than high school (HS)/some HS; General Educational Development credential or HS equivalency; HS graduate; vocational, trade, or business school after HS; less than 2 years of college; associate's degree; college or more |
| DHSAGE | BQ1030, BQ1205, BQ1220 | Age upon graduating from high school | 16–19; 20 or more years; not applicable |
| DDTYPEC | BA1224 | Type of high school degree | Regular diploma from school in United States, or U.S. government school outside United States; General Educational Development credential or certificate of completion; "Did not receive HS diploma"; "Regular from school outside U.S."; missing |
| DRFSSCHC | BQ1230 | Reason for stopping school before college degree | Financial problems; did not do well in school; did not like school or was bored in school; expelled from school or asked to leave; wanted to work; wanted to go into the military; personal reasons; sent to jail or detention or prison; other |
| DPVOC | BQ1421, BQ1423 | Length of time in prison Vocational training programs | Less than 1 year; 1 or more years |
| DPCLSHR | BQ1450 | How many hours spent in prison classes last week | 0, 1–19; 20–49; 50 or more hours |
| DOFFENS1-3 | BQ1475a-e | Offenses for which respondent is in prison | Violent; property; drugs; public order; other |

See note at end of table.

**Table 10-6. Source, description, and categories of recoded variables for prison sample public-use microdata file: 2003—Continued**

| Variable name | Source variables | Variable description | Categories of recoded variable |
|---|---|---|---|
| DEMPTYPC | BQ1620 | Type of employer in past 3 years | Private employer; self-employed; other |
| MEDC | BQ1830 | Mother's educational attainment | Less than HS/some HS; General Educational Development credential or HS equivalency; HS graduate; vocational, trade, or business school after HS; less than 2 years of college; associate's degree; college or more |
| FEDC | BQ1840 | Father's educational attainment | Less than HS/some HS; General Educational Development credential or HS equivalency; HS graduate; vocational, trade, or business school after HS; less than 2 years of college; associate's degree; college or more |
| DMCBIRTH | BQ1825 | Mother's country of birth | United States; other |
| DFCBIRTH | BQ1835 | Father's country of birth | United States; other |
| DMARITAL | BQ1845 | Marital status | Never married; married or living as married; separated or divorced or widowed |
| DWLFLSTC | BQ1880 | Last received welfare payments | 3 years or less; more than 3 years |
| DRACE | BQ2000, BQ2010a-e | Race/ethnicity | White; Black; Hispanic; other (including multiracial) |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

## 10.3 DATA SWAPPING

To ensure that an individual respondent is not identified, the risk of data disclosure was further reduced by using a data swapping technique. Data swapping is an NCES requirement that reduces risk by modifying microdata. In data swapping, a probability sampling of records are paired with other records on the file using selected characteristics, and then some identifying variables are swapped between the two records (refer to Kaufman et al. [2005] for further discussion). The sampling rate for NAAL swapping was designed to protect the confidentiality of the data without affecting the usability of the dataset. This method is an effective way of keeping as much valuable data as possible while protecting respondent identity.

Swapping preserves the univariate frequencies, means, and variances, although it may affect multivariate relationships. Pre- and post-swapping percentage distributions (unweighted and weighted) and correlations were reviewed to ensure data quality was maintained.

*This page is intentionally left blank.*

# CHAPTER 11

## RESPONSE RATES AND NONRESPONSE BIAS ANALYSIS

Thomas Krenzke and Leyla Mohadjer, Westat

This chapter provides weighted response rates and a systematic analysis of the potential for nonresponse bias for the household sample and the prison sample of the National Assessment of Adult Literacy (NAAL), separately. The analyses focus on the impact of nonresponse on survey results.

Total survey error has two components: variable error (measured through the calculation of variances) and bias. The variance is the first term in the following equation for total survey error in a survey estimate:

$$\text{Total survey error} = \text{variance} + \text{bias}^2. \qquad (1)$$

Bias, the second term in the equation, contains all sources of error other than variable error. A major component of bias is nonresponse, that is, the bias owing to the failure of some selected persons in the sample to respond to the survey. Nonresponse bias can be substantial when two conditions hold: (1) the response rate is relatively low and (2) the difference between the characteristics of respondents and nonrespondents is relatively large.

An estimate for nonresponse bias, assuming that nonresponse is the only source of bias, is expressed in Cochran (1977) as

$$Bias(\bar{y}_R) = (1 - W_R)(\bar{Y}_R - \bar{Y}_N), \qquad (2)$$

where $W_R$ is the response rate and $\bar{Y}_R$ and $\bar{Y}_N$ are the mean values of the survey items estimated among the respondents and nonrespondents, respectively. Thus, the estimates from any survey are subject to bias when some selected persons fail to participate in the survey. Because we do not have survey values for nonrespondents, nonresponse bias is not known and can only be estimated.

The following sections provide insights into the effects of nonresponse on the NAAL survey. The unweighted and weighted unit and item response rates are provided for the household study (section 11.1) and the correctional institution sample, known as the prison study (section 11.2). Unweighted response rates are indicators of the success of the data collection effort. Weighted response rates are more appropriate in examining the potential effect of nonresponse on population parameters. Bivariate and multivariate analyses of the potential for nonresponse bias are provided for both the household study and the prison study.

## 11.1 HOUSEHOLD LITERACY STUDY

Data from respondents were collected through a screener, a background questionnaire, an assessment, and an oral module. In the nonresponse follow-up strategies, efforts were made to reduce the potential for nonresponse bias by targeting interviewer resources in areas with low response rates. To identify target areas, a multivariate analysis was conducted using a Chi-Squared Automatic Interaction Detector (CHAID) analysis (for more information on CHAID, refer to section 11.1.4.1.2). The resulting classification tree revealed the domains, as defined by combinations of variables, with the most differential response rates, thereby leading to domains with a high potential for nonresponse bias. Overall, the results of the analysis showed acceptable response rates for most of the cells identified by the CHAID program. The analysis, which was conducted for both the screener and the combined background questionnaire/assessment response rates, identified the primary sampling units (PSUs) that included the domains with less than a 70 percent response rate. Field activities and resources were focused on these PSUs in the remaining weeks of the data collection.

After data had been collected and weights produced, a systematic analysis was conducted to examine the impact of bias owing to the remaining nonresponding dwelling units and persons in the household sample. The sections that follow report on the nonresponse bias analysis. Section 11.1.1 gives an overview of the analysis weights. Section 11.1.2 provides unweighted and weighted response rates at the unit level. Section 11.1.3 summarizes response rates at the item level. Section 11.1.4 provides a detailed nonresponse bias analysis for the household study.

### 11.1.1 Analysis Weights

The systematic analysis of nonresponse bias in the household sample, which includes the computation of weighted response rates, used survey weights that were specially created for this analysis. The remainder of this section provides a brief overview of the weighting process for the household study, followed by a brief overview of the nonresponse bias weights used in this analysis.

In the NAAL household study, the nonresponse-adjusted weights were created separately for each of the seven independent samples (NAAL and the six State Assessment of Adult Literacy [SAAL] states). A composite weighting procedure was conducted to combine the NAAL and SAAL samples to improve the survey estimates for the nation and the six SAAL states. More details are provided in section 12.1.

National Center for Education Statistics (NCES) standards for a nonresponse bias analysis require the use of base weights. Because composited screener (dwelling unit-level) and background questionnaire base weights were not created as part of the NAAL household study weighting process, composited base weights were created for this nonresponse bias analysis. An objective of the weight modification process used in the nonresponse bias analysis (as required by the standards) was to minimize the effects of the nonresponse adjustments carried out for the final NAAL weights. For the nonresponse bias analysis, a minor adjustment was made to the NAAL and SAAL screener base weights and replicate weights for sample persons whose eligibility status was unknown (e.g., those who were unavailable after multiple attempts during the field period), in order to represent the eligible population only. Next, the weights were poststratified to one number for each sample (six SAAL states and all other states combined), and then compositing factors from the NAAL household study weighting process were applied to the screener weights to combine the state samples. The resulting composited screener weights were used in the nonresponse bias analysis at the screener level. The background questionnaire base weights for the nonresponse bias analysis were computed by applying the within-household sampling fraction to the nonresponse bias screener base weights.

### 11.1.2 Unit Response Rates

NAAL had four stages of data collection where unit nonresponse occurred: the screener, background questionnaire, assessment, and oral module. Both unweighted and weighted response rates were computed for each stage. Screener composited base weights (discussed in section 11.1.1) were used in the screener response rate calculations, and background questionnaire composited base weights were used for the background questionnaire, assessment, and oral module calculations.

Response rates were calculated as follows:

$$RR = \frac{\sum_{i \in SR} W_i}{\sum_{i \in SR \cup SNR} W_i},$$

where

$W_i$ = the weight of unit $i$;

$SR$ = the set of participating units; and

$SNR$ = the set of eligible nonparticipating units.

Table 11-1 contains response rates for each stage and for the survey overall. The weighted response rates are 82 percent, 76 percent, 97 percent, and 95 percent for the screener, background questionnaire, assessment, and oral module, respectively. The overall weighted response rate—the product of the screener, background questionnaire, and assessment response rates—is 60 percent. Table 11-1 also shows response rates by selected analysis variable domains (defined later in table 11-14). As the table shows, there are differential response rates among subgroups. For instance, the weighted overall survey response rate in the Northeast is 55 percent, compared with 60 to 62 percent in the other census regions.

### 11.1.3 Item Nonresponse

Item response rates were computed for all 361 items in the background questionnaire. There were numerous reasons for item nonresponse: The respondent did not know the answer to the item or did not wish to respond, or the interview was terminated before completion and items in the latter part of the questionnaire were not asked. The numerator of the response rate consists of all item respondents; the denominator contains all unit respondents, excluding those for whom the item was skipped by the computer-assisted personal interviewing (CAPI) instrument because it was not applicable. This approach is consistent with NCES standard 1-3-5 (U.S. Department of Education, National Center for Education Statistics 2002); that is, item response rates were computed among persons who were asked the question. Westat computed both unweighted item response rates and response rates weighted with background questionnaire composited base weights.

**Table 11-1. Household Study unweighted (UW) and weighted (W) unit response rates, by analysis variable, in percent: 2003**

| Analysis variable | Screener | | Background questionnaire | | Assessment | | Oral module | | Overall[1] | |
|---|---|---|---|---|---|---|---|---|---|---|
| | UW | W | UW | W | UW | W | UW | W | UW | W |
| Total | 81.8 | 82.2 | 78.1 | 75.6 | 97.2 | 96.7 | 95.1 | 94.6 | 62.1 | 60.1 |
| Region | | | | | | | | | | |
| Northeast | 74.0 | 75.6 | 76.3 | 74.8 | 97.4 | 97.1 | 95.1 | 95.2 | 55.0 | 54.9 |
| Midwest | 82.5 | 82.9 | 80.3 | 78.1 | 96.5 | 95.7 | 94.3 | 93.8 | 64.0 | 62.0 |
| South | 84.2 | 83.4 | 78.7 | 74.8 | 98.0 | 98.3 | 96.5 | 96.6 | 64.9 | 61.3 |
| West | 85.2 | 84.7 | 76.4 | 74.6 | 95.2 | 94.9 | 92.2 | 91.9 | 62.0 | 60.0 |
| MSA[2] status | | | | | | | | | | |
| Non-MSA | 85.3 | 85.1 | 80.6 | 79.2 | 97.4 | 97.0 | 95.2 | 94.3 | 67.0 | 65.4 |
| MSA | 81.1 | 81.4 | 77.5 | 74.6 | 97.1 | 96.6 | 95.1 | 94.7 | 61.0 | 58.7 |
| Average household size | | | | | | | | | | |
| 2.42 or less | 80.0 | 80.5 | 79.4 | 77.1 | 97.5 | 97.4 | 95.6 | 95.3 | 61.9 | 60.5 |
| 2.43–2.80 | 81.8 | 81.8 | 77.8 | 75.1 | 97.3 | 96.9 | 95.4 | 95.1 | 62.0 | 59.5 |
| Greater than 2.80 | 83.8 | 84.5 | 77.3 | 74.7 | 96.7 | 95.9 | 94.4 | 93.6 | 62.6 | 60.5 |
| Percent with less than high school education | | | | | | | | | | |
| 10.4 or less | 77.6 | 78.2 | 75.5 | 73.2 | 97.4 | 96.6 | 95.7 | 94.9 | 57.0 | 55.3 |
| 10.5–20.3 | 80.9 | 81.8 | 76.8 | 74.6 | 96.9 | 96.7 | 94.9 | 94.6 | 60.2 | 59.0 |
| 20.4–32.0 | 83.1 | 84.7 | 78.7 | 77.1 | 97.1 | 96.7 | 94.9 | 94.4 | 63.5 | 63.1 |
| Greater than 32.0 | 86.5 | 87.3 | 82.0 | 79.2 | 97.3 | 97.0 | 94.9 | 94.5 | 69.0 | 67.1 |
| Percent speaking Spanish but not English | | | | | | | | | | |
| 0 | 80.8 | 81.0 | 78.0 | 75.6 | 97.3 | 96.8 | 95.5 | 95.0 | 61.3 | 59.3 |
| 1–28 | 81.5 | 81.7 | 76.8 | 74.4 | 97.1 | 96.6 | 95.2 | 94.5 | 60.7 | 58.7 |
| Greater than 28 | 84.0 | 85.6 | 80.1 | 77.4 | 97.1 | 96.8 | 94.4 | 94.1 | 65.3 | 64.2 |
| Percent below 150 percent of poverty | | | | | | | | | | |
| 10.7 or less | 78.0 | 78.6 | 74.4 | 72.0 | 96.9 | 96.2 | 95.2 | 94.6 | 56.3 | 54.4 |
| 10.8–20.0 | 79.7 | 80.8 | 76.6 | 75.3 | 97.1 | 96.7 | 95.2 | 94.8 | 59.3 | 58.8 |
| 20.1–33.3 | 83.6 | 84.9 | 79.2 | 77.3 | 96.9 | 96.7 | 94.6 | 93.9 | 64.2 | 63.4 |
| Greater than 33.3 | 87.0 | 88.3 | 82.9 | 81.0 | 97.7 | 97.8 | 95.5 | 95.5 | 70.4 | 70.0 |
| Median income (in dollars) | | | | | | | | | | |
| 28,400 or less | 87.4 | 88.5 | 83.5 | 81.7 | 97.8 | 97.7 | 95.6 | 95.3 | 71.3 | 70.6 |
| 28,401–37,850 | 83.9 | 84.7 | 78.9 | 77.8 | 97.0 | 96.9 | 94.7 | 94.4 | 64.2 | 63.9 |
| 37,851–52,100 | 79.5 | 80.4 | 76.7 | 74.8 | 97.1 | 96.7 | 95.3 | 94.9 | 59.2 | 58.2 |
| Greater than 52,100 | 77.5 | 78.4 | 73.9 | 71.2 | 96.8 | 96.0 | 94.9 | 94.2 | 55.4 | 53.6 |

See notes at end of table.

**Table 11-1. Household Study unweighted (UW) and weighted (W) unit response rates, by analysis variable, in percent: 2003—Continued**

| Analysis variable | Screener | | Background questionnaire | | Assessment | | Oral Module | | Overall[1] | |
|---|---|---|---|---|---|---|---|---|---|---|
| | UW | W | UW | W | UW | W | UW | W | UW | W |
| Percent who rent | | | | | | | | | | |
| 16 or less | 79.6 | 80.6 | 74.8 | 72.7 | 96.8 | 96.0 | 94.9 | 94.1 | 57.6 | 56.2 |
| 17–31 | 83.2 | 83.4 | 77.9 | 75.9 | 97.0 | 97.0 | 95.0 | 94.8 | 62.9 | 61.4 |
| 32–59 | 83.6 | 82.6 | 79.7 | 76.9 | 97.6 | 97.4 | 95.8 | 95.6 | 65.0 | 61.9 |
| Greater than 59 | 81.2 | 82.5 | 81.0 | 79.3 | 97.3 | 96.8 | 94.8 | 94.3 | 63.9 | 63.3 |
| Age (years) | | | | | | | | | | |
| 16–29 | — | — | 82.3 | 80.5 | 98.0 | 97.8 | 96.8 | 96.6 | — | — |
| 30–49 | — | — | 77.1 | 73.9 | 97.1 | 96.6 | 95.2 | 94.5 | — | — |
| 50–69 | — | — | 77.0 | 73.9 | 96.7 | 96.3 | 94.6 | 94.1 | — | — |
| 70+ | — | — | 75.4 | 74.2 | 96.3 | 95.5 | 91.8 | 91.3 | — | — |
| Gender | | | | | | | | | | |
| Male | — | — | 75.3 | 72.7 | 96.5 | 96.0 | 94.2 | 93.6 | — | — |
| Female | — | — | 80.4 | 78.2 | 97.7 | 97.4 | 95.9 | 95.5 | — | — |
| Race/ethnicity | | | | | | | | | | |
| Hispanic | — | — | 81.0 | 79.4 | 96.9 | 96.6 | 94.4 | 94.2 | — | — |
| Non-Hispanic Black only | — | — | 81.0 | 79.3 | 97.9 | 98.0 | 96.0 | 96.3 | — | — |
| Other[3] | — | — | 76.6 | 74.3 | 97.0 | 96.5 | 95.1 | 94.5 | — | — |

— Not available.
[1] Overall response rate is the product of the screener, background questionnaire, and assessment response rates.
[2] Metropolitan Statistical Area.
[3] Includes non-Hispanic White, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races.
NOTE: The following are segment-level variables derived from block group data from Census 2000: average household size, percent with less than high school education, percent speaking Spanish but not English, percent below 150% of poverty, median income, and percent who rent. The following are person-level variables captured during the screening: age, gender, race/ethnicity.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

In the background questionnaire, the interviewer asked income items in several formats, first asking for income in 1 of 13 categories and then, if the sample person refused, asking a series of questions involving broader income classifications, with the goal of placing the sample person in one of eight income categories. For the computation of response rates, the eight personal income items were combined into one personal income variable, and the eight household income items were combined into one household income variable. A sample person is considered a respondent to the income item if he or she can be put into one of the eight income categories on the basis of the series of questions. The overall personal income weighted response rate is 92.7 percent, and the overall household income weighted response rate is 88.3 percent. Across all background questionnaire items, weighted item response rates range from 87.9 percent to 100 percent, with a median of 99.9 percent.

### 11.1.4 Nonresponse Bias Analysis

The analysis in this section is in accordance with NCES Standard 4-4. Standard 4-4-1 states that "any survey stage of data collection with a unit or item response rate less than 85 percent must be evaluated for the potential magnitude of nonresponse bias before the data or any analysis using the data may be released." As described in section 11.1.3, all items had a response rate of more than 85 percent, so an item nonresponse bias analysis was not carried out. Two data collection stages had weighted unit response rates below 85 percent: the Screener at 82 percent and the background questionnaire at 76 percent (see section 11.1.2). Section 11.1.4.1 presents the nonresponse bias analysis for the screener, and section 11.1.4.2 provides the nonresponse bias analysis for the background questionnaire.

#### 11.1.4.1 Evaluating Bias Owing to Screener Nonresponse

A comparison of screener respondents and nonrespondents using variables known for both groups provides some indication of the potential for nonresponse bias in resulting survey estimates. The variables selected for the screener nonresponse bias analysis are displayed in table 11-2 and come from two sources: Census 2000 Public Law (PL) 94 county-level data and Census 2000 Summary File 3A (SF3A) block group-level data. The continuous variables from the SF3A were recoded into categories of approximately equal sample size.

Section 11.1.4.1.1 describes chi-square tests that may detect a significant relationship between response indicator and the analysis variable of interest. It also includes calculations of bias used in estimating the distribution of analysis variables. Section 11.1.4.1.2 provides a multivariate analysis of the relationship between response indicator and analysis variables that may reveal the areas with the greatest potential for bias before weighting adjustments. Finally, section 11.1.4.1.3 shows the effect of the weighting adjustments on the potential for nonresponse bias.

**Table 11-2. Household Study variables used in screener nonresponse bias analysis, by source and values: 2003**

| Variable description | Source[1] | Values |
|---|---|---|
| Region | PL-94 | 1: Northeast<br>2: Midwest<br>3: South<br>4: West |
| MSA[2] status | PL-94 | 1: MSA<br>2: Non-MSA |
| Average household size | SF3A | 1: 2.42 or less<br>2: 2.43–2.80<br>3: Greater than 2.80 |
| Percent aged 25+ with less than high school education | SF3A | 1: 10.4 or less<br>2: 10.5–20.3<br>3: 20.4–32.0<br>4: Greater than 32.0 |
| Percent aged 5–64 speaking Spanish at home and English not well or not at all | SF3A | 1: 0<br>2: 1–28<br>3: Greater than 28 |
| Percent below 150 percent of poverty | SF3A | 1: 10.7 or less<br>2: 10.8–20.0<br>3: 20.1–33.3<br>4: Greater than 33.3 |
| Median income (in dollars) | SF3A | 1: 28,400 or less<br>2: 28,401–37,850<br>3: 37,851–52,100<br>4: Greater than 52,100 |
| Percent who rent | SF3A | 1: 16 or less<br>2: 17–31<br>3: 32–59<br>4: Greater than 59 |

[1] The SF3A (Summary File 3A) and PL-94 (county-level Public Law 94) variables provide relevant statistics for the block group or the county of the sampled dwelling unit.
[2] Metropolitan Statistical Area.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy; U.S. Department of Commerce, U.S. Census Bureau, Decennial Census, 2000.

#### 11.1.4.1.1 Screener Bivariate Analysis

The distribution of screener respondents was compared with the distribution of all eligible sampled dwelling units for each of the table 11-2 variables. Weighted percentages and standard errors (SEs) were calculated in the WesVar software, using replicated composite screener base weights. To test the significance of the relationship between response status and each of the table 11-2 variables, a Rao-Scott chi-square (RS3) test of independence (Rao and Scott 1984) was performed. In addition, an estimate of bias was calculated for each domain. Bias was estimated as

$$Bias(\bar{y}_R) = (1 - W_R)(\bar{Y}_R - \bar{Y}_N), \tag{3}$$

where $W_R$ is the weighted unit screener response rate (82.2 percent), $\bar{Y}_R$ is the weighted estimate of the domain percentage for respondents, and $\bar{Y}_N$ is the weighted estimate of the domain percentage for nonrespondents. A *t* test[1] was performed to determine whether the bias was significantly different from 0. In accordance with NCES Guideline 5-1-4A, the *t* tests used a simple Bonferroni adjustment to control the overall α-level (0.05) for each domain variable. The Bonferroni adjustment is appropriate for a small number of comparisons (Miller 1981). The adjustment was computed as $\alpha' = \alpha/g$, where *g* is the number of comparisons. For example, for average household size, three *t* tests were conducted. The Bonferroni adjustment was $\alpha'=0.05/3=0.0166$. Therefore, any *p* values of less than α′ in the table were considered statistically significant.

The results of the Rao-Scott chi-square analysis are presented in table 11-3. At the 5 percent α-level, all analysis variables have a significant relationship to screener response status except the percentage of householders who rent their homes. The results of the *t* tests for bias (shown in table 11-4) are consistent with the chi-square analysis. After the Bonferroni adjustment, the bias in estimating the domain percentages is significantly different from 0 for at least one domain of each variable in table 11-2, with the exception of Metropolitan Statistical Area (MSA) status and the percentage of householders who rent their homes. This finding is supported by the evidence of differential response rates among the subgroup domains shown in table 11-1. For instance, dwelling units in segments with a median income of less than $28,400 had a relatively high weighted screener response rate of 88.5 percent. Using only

respondents, without weighting adjustments, would result in an overestimate of this domain percentage by 1.3 (or 7.68 percent), as shown in table 11-4. For this domain, the bias is fairly minor in relation to the

[1]A *t* test is used to compare the means of two domain-level estimates.

**Table 11-3. Household Study sample distribution of screener respondents versus eligible dwelling units, by analysis domain: 2003**

| Analysis domain | Respondents | | | Eligibles | | | Chi-square | |
|---|---|---|---|---|---|---|---|---|
| | Number of respondents | Domain percent | Standard error | Number of eligibles | Domain percent | Standard error | Statistic | *p* value |
| Region | | | | | | | | |
| Northeast | 4,930 | 16.1 | 1.49 | 6,662 | 17.5 | 1.61 | 20.59 | 0.000 |
| Midwest | 4,612 | 24.7 | 1.23 | 5,587 | 24.5 | 1.29 | | |
| South | 11,134 | 37.2 | 1.74 | 13,226 | 36.6 | 1.74 | | |
| West | 4,447 | 22.0 | 0.95 | 5,219 | 21.4 | 0.96 | | |
| MSA[1] status | | | | | | | | |
| Non-MSA | 4,875 | 20.8 | 1.40 | 5,713 | 20.0 | 1.37 | 4.45 | 0.035 |
| MSA | 20,248 | 79.3 | 1.40 | 24,981 | 80.0 | 1.37 | | |
| Average household size | | | | | | | | |
| 2.42 or less | 8,125 | 33.4 | 1.58 | 10,155 | 34.1 | 1.48 | 10.80 | 0.004 |
| 2.43–2.80 | 8,793 | 35.1 | 1.35 | 10,748 | 35.3 | 1.31 | | |
| Greater than 2.80 | 8,205 | 31.5 | 1.87 | 9,791 | 30.7 | 1.79 | | |
| Percent with less than high school education | | | | | | | | |
| 10.4 or less | 6,378 | 31.5 | 1.70 | 8,220 | 33.1 | 1.66 | 43.12 | 0.000 |
| 10.5–20.3 | 6,445 | 27.8 | 1.37 | 7,962 | 27.9 | 1.33 | | |
| 20.4–32.0 | 6,240 | 22.9 | 1.29 | 7,508 | 22.2 | 1.20 | | |
| Greater than 32.0 | 6,060 | 17.8 | 1.32 | 7,004 | 16.8 | 1.23 | | |
| Percent speaking Spanish but not English | | | | | | | | |
| 0 | 10,021 | 42.3 | 2.08 | 12,403 | 42.9 | 2.05 | 13.42 | 0.001 |
| 1–28 | 8,439 | 36.5 | 1.47 | 10,359 | 36.8 | 1.52 | | |
| Greater than 28 | 6,663 | 21.2 | 1.61 | 7,932 | 20.4 | 1.50 | | |
| Percent below 150 percent of poverty | | | | | | | | |
| 10.7 or less | 6,693 | 32.7 | 1.42 | 8,576 | 34.2 | 1.39 | 51.95 | 0.000 |
| 10.8–20.0 | 6,218 | 26.8 | 1.60 | 7,803 | 27.2 | 1.57 | | |
| 20.1–33.3 | 5,952 | 23.2 | 1.62 | 7,117 | 22.4 | 1.51 | | |
| Greater than 33.3 | 6,260 | 17.4 | 1.18 | 7,198 | 16.2 | 1.10 | | |
| Median income (in dollars) | | | | | | | | |
| 28,400 or less | 6,264 | 18.8 | 1.41 | 7,168 | 17.5 | 1.31 | 72.17 | 0.000 |
| 28,401–37,850 | 6,129 | 24.4 | 1.34 | 7,302 | 23.6 | 1.27 | | |
| 37,851–52,100 | 6,195 | 25.6 | 1.47 | 7,797 | 26.2 | 1.39 | | |
| Greater than 52,100 | 6,535 | 31.2 | 1.68 | 8,427 | 32.7 | 1.61 | | |
| Percent who rent | | | | | | | | |
| 16 or less | 6,636 | 31.8 | 1.54 | 8,336 | 32.4 | 1.52 | 6.13 | 0.084 |
| 17–31 | 6,470 | 27.5 | 1.55 | 7,780 | 27.0 | 1.53 | | |
| 32–59 | 6,281 | 22.5 | 1.24 | 7,514 | 22.4 | 1.25 | | |
| Greater than 59 | 5,736 | 18.3 | 0.84 | 7,064 | 18.2 | 0.83 | | |

[1] Metropolitan Statistical Area.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

## Table 11-4. Household Study estimates of screener nonresponse bias, by analysis domain: 2003

| Analysis domain | Eligibles: Domain percent (1) | Eligibles: SE[1] (2) | Respondent domain percent (3) | Non-respondent domain percent (4) | Bias: Estimate (5) | Bias: SE[1] (6) | Bias: *p* value (7) | Relative bias (5)/(1)*100 | Bias ratio (5)/(2) |
|---|---|---|---|---|---|---|---|---|---|
| Region | | | | | | | | | |
| Northeast | 17.5 | 1.61 | 16.1 | 24.0 | –1.4 | 0.24 | 0.000* | –7.93 | –0.86 |
| Midwest | 24.5 | 1.29 | 24.7 | 23.5 | 0.2 | 0.40 | 0.610 | 0.86 | 0.16 |
| South | 36.6 | 1.74 | 37.2 | 34.2 | 0.5 | 0.41 | 0.204 | 1.45 | 0.30 |
| West | 21.4 | 0.96 | 22.0 | 18.3 | 0.7 | 0.27 | 0.016 | 3.09 | 0.69 |
| MSA[2] status | | | | | | | | | |
| Non-MSA | 20.0 | 1.37 | 20.8 | 16.8 | 0.7 | 0.34 | 0.041 | 3.54 | 0.52 |
| MSA | 80.0 | 1.37 | 79.3 | 83.2 | –0.7 | 0.34 | 0.041 | –0.89 | –0.52 |
| Average household size | | | | | | | | | |
| 2.42 or less | 34.1 | 1.48 | 33.4 | 37.4 | –0.7 | 0.29 | 0.019 | –2.08 | –0.48 |
| 2.43–2.80 | 35.3 | 1.31 | 35.1 | 35.9 | –0.2 | 0.27 | 0.588 | –0.43 | –0.11 |
| Greater than 2.80 | 30.7 | 1.79 | 31.5 | 26.7 | 0.9 | 0.28 | 0.003* | 2.81 | 0.48 |
| Percent with less than high school education | | | | | | | | | |
| 10.4 or less | 33.1 | 1.66 | 31.5 | 40.6 | –1.6 | 0.32 | 0.000* | –4.89 | –0.98 |
| 10.5–20.3 | 27.9 | 1.33 | 27.8 | 28.5 | –0.1 | 0.20 | 0.527 | –0.47 | –0.10 |
| 20.4–32 | 22.2 | 1.20 | 22.9 | 19.0 | 0.7 | 0.25 | 0.006* | 3.11 | 0.58 |
| Greater than 32 | 16.8 | 1.23 | 17.8 | 11.9 | 1.1 | 0.20 | 0.000* | 6.27 | 0.85 |
| Percent speaking Spanish but not English | | | | | | | | | |
| 0 | 42.9 | 2.05 | 42.3 | 45.7 | –0.6 | 0.29 | 0.036 | –1.45 | –0.30 |
| 1–28 | 36.8 | 1.52 | 36.5 | 37.8 | –0.2 | 0.29 | 0.425 | –0.63 | –0.15 |
| Greater than 28 | 20.4 | 1.50 | 21.2 | 16.4 | 0.9 | 0.22 | 0.000* | 4.18 | 0.57 |
| Percent below 150 percent of poverty | | | | | | | | | |
| 10.7 or less | 34.2 | 1.39 | 32.7 | 41.1 | –1.5 | 0.31 | 0.000* | –4.35 | –1.07 |
| 10.8–20.0 | 27.2 | 1.57 | 26.8 | 29.3 | –0.4 | 0.25 | 0.079 | –1.65 | –0.29 |
| 20.1–33.3 | 22.4 | 1.51 | 23.2 | 19.0 | 0.7 | 0.24 | 0.003* | 3.30 | 0.49 |
| Greater than 33.3 | 16.2 | 1.10 | 17.4 | 10.6 | 1.2 | 0.16 | 0.000* | 7.43 | 1.09 |
| Median income (in dollars) | | | | | | | | | |
| 28,400 or less | 17.5 | 1.31 | 18.8 | 11.3 | 1.3 | 0.17 | 0.000* | 7.68 | 1.02 |
| 28,401–37,850 | 23.6 | 1.27 | 24.4 | 20.3 | 0.7 | 0.23 | 0.003* | 3.09 | 0.57 |
| 37,851–52,100 | 26.2 | 1.39 | 25.6 | 28.8 | –0.6 | 0.22 | 0.012* | –2.17 | –0.41 |
| Greater than 52,100 | 32.7 | 1.61 | 31.2 | 39.7 | –1.5 | 0.26 | 0.000* | –4.62 | –0.94 |
| Percent who rent | | | | | | | | | |
| 16 or less | 32.4 | 1.52 | 31.8 | 35.2 | –0.6 | 0.26 | 0.023 | –1.85 | –0.39 |
| 17–31 | 27.0 | 1.53 | 27.5 | 25.1 | 0.4 | 0.25 | 0.104 | 1.55 | 0.27 |
| 32–59 | 22.4 | 1.25 | 22.5 | 21.8 | 0.1 | 0.22 | 0.596 | 0.54 | 0.10 |
| Greater than 59 | 18.2 | 0.83 | 18.3 | 17.9 | 0.1 | 0.16 | 0.673 | 0.38 | 0.08 |

* Statistically significant with simple Bonferroni adjustment at $\alpha = 0.05$.
[1] Standard error.
[2] Metropolitan Statistical Area.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

standard error; the ratio of bias to the standard error is 1.02. A bias ratio over 1.96 would provide a strong indication of potential bias, since the estimate based on respondents would differ from the estimate based on eligibles by more than 1.96 times the standard error of the estimate.

Although the relationships between response status and the table 11-2 variables are significant, the differences between the distributions of respondents and eligible dwelling units are minor. The absolute bias is less than 2 for all estimated domain percentages. In addition, many of the table 11-2 variables were used in weighting adjustments, and so differences between respondents and eligible dwelling units were reduced through the weighting process (refer to section 11.1.4.1.3). Therefore, the bivariate analysis indicates minimal potential for bias at the screener level, and thus minimal impact of screener nonresponse on literacy scores, assuming that literacy scores are highly correlated with the variables used in the weighting adjustments.

#### 11.1.4.1.2 Screener Multivariate Analysis

The bivariate analysis described in section 11.1.4.1.1 is useful in explaining each variable individually. A multivariate analysis is useful in showing relationships among a number of variables. One approach is to provide a CHAID analysis. CHAID is a classification algorithm that uses chi-square tests to divide a sample into subgroups that best explain differential response rates.

The analysis in CHAID begins by dividing the sample into two or more groups on the basis of the categories of the best predictor. Each of these groups is divided into smaller subgroups on the basis of the best available predictor at each level. The splitting process continues until either no significant predictor remains or the minimum cell size requirement is met. The CHAID software displays the final subgroups in the form of a tree diagram whose branches (nodes) correspond to the groups. The resulting classification tree reveals the domains, as defined by combinations of variables, with the most differential response rates, thereby leading to domains with the highest potential for nonresponse bias.

CHAID was run with screener response status as the dependent variable and the table 11-2 variables as the independent variables. Cell sizes were limited to 300 or more dwelling units (approximately 1 percent of the sample), and up to three-way interactions were allowed (three tree levels). The resulting tree is shown in figure 11-1 and summarized in table 11-5. Twenty-five cells were formed,

**Figure 11-1. Household Study multivariate CHAID analysis of screener response indicators: 2003**

Overall weighted response rate = 82.2 percent
Total number of eligibles = 30,694

| Median income (in dollars) | | Region | | Split variable | Category | | Cell |
|---|---|---|---|---|---|---|---|
| 28,400 or less | 88.5% (7,168) | Northeast | 81.1% (1,309) | | | | Cell 1 |
| | | Midwest, South, West | 89.8% (5,859) | Household size | 2.42 or less | 87.5% (2,336) | Cell 2 |
| | | | | | 2.43–2.80 | 90.5% (1,938) | Cell 3 |
| | | | | | Greater than 2.80 | 93.1% (1,585) | Cell 4 |
| 28,401–37,850 | 84.7% (7,302) | Northeast | 77.0% (1,251) | | | | Cell 5 |
| | | Midwest, South, West | 86.1% (6,051) | Household size | 2.42 or less | 84.6% (2,355) | Cell 6 |
| | | | | | Greater than 2.42 | 87.2% (3,696) | Cell 7 |
| 37,851–52,100 | 80.4% (7,797) | Northeast | 73.2% (1,652) | Percent who rent | 59% or less | 74.8% (1,235) | Cell 8 |
| | | | | | Greater than 59% | 64.4% (417) | Cell 9 |
| | | Midwest | 79.6% (1,677) | | | | Cell 10 |
| | | South | 82.1% (3,251) | Percent below poverty | 10.7% or less | 74.3% (491) | Cell 11 |
| | | | | | 10.8–20.0% | 81.7% (2,156) | Cell 12 |
| | | | | | Greater than 20.0% | 91.3% (604) | Cell 13 |
| | | West | 84.6% (1,217) | Percent who speak Spanish, not English | 28% or less | 81.6% (674) | Cell 14 |
| | | | | | Greater than 28% | 91.7% (543) | Cell 15 |

See notes at end of figure.

**Figure 11-1. Household Study multivariate CHAID analysis of screener response indicators: 2003—Continued**

| Median income (in dollars) | | Household size | | Percent who rent / Region | | |
|---|---|---|---|---|---|---|
| Greater than 52,100 | 78.4% (8,427) | 2.42 or less | 72.5% (1,698) | **Percent who rent** | | |
| | | | | 16% or less | 74.4% (378) | Cell 16 |
| | | | | 17–31% | 83.3% (423) | Cell 17 |
| | | | | 32–59% | 68.0% (581) | Cell 18 |
| | | | | Greater than 59% | 59.7% (316) | Cell 19 |
| | | 2.43–2.80 | 77.5% (2,694) | **Region** | | |
| | | | | Northeast, West | 77.4% (1,197) | Cell 20 |
| | | | | Midwest | 81.7% (622) | Cell 21 |
| | | | | South | 72.1% (875) | Cell 22 |
| | | Greater than 2.80 | 81.5% (4,035) | **Percent who rent** | | |
| | | | | 16% or less | 81.7% (3,069) | Cell 23 |
| | | | | 17–31% | 78.1% (601) | Cell 24 |
| | | | | Greater than 31% | 88.9% (365) | Cell 25 |

NOTE: CHAID software uses a classification algorithm to divide the sample into subgroups that best explain differential response rates. All percentages are weighted response rates and the numbers inside the parentheses are sample sizes.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 11-5. Household Study multivariate CHAID analysis of screener response indicators, by response cell: 2003**

| Response cell | Number of eligibles | Number of respondents | Unweighted response rate (percent) | Weighted response rate (percent) |
|---|---|---|---|---|
| Overall | 30,694 | 25,123 | 81.9 | 82.2 |
| | | | | |
| 1 | 1,309 | 1,036 | 79.1 | 81.1 |
| 2 | 2,336 | 2,041 | 87.4 | 87.5 |
| 3 | 1,938 | 1,729 | 89.2 | 90.5 |
| 4 | 1,585 | 1,458 | 92.0 | 93.1 |
| 5 | 1,251 | 945 | 75.5 | 77.0 |
| 6 | 2,355 | 1,980 | 84.1 | 84.6 |
| 7 | 3,696 | 3,204 | 86.7 | 87.2 |
| 8 | 1,235 | 915 | 74.1 | 74.8 |
| 9 | 417 | 264 | 63.3 | 64.4 |
| 10 | 1,677 | 1,317 | 78.5 | 79.6 |
| 11 | 491 | 374 | 76.2 | 74.3 |
| 12 | 2,156 | 1,748 | 81.1 | 81.7 |
| 13 | 604 | 534 | 88.4 | 91.3 |
| 14 | 674 | 551 | 81.8 | 81.6 |
| 15 | 543 | 492 | 90.6 | 91.7 |
| 16 | 378 | 276 | 73.0 | 74.4 |
| 17 | 423 | 341 | 80.6 | 83.3 |
| 18 | 581 | 403 | 69.3 | 68.0 |
| 19 | 316 | 187 | 59.2 | 59.7 |
| 20 | 1,197 | 907 | 75.8 | 77.4 |
| 21 | 622 | 493 | 79.3 | 81.7 |
| 22 | 875 | 674 | 77.0 | 72.1 |
| 23 | 3,069 | 2,474 | 80.6 | 81.7 |
| 24 | 601 | 467 | 77.7 | 78.1 |
| 25 | 365 | 313 | 85.8 | 88.9 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

with weighted response rates ranging from 59.7 percent to 93.1 percent. The lowest response rate was for the group within segments with high median income (greater than $52,100), small average household size (2.42 or less), and a large proportion of renters (greater than 59 percent). The highest response rate was for the group within segments with low median income ($28,400 or less) and large average household size (greater than 2.8), in the Midwest, South, or West. Median income was the dominant variable in distinguishing response rate groups, which is consistent with the results of the bivariate analysis. Region, household size, percentage of householders who rent, percentage below 150 percent of poverty, and percentage speaking Spanish but not English were also significant contributors to the CHAID tree.

Although the CHAID tree is useful for dissecting the sample into fine groups of dwelling units with response patterns as different as possible, it should be used with caution because CHAID does not take into account the complex design of the sample. Consequently, the significance level of the test may be lower than the 0.05 α-level indicated. If the appropriate significance level could be used, then the tree might have fewer significant response cells. Thus, the tree shown in figure 11-1 is a conservative picture because any indication of nonresponse bias shown by the CHAID results may be overstated.

Logistic regression models are also useful in identifying significant effects on response propensity. Screener response status was used as the binary dependent variable, and the table 11-2 variables were used as the predictors. The main effects model had the form

$$\log\left[\frac{\Pr(\text{Response})}{1-\Pr(\text{Response})}\right] = A + \sum B_i X_{ij} ,$$

where the $X_{ij}$'s are indicator variables for the table 11-2 variables. An $F$ test was performed on each table 11-2 variable to determine whether it was significantly related to response propensity.

The results of the logistic regression analysis are presented in table 11-6. Three variables—region, average household size, and median income—were significantly related to response propensity at the 5 percent α-level. Response propensity is significantly lower in the Northeast than in the West, in segments with small average household sizes compared to those with large average household sizes, and in high median income segments compared to lower median income segments. The results are consistent with the CHAID analysis: The same three variables were selected into the first two levels of the CHAID tree. All three variables were used in adjusting the screener weights for nonresponse for at least one SAAL state or for the NAAL sample. Thus, the potential for nonresponse bias suggested by the multivariate analysis was reduced through the weighting adjustments, as shown in section 11.1.4.1.3.

**Table 11-6. Household Study multivariate logistic regression analysis of screener response indicators, by predictor: 2003**

| Predictor | *F* test: *F* statistic | *F* test: Numerator *df*[1] | *F* test: Denominator *df*[1] | *F* test: *p* value | Regression coefficient: Estimate | Regression coefficient: *p* value |
|---|---|---|---|---|---|---|
| Overall fit | 9.98 | 20 | 42 | 0.000 | † | † |
| Region | 14.71 | 3 | 59 | 0.000 | † | † |
| Northeast | † | † | † | † | –0.52 | 0.000 |
| Midwest | † | † | † | † | –0.07 | 0.602 |
| South | † | † | † | † | –0.15 | 0.216 |
| MSA[2] status | 0.11 | 1 | 61 | 0.745 | † | † |
| Average household size | 12.27 | 2 | 60 | 0.000 | † | † |
| 2.42 or less | † | † | † | † | –0.41 | 0.000 |
| 2.43–2.80 | † | † | † | † | –0.28 | 0.001 |
| Percent with less than high school | 0.45 | 3 | 59 | 0.718 | † | † |
| Percent speaking Spanish but not | 0.06 | 2 | 60 | 0.941 | † | † |
| Percent below 150 percent of | 0.10 | 3 | 59 | 0.961 | † | † |
| Median income (in dollars) | 5.88 | 3 | 59 | 0.001 | † | † |
| 28,400 or less | † | † | † | † | 0.74 | 0.000 |
| 28,401–37,850 | † | † | † | † | 0.43 | 0.012 |
| 37,851–52,100 | † | † | † | † | 0.12 | 0.299 |
| Percent who rent | 2.01 | 3 | 59 | 0.122 | † | † |

† Not applicable.
[1] Degrees of freedom.
[2] Metropolitan Statistical Area.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

#### 11.1.4.1.3 Potential for Screener Nonresponse Bias Remaining After Weighting Procedures

As described in section 12.1.3, weighting procedures were implemented to reduce the potential for nonresponse bias by creating nonresponse adjustment classes for which the respondents' literacy-related characteristics are similar to those of nonrespondents. The extent of the reduction in nonresponse bias depends on the correlation of the weighting class variables with literacy scores.

Tables 11-7 through 11-13 show *t* test results for the change in the distribution of sample cases after each screener weighting stage for the national sample and each of the six participating states. The *t* tests were performed to determine whether the change in the estimated domain percentage is significantly different from 0. A simple Bonferroni adjustment was used to control the overall α-level at 0.05 for each domain. The adjustment is computed as $\alpha' = \alpha/g$, where $g$ is the number of comparisons. Therefore, any $p$ values less than $\alpha'$ in the tables can be considered statistically significant. Calculations of the bias ratio of estimated percentages are also included. Although the *t* test may indicate a statistically significant difference, the difference may not be important. Thus, it is also useful to look at the bias ratio, the ratio of the bias to the standard error of the estimate, to gauge the importance of the potential bias. That is, if the absolute value of the bias to standard error ratio is greater than 1.96, then the ratio provides a strong indication of potential bias.

The checks were performed separately for the national NAAL household sample and each of the six SAAL states to reflect the weighting process. Unlike the results in table 11-3, these comparisons use the actual survey weights, which were processed separately for each household sample. The following comparisons were made for each of the analysis domains in table 11-2:

- Comparison of distributions from screener base weights for the estimated eligible population with those for the screener respondents only, to check for differences owing to screener nonresponse, and

- Comparison of distributions from screener base weights for the estimated eligible population with those from the screener nonresponse-adjusted weights, to check for differences even after the nonresponse adjustment to the screener.

The $p$ values resulting from the first set of comparisons indicate a significant difference between the eligible dwelling units and respondents for most of the subgroups when base weights are used. This result is comparable to those obtained for the bivariate and multivariate analyses described in sections 11.1.4.1.1 and 11.1.4.1.2, respectively. A nonresponse adjustment was necessary to reduce the bias in estimates based on data from respondents only.

The *p* values resulting from the second set of comparisons show that for most of the subgroups, there is no significant difference between the weighted distribution of eligible dwelling units and the respondents after the nonresponse adjustment. For the national NAAL sample, the one variable with a significant difference is average household size, but the bias is minor, at less than half the standard error of the estimated percentage. Therefore, the nonresponse adjustment appears to have been effective in reducing the bias owing to screener nonresponse, to the extent that table 11-2 variables are related to literacy. For the state samples, fewer variables were available for use in nonresponse adjustments because of the smaller sample sizes; therefore, the bias estimates are generally higher. However, the bias ratio remains less than 1.00 for all estimates in the tables.

**Table 11-7. Household Study screener weighting effects for the national NAAL household sample, by subgroup: 2003**

| | Base weight —eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio | | Bias | | | Bias ratio |
| Subgroup | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | *p* value | (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | *p* value | (4)-(1) / (2) |
| Total | 100.0 | † | 100.0 | † | † | † | † | 100.0 | † | † | † | † |
| Region | | | | | | | | | | | | |
| Northeast | 19.3 | 0.51 | 18.0 | –1.3 | 0.24 | 0.000* | –2.48 | 19.3 | # | 0.04 | 0.371 | –0.06 |
| Midwest | 23.3 | 0.81 | 23.3 | 0.1 | 0.28 | 0.827 | 0.09 | 23.2 | # | 0.09 | 0.607 | –0.05 |
| South | 36.3 | 0.85 | 37.0 | 0.8 | 0.28 | 0.006* | 0.92 | 36.4 | 0.1 | 0.07 | 0.148 | 0.11 |
| West | 21.2 | 0.63 | 21.6 | 0.4 | 0.23 | 0.070 | 0.65 | 21.2 | # | 0.07 | 0.806 | –0.03 |
| MSA[2] status | | | | | | | | | | | | |
| Non-MSA | 20.6 | 0.59 | 20.8 | 0.3 | 0.23 | 0.226 | 0.48 | 20.5 | # | 0.15 | 0.931 | –0.03 |
| MSA | 79.4 | 0.59 | 79.2 | –0.3 | 0.23 | 0.226 | –0.48 | 79.5 | # | 0.15 | 0.931 | 0.03 |
| Average household size | | | | | | | | | | | | |
| 2.42 or less | 33.9 | 1.63 | 33.1 | –0.8 | 0.27 | 0.006* | –0.48 | 33.3 | –0.6 | 0.21 | 0.005* | –0.38 |
| 2.43–2.80 | 35.6 | 1.47 | 35.5 | –0.1 | 0.23 | 0.637 | –0.08 | 35.7 | 0.1 | 0.20 | 0.661 | 0.06 |
| Greater than 2.80 | 30.5 | 1.95 | 31.4 | 0.9 | 0.24 | 0.001* | 0.45 | 31.0 | 0.5 | 0.16 | 0.002* | 0.27 |
| Percent with less than high school education | | | | | | | | | | | | |
| 10.4 or less | 33.0 | 1.64 | 31.4 | –1.6 | 0.29 | 0.000* | –0.98 | 32.7 | –0.3 | 0.21 | 0.165 | –0.18 |
| 10.5–20.3 | 28.2 | 1.37 | 28.1 | –0.1 | 0.21 | 0.511 | –0.09 | 28.2 | # | 0.19 | 0.898 | 0.02 |
| 20.4–32.0 | 22.0 | 1.12 | 22.6 | 0.7 | 0.19 | 0.001* | 0.59 | 22.1 | 0.2 | 0.14 | 0.165 | 0.17 |
| Greater than 32.0 | 16.8 | 1.31 | 17.9 | 1.1 | 0.16 | 0.000* | 0.83 | 16.9 | 0.1 | 0.08 | 0.353 | 0.05 |
| Percent speaking Spanish but not English | | | | | | | | | | | | |
| 0 | 43.4 | 2.10 | 42.3 | –1.1 | 0.24 | 0.000* | –0.53 | 43.2 | –0.2 | 0.21 | 0.313 | –0.10 |
| 1–28 | 36.6 | 1.56 | 36.4 | –0.2 | 0.24 | 0.371 | –0.13 | 36.7 | # | 0.22 | 0.953 | 0.01 |
| Greater than 28 | 19.9 | 1.56 | 21.2 | 1.3 | 0.20 | 0.000* | 0.84 | 20.1 | 0.2 | 0.11 | 0.081 | 0.12 |

See notes at end of table.

**Table 11-7. Household Study screener weighting effects for the national NAAL household sample, by subgroup: 2003—Continued**

| Subgroup | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio | | Bias | | | Bias ratio |
| | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | p value | (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | p value | (4)-(1) / (2) |
| Percent below 150 percent of poverty | | | | | | | | | | | | |
| 10.7 or less | 34.4 | 1.41 | 32.9 | –1.5 | 0.25 | 0.000* | –1.06 | 34.5 | # | 0.03 | 0.224 | 0.03 |
| 10.8–20.0 | 27.2 | 1.56 | 26.9 | –0.3 | 0.21 | 0.218 | –0.17 | 27.1 | # | 0.03 | 0.224 | –0.03 |
| 20.1–33.3 | 22.0 | 1.55 | 22.6 | 0.6 | 0.18 | 0.002* | 0.37 | 22.0 | # | 0.00 | 0.917 | 0.00 |
| Greater than 33.3 | 16.4 | 1.14 | 17.6 | 1.2 | 0.14 | 0.000* | 1.04 | 16.4 | # | 0.00 | 0.854 | 0.00 |
| Median income (in dollars) | | | | | | | | | | | | |
| 28,400 or less | 18.0 | 1.35 | 19.1 | 1.1 | 0.15 | 0.000* | 0.82 | 18.1 | # | 0.07 | 0.595 | 0.03 |
| 28,401–37,850 | 23.3 | 1.31 | 24.0 | 0.7 | 0.22 | 0.005* | 0.50 | 23.5 | 0.2 | 0.15 | 0.299 | 0.12 |
| 37,851–52,100 | 25.5 | 1.48 | 25.3 | –0.2 | 0.21 | 0.318 | –0.14 | 25.4 | –0.1 | 0.17 | 0.663 | –0.05 |
| Greater than 52,100 | 33.2 | 1.59 | 31.6 | –1.5 | 0.24 | 0.000* | –0.96 | 33.0 | –0.1 | 0.18 | 0.508 | –0.07 |
| Percent who rent | | | | | | | | | | | | |
| 16 or less | 33.1 | 1.69 | 32.4 | –0.7 | 0.23 | 0.003* | –0.42 | 33.2 | 0.1 | 0.21 | 0.785 | 0.04 |
| 17–31 | 27.3 | 1.58 | 27.4 | 0.1 | 0.21 | 0.696 | 0.05 | 27.4 | 0.1 | 0.18 | 0.775 | 0.03 |
| 32–59 | 22.1 | 1.40 | 22.3 | 0.2 | 0.20 | 0.276 | 0.16 | 22.1 | # | 0.18 | 0.955 | 0.01 |
| Greater than 59 | 17.5 | 0.84 | 18.0 | 0.4 | 0.14 | 0.005* | 0.49 | 17.4 | –0.1 | 0.11 | 0.295 | –0.14 |

† Not applicable.
# Rounds to zero.
* Statistically significant with Bonferroni adjustment at $\alpha = 0.05$.
[1] Standard error.
[2] Metropolitan Statistical Area.
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 11-8. Household Study screener weighting effects for Kentucky, by subgroup: 2003**

| Subgroup | Base weight—eligible Percent (1) | Base weight—eligible SE[1] (2) | Base weight—respondents Percent (3) | Bias Estimate (3)-(1) | Bias SE[1] | Bias p value | Bias ratio (3)-(1)/(2) | Nonresponse-adjusted weight—respondents Percent (4) | Bias Estimate (4)-(1) | Bias SE[1] | Bias p value | Bias ratio (4)-(1)/(2) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Total | 100.0 | † | 100.0 | † | † | † | † | 100.0 | † | † | † | † |
| MSA[2] status | | | | | | | | | | | | |
| Non-MSA | 52.0 | 1.34 | 53.3 | 1.3 | 0.60 | 0.038 | 0.99 | 52.4 | 0.3 | 0.41 | 0.429 | 0.25 |
| MSA | 48.0 | 1.34 | 46.7 | –1.3 | 0.60 | 0.038 | –0.99 | 47.6 | –0.3 | 0.41 | 0.429 | –0.25 |
| Average household size | | | | | | | | | | | | |
| 2.42 or less | 34.4 | 3.68 | 34.2 | –0.1 | 0.43 | 0.771 | –0.04 | 33.8 | –0.6 | 0.38 | 0.150 | –0.16 |
| 2.43–2.80 | 53.4 | 3.82 | 53.3 | –0.1 | 0.49 | 0.841 | –0.03 | 53.4 | 0.1 | 0.50 | 0.862 | 0.02 |
| Greater than 2.80 | 12.3 | 3.29 | 12.5 | 0.2 | 0.36 | 0.538 | 0.07 | 12.8 | 0.5 | 0.39 | 0.224 | 0.15 |
| Percent with less than high school education | | | | | | | | | | | | |
| 10.4 or less | 17.3 | 3.23 | 16.4 | –0.9 | 0.47 | 0.059 | –0.29 | 17.2 | –0.1 | 0.14 | 0.464 | –0.03 |
| 10.5–20.3 | 20.3 | 3.48 | 19.5 | –0.8 | 0.42 | 0.077 | –0.23 | 20.4 | 0.1 | 0.18 | 0.623 | 0.03 |
| 20.4–32.0 | 36.6 | 4.53 | 37.2 | 0.6 | 0.39 | 0.134 | 0.14 | 36.6 | # | 0.12 | 0.737 | 0.01 |
| Greater than 32.0 | 25.8 | 3.00 | 26.9 | 1.1 | 0.38 | 0.008* | 0.37 | 25.8 | # | 0.10 | 0.843 | –0.01 |
| Percent speaking Spanish but not English | | | | | | | | | | | | |
| 0 | 61.9 | 3.95 | 62.4 | 0.5 | 0.61 | 0.448 | 0.12 | 61.9 | # | 0.00 | 0.931 | 0.00 |
| 1–28 | 32.0 | 3.00 | 31.9 | –0.1 | 0.50 | 0.816 | –0.04 | 32.2 | 0.1 | 0.19 | 0.476 | 0.05 |
| Greater than 28 | 6.1 | 1.77 | 5.7 | –0.4 | 0.23 | 0.137 | –0.20 | 5.9 | –0.1 | 0.19 | 0.476 | –0.08 |
| Percent below 150 percent of poverty | | | | | | | | | | | | |
| 10.7 or less | 15.7 | 2.78 | 14.2 | –1.5 | 0.39 | 0.001* | –0.54 | 15.2 | –0.5 | 0.19 | 0.021 | –0.17 |
| 10.8–20.0 | 27.3 | 2.80 | 26.8 | –0.5 | 0.50 | 0.311 | –0.19 | 27.9 | 0.6 | 0.36 | 0.118 | 0.21 |
| 20.1–33.3 | 26.0 | 4.38 | 26.6 | 0.7 | 0.33 | 0.053 | 0.16 | 26.0 | # | 0.35 | 0.982 | 0.00 |
| Greater than 33.3 | 31.0 | 4.33 | 32.3 | 1.3 | 0.39 | 0.003* | 0.31 | 30.9 | –0.1 | 0.20 | 0.592 | –0.03 |

See notes at end of table.

**Table 11-8. Household Study screener weighting effects for Kentucky, by subgroup: 2003—Continued**

| Subgroup | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio | | Bias | | | Bias ratio |
| | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | $p$ value | (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | $p$ value | (4)-(1) / (2) |
| Median income (in dollars) | | | | | | | | | | | | |
| 28,400 or less | 30.2 | 4.13 | 31.8 | 1.6 | 0.35 | 0.000* | 0.38 | 30.2 | # | 0.00 | 0.946 | 0.00 |
| 28,401–37,850 | 27.6 | 4.23 | 28.7 | 1.0 | 0.37 | 0.013 | 0.24 | 27.6 | # | 0.00 | 0.883 | 0.00 |
| 37,851–52,100 | 28.3 | 2.53 | 26.7 | –1.7 | 0.57 | 0.009* | –0.66 | 28.3 | # | 0.00 | 0.824 | 0.00 |
| Greater than 52,100 | 13.8 | 2.44 | 12.9 | –0.9 | 0.52 | 0.089 | –0.38 | 13.8 | # | 0.00 | 0.827 | 0.00 |
| Percent who rent | | | | | | | | | | | | |
| 16 or less | 28.8 | 2.93 | 27.9 | –0.9 | 0.47 | 0.081 | –0.29 | 28.6 | –0.2 | 0.37 | 0.605 | –0.06 |
| 17–31 | 39.4 | 3.46 | 39.3 | –0.1 | 0.47 | 0.867 | –0.02 | 39.4 | # | 0.36 | 0.964 | 0.00 |
| 32–59 | 22.0 | 2.40 | 22.7 | 0.7 | 0.34 | 0.068 | 0.28 | 22.2 | 0.2 | 0.26 | 0.534 | 0.07 |
| Greater than 59 | 9.8 | 1.72 | 10.1 | 0.3 | 0.24 | 0.251 | 0.17 | 9.8 | # | 0.16 | 0.935 | 0.01 |

†Not applicable.
# Rounds to zero.
*Statistically significant with Bonferroni adjustment at $\alpha = 0.05$.
[1]Standard error.
[2]Metropolitan Statistical Area.
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 11-9. Household Study screener weighting effects for Maryland, by subgroup: 2003**

| | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | | | Bias | | | |
| Subgroup | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | *p* value | Bias ratio (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | *p* value | Bias ratio (4)-(1) / (2) |
| Total | 100.0 | † | 100.0 | † | † | † | † | 100.0 | † | † | † | † |
| MSA[2] status | | | | | | | | | | | | |
| Non-MSA | 6.8 | 1.30 | 6.8 | # | 0.21 | 0.997 | 0.00 | 6.8 | # | 0.22 | 0.977 | –0.01 |
| MSA | 93.2 | 1.30 | 93.2 | # | 0.21 | 0.997 | 0.00 | 93.2 | # | 0.22 | 0.977 | 0.01 |
| Average household size | | | | | | | | | | | | |
| 2.42 or less | 38.7 | 4.47 | 38.6 | –0.2 | 0.51 | 0.772 | –0.04 | 38.4 | –0.4 | 0.49 | 0.455 | –0.08 |
| 2.43–2.80 | 30.0 | 3.37 | 30.4 | 0.4 | 0.57 | 0.499 | 0.12 | 30.6 | 0.6 | 0.42 | 0.141 | 0.19 |
| Greater than 2.80 | 31.3 | 4.61 | 31.1 | –0.2 | 0.49 | 0.624 | –0.05 | 31.0 | –0.3 | 0.58 | 0.645 | –0.06 |
| Percent with less than high school education | | | | | | | | | | | | |
| 10.4 or less | 44.9 | 3.35 | 45.2 | 0.3 | 0.92 | 0.740 | 0.09 | 44.8 | –0.1 | 0.46 | 0.814 | –0.03 |
| 10.5–20.3 | 28.0 | 4.35 | 27.5 | –0.6 | 0.78 | 0.471 | –0.13 | 27.8 | –0.2 | 0.59 | 0.707 | –0.05 |
| 20.4–32.0 | 19.3 | 3.93 | 19.1 | –0.3 | 0.71 | 0.718 | –0.07 | 19.2 | –0.1 | 0.57 | 0.874 | –0.02 |
| Greater than 32.0 | 7.7 | 1.47 | 8.3 | 0.5 | 0.42 | 0.233 | 0.36 | 8.2 | 0.4 | 0.25 | 0.106 | 0.29 |
| Percent speaking Spanish but not English | | | | | | | | | | | | |
| 0 | 56.4 | 5.48 | 56.6 | 0.2 | 0.68 | 0.765 | 0.04 | 56.4 | –0.1 | 0.64 | 0.930 | –0.01 |
| 1–28 | 35.5 | 5.32 | 35.2 | –0.4 | 0.80 | 0.669 | –0.07 | 35.6 | 0.1 | 0.72 | 0.903 | 0.02 |
| Greater than 28 | 8.1 | 1.49 | 8.2 | 0.1 | 0.48 | 0.776 | 0.09 | 8.0 | # | 0.42 | 0.941 | –0.02 |
| Percent below 150 percent of poverty | | | | | | | | | | | | |
| 10.7 or less | 59.0 | 3.68 | 59.1 | 0.1 | 0.77 | 0.951 | 0.01 | 58.6 | –0.5 | 0.56 | 0.432 | –0.12 |
| 10.8–20.0 | 25.2 | 4.10 | 24.8 | –0.4 | 0.80 | 0.600 | –0.10 | 25.2 | # | 0.68 | 0.966 | –0.01 |
| 20.1–33.3 | 10.3 | 1.81 | 10.3 | –0.1 | 0.64 | 0.945 | –0.03 | 10.4 | 0.1 | 0.58 | 0.861 | 0.06 |
| Greater than 33.3 | 5.4 | 1.31 | 5.8 | 0.4 | 0.18 | 0.033 | 0.33 | 5.8 | 0.4 | 0.37 | 0.324 | 0.29 |

See notes at end of table.

**Table 11-9. Household Study screener weighting effects for Maryland, by subgroup: 2003—Continued**

| Subgroup | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | | | Bias | | | |
| | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | p value | Bias ratio (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | p value | Bias ratio (4)-(1) / (2) |
| Median income (in dollars) | | | | | | | | | | | | |
| 28,400 or less | 7.9 | 1.79 | 8.3 | 0.4 | 0.37 | 0.246 | 0.25 | 8.5 | 0.6 | 0.32 | 0.067 | 0.35 |
| 28,401–37,850 | 13.4 | 3.76 | 13.1 | –0.4 | 0.66 | 0.582 | –0.10 | 13.2 | –0.3 | 0.36 | 0.460 | –0.07 |
| 37,851–52,100 | 22.5 | 3.06 | 22.0 | –0.5 | 0.91 | 0.611 | –0.15 | 22.3 | –0.2 | 0.70 | 0.765 | –0.07 |
| Greater than 52,100 | 56.2 | 2.80 | 56.6 | 0.4 | 0.79 | 0.622 | 0.14 | 56.1 | –0.1 | 0.48 | 0.782 | –0.05 |
| Percent who rent | | | | | | | | | | | | |
| 16 or less | 35.2 | 4.46 | 35.3 | 0.1 | 0.49 | 0.861 | 0.02 | 35.2 | # | 0.59 | 0.980 | 0.00 |
| 17–31 | 25.7 | 3.27 | 25.0 | –0.7 | 0.63 | 0.260 | –0.22 | 25.3 | –0.5 | 0.81 | 0.588 | –0.14 |
| 32–59 | 24.0 | 4.70 | 24.1 | 0.1 | 0.62 | 0.865 | 0.02 | 24.1 | 0.1 | 0.45 | 0.875 | 0.01 |
| Greater than 59 | 15.1 | 4.12 | 15.6 | 0.5 | 0.48 | 0.280 | 0.13 | 15.5 | 0.4 | 0.70 | 0.611 | 0.09 |

† Not applicable.
# Rounds to zero.
[1] Standard error.
[2] Metropolitan Statistical Area.
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

## Table 11-10. Household Study screener weighting effects for Massachusetts, by subgroup: 2003

| Subgroup | Base weight—eligible: Percent (1) | Base weight—eligible: SE[1] (2) | Base weight—respondents: Percent (3) | Base weight—respondents: Bias: Estimate (3)-(1) | Base weight—respondents: Bias: SE[1] | Base weight—respondents: Bias: $p$ value | Base weight—respondents: Bias ratio (3)-(1) / (2) | Nonresponse-adjusted weight—respondents: Percent (4) | Nonresponse-adjusted weight—respondents: Bias: Estimate (4)-(1) | Nonresponse-adjusted weight—respondents: Bias: SE[1] | Nonresponse-adjusted weight—respondents: Bias: $p$ value | Nonresponse-adjusted weight—respondents: Bias ratio (4)-(1) / (2) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Total | 100.0 | † | 100.0 | † | † | † | † | 100.0 | † | † | † | † |
| MSA[2] status | | | | | | | | | | | | |
| Non-MSA | 1.6 | 0.45 | 1.5 | –0.1 | 0.20 | 0.602 | –0.24 | 1.5 | # | 0.23 | 0.870 | –0.09 |
| MSA | 98.4 | 0.45 | 98.6 | 0.1 | 0.20 | 0.602 | 0.24 | 98.5 | # | 0.23 | 0.870 | 0.09 |
| Average household size | | | | | | | | | | | | |
| 2.42 or less | 37.7 | 4.66 | 37.3 | –0.5 | 0.61 | 0.470 | –0.10 | 37.6 | –0.1 | 0.43 | 0.787 | –0.03 |
| 2.43–2.80 | 33.0 | 2.93 | 32.9 | –0.1 | 0.73 | 0.879 | –0.04 | 33.1 | 0.1 | 0.47 | 0.836 | 0.03 |
| Greater than 2.80 | 29.3 | 4.29 | 29.9 | 0.6 | 0.58 | 0.346 | 0.13 | 29.3 | # | 0.41 | 0.960 | 0.00 |
| Percent with less than high school education | | | | | | | | | | | | |
| 10.4 or less | 45.1 | 6.09 | 44.6 | –0.5 | 1.04 | 0.638 | –0.08 | 45.1 | # | 0.51 | 0.998 | 0.00 |
| 10.5–20.3 | 34.2 | 5.45 | 33.8 | –0.4 | 0.98 | 0.709 | –0.07 | 34.0 | –0.2 | 0.69 | 0.764 | –0.04 |
| 20.4–32.0 | 9.4 | 2.32 | 9.5 | 0.1 | 0.51 | 0.875 | 0.03 | 9.6 | 0.2 | 0.45 | 0.691 | 0.08 |
| Greater than 32.0 | 11.4 | 1.98 | 12.1 | 0.8 | 0.25 | 0.005* | 0.40 | 11.4 | # | 0.13 | 0.834 | 0.02 |
| Percent speaking Spanish but not English | | | | | | | | | | | | |
| 0 | 61.3 | 2.78 | 59.7 | –1.6 | 0.91 | 0.093 | –0.57 | 60.6 | –0.7 | 0.90 | 0.435 | –0.26 |
| 1–28 | 31.0 | 2.73 | 31.4 | 0.4 | 0.99 | 0.704 | 0.14 | 31.5 | 0.5 | 0.93 | 0.610 | 0.18 |
| Greater than 28 | 7.7 | 1.70 | 9.0 | 1.2 | 0.34 | 0.002* | 0.72 | 8.0 | 0.2 | 0.16 | 0.160 | 0.14 |
| Percent below 150 percent of poverty | | | | | | | | | | | | |
| 10.7 or less | 54.4 | 5.20 | 53.4 | –1.1 | 1.08 | 0.344 | –0.20 | 54.1 | –0.4 | 0.65 | 0.589 | –0.07 |
| 10.8–20.0 | 19.3 | 5.29 | 19.3 | # | 0.60 | 0.993 | 0.00 | 19.4 | 0.1 | 0.67 | 0.859 | 0.02 |
| 20.1–33.3 | 16.8 | 3.34 | 17.4 | 0.6 | 0.79 | 0.446 | 0.18 | 17.3 | 0.5 | 0.34 | 0.173 | 0.14 |
| Greater than 33.3 | 9.51 | 1.81 | 10.0 | 0.4 | 0.53 | 0.413 | 0.24 | 9.3 | –0.2 | 0.24 | 0.340 | –0.13 |

See notes at end of table.

**Table 11-10. Household Study screener weighting effects for Massachusetts, by subgroup: 2003—Continued**

| | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio | | Bias | | | Bias ratio |
| Subgroup | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | *p* value | (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | *p* value | (4)-(1) / (2) |
| Median income (in dollars) | | | | | | | | | | | | |
| 28,400 or less | 9.2 | 2.47 | 9.6 | 0.4 | 0.62 | 0.470 | 0.18 | 9.4 | 0.2 | 0.34 | 0.491 | 0.09 |
| 28,401–37,850 | 10.0 | 1.50 | 10.7 | 0.7 | 0.23 | 0.006* | 0.49 | 10.4 | 0.4 | 0.36 | 0.259 | 0.28 |
| 37,851–52,100 | 32.4 | 4.16 | 31.7 | –0.7 | 0.97 | 0.451 | –0.18 | 31.8 | –0.7 | 0.34 | 0.062 | –0.16 |
| Greater than 52,100 | 48.4 | 3.99 | 48.0 | –0.4 | 1.07 | 0.683 | –0.11 | 48.5 | # | 0.09 | 0.862 | 0.00 |
| Percent who rent | | | | | | | | | | | | |
| 16 or less | 34.0 | 4.07 | 33.2 | –0.9 | 0.70 | 0.224 | –0.21 | 33.4 | –0.6 | 0.69 | 0.404 | –0.14 |
| 17–31 | 17.8 | 1.99 | 17.9 | 0.1 | 0.64 | 0.946 | 0.03 | 18.1 | 0.3 | 0.54 | 0.553 | 0.17 |
| 32–59 | 21.6 | 3.71 | 21.8 | 0.1 | 0.65 | 0.860 | 0.03 | 22.0 | 0.4 | 0.69 | 0.589 | 0.10 |
| Greater than 59 | 26.5 | 2.45 | 27.2 | 0.7 | 0.84 | 0.403 | 0.29 | 26.4 | –0.1 | 0.50 | 0.818 | –0.04 |

† Not applicable.
# Rounds to zero.
* Statistically significant with Bonferroni adjustment at $\alpha = 0.05$.
[1] Standard error.
[2] Metropolitan Statistical Area.
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 11-11. Household Study screener weighting effects for Missouri, by subgroup: 2003**

| Subgroup | Base weight—eligible Percent (1) | Base weight—eligible SE[1] (2) | Base weight—respondents Percent (3) | Base weight—respondents Bias Estimate (3)-(1) | Base weight—respondents Bias SE[1] | Base weight—respondents Bias p value | Base weight—respondents Bias ratio (3)-(1) / (2) | Nonresponse-adjusted weight—respondents Percent (4) | Nonresponse-adjusted weight—respondents Bias Estimate (4)-(1) | Nonresponse-adjusted weight—respondents Bias SE[1] | Nonresponse-adjusted weight—respondents Bias p value | Nonresponse-adjusted weight—respondents Bias ratio (4)-(1) / (2) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Total | 100.0 | † | 100.0 | † | † | † | † | 100.0 | † | † | † | † |
| MSA[2] status | | | | | | | | | | | | |
| Non-MSA | 32.7 | 1.86 | 34.0 | 1.3 | 0.45 | 0.008* | 0.72 | 32.8 | 0.1 | 0.26 | 0.718 | 0.05 |
| MSA | 67.3 | 1.86 | 66.0 | –1.3 | 0.45 | 0.008* | –0.72 | 67.3 | –0.1 | 0.26 | 0.718 | –0.05 |
| Average household size | | | | | | | | | | | | |
| 2.42 or less | 41.3 | 3.85 | 41.8 | 0.5 | 0.43 | 0.233 | 0.14 | 41.6 | 0.3 | 0.52 | 0.612 | 0.07 |
| 2.43–2.80 | 44.4 | 4.45 | 43.8 | –0.6 | 0.47 | 0.216 | –0.14 | 43.8 | –0.7 | 0.49 | 0.188 | –0.15 |
| Greater than 2.80 | 14.3 | 2.92 | 14.4 | 0.1 | 0.28 | 0.781 | 0.03 | 14.7 | 0.4 | 0.38 | 0.301 | 0.14 |
| Percent with less than high school education | | | | | | | | | | | | |
| 10.4 or less | 29.5 | 6.20 | 27.8 | –1.8 | 0.58 | 0.007* | –0.28 | 29.1 | –0.5 | 0.25 | 0.070 | –0.08 |
| 10.5–20.3 | 31.2 | 4.04 | 31.9 | 0.7 | 0.37 | 0.086 | 0.17 | 31.7 | 0.5 | 0.39 | 0.208 | 0.13 |
| 20.4–32.0 | 27.4 | 3.97 | 28.1 | 0.6 | 0.56 | 0.260 | 0.16 | 27.3 | –0.1 | 0.41 | 0.738 | –0.03 |
| Greater than 32.0 | 11.9 | 3.48 | 12.3 | 0.4 | 0.28 | 0.145 | 0.12 | 12.0 | 0.1 | 0.22 | 0.642 | 0.03 |
| Percent speaking Spanish but not English | | | | | | | | | | | | |
| 0 | 56.0 | 5.07 | 56.1 | 0.2 | 0.41 | 0.724 | 0.03 | 56.1 | 0.1 | 0.59 | 0.831 | 0.02 |
| 1–28 | 40.0 | 4.83 | 39.8 | –0.1 | 0.41 | 0.754 | –0.03 | 40.0 | 0.1 | 0.53 | 0.922 | 0.01 |
| Greater than 28 | 4.1 | 2.23 | 4.1 | # | 0.05 | 0.698 | –0.01 | 3.90 | –0.2 | 0.12 | 0.150 | –0.08 |
| Percent below 150 percent of poverty | | | | | | | | | | | | |
| 10.7 or less | 27.5 | 5.14 | 26.1 | 1.4 | 0.72 | 0.066 | –0.28 | 27.5 | # | 0.14 | 0.826 | 0.01 |
| 10.8–20.0 | 26.3 | 3.69 | 26.8 | 0.5 | 0.68 | 0.495 | 0.13 | 26.6 | 0.3 | 0.39 | 0.448 | 0.08 |
| 20.1–33.3 | 29.6 | 5.25 | 30.0 | 0.4 | 0.37 | 0.259 | 0.08 | 29.1 | –0.5 | 0.42 | 0.252 | –0.10 |
| Greater than 33.3 | 16.7 | 2.97 | 17.2 | 0.5 | 0.36 | 0.171 | 0.18 | 16.8 | 0.2 | 0.27 | 0.544 | 0.06 |

See notes at end of table.

**Table 11-11. Household Study screener weighting effects for Missouri, by subgroup: 2003—Continued**

| Subgroup | Base weight—eligible: Percent (1) | Base weight—eligible: SE[1] (2) | Base weight—respondents: Percent (3) | Base weight—respondents: Bias: Estimate (3)-(1) | Base weight—respondents: Bias: SE[1] | Base weight—respondents: Bias: *p* value | Base weight—respondents: Bias ratio (3)-(1) (2) | Nonresponse-adjusted weight—respondents: Percent (4) | Nonresponse-adjusted weight—respondents: Bias: Estimate (4)-(1) | Nonresponse-adjusted weight—respondents: Bias: SE[1] | Nonresponse-adjusted weight—respondents: Bias: *p* value | Nonresponse-adjusted weight—respondents: Bias ratio (4)-(1) (2) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Median income (in dollars) | | | | | | | | | | | | |
| 28,400 or less | 25.2 | 3.55 | 25.7 | 0.5 | 0.26 | 0.0[illegible] | 0.14 | 25.2 | # | 0.00 | 0.835 | 0.00 |
| 28,401–37,850 | 25.2 | 3.30 | 26.5 | 1.3 | 0.35 | 0.00[illegible] | 0.39 | 25.2 | # | 0.00 | 0.549 | 0.00 |
| 37,851–52,100 | 28.0 | 2.69 | 27.6 | –0.4 | 0.55 | 0.5[illegible] | –0.13 | 28.0 | # | 0.00 | 0.660 | 0.00 |
| Greater than 52,100 | 21.6 | 4.81 | 20.1 | –1.5 | 0.72 | 0.0[illegible] | –0.31 | 21.6 | # | 0.00 | 0.818 | 0.00 |
| Percent who rent | | | | | | | | | | | | |
| 16 or less | 37.1 | 4.54 | 35.8 | –1.3 | 0.72 | 0.0[illegible] | –0.28 | 36.7 | –0.4 | 0.33 | 0.211 | –0.10 |
| 17–31 | 25.2 | 4.13 | 26.2 | 0.9 | 0.56 | 0.1[illegible] | 0.23 | 25.4 | 0.2 | 0.26 | 0.395 | 0.06 |
| 32–59 | 27.8 | 4.18 | 28.1 | 0.3 | 0.49 | 0.5[illegible] | 0.07 | 28.0 | 0.2 | 0.40 | 0.629 | 0.05 |
| Greater than 59 | 9.9 | 1.94 | 10.0 | # | 0.31 | 0.8[illegible] | 0.03 | 9.9 | # | 0.25 | 0.962 | 0.01 |

† Not applicable.
# Rounds to zero.
* Statistically significant with Bonferroni adjustment at $\alpha = 0.05$.
[1] Standard error.
[2] Metropolitan Statistical Area.
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 11-12. Household Study screener weighting effects for New York, by subgroup: 2003**

| Subgroup | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio | | Bias | | | Bias ratio |
| | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | *p* value | (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | *p* value | (4)-(1) / (2) |
| Total | 100.0 | † | 100.0 | † | † | † | † | 100.0 | † | † | † | † |
| MSA[2] status | | | | | | | | | | | | |
| Non-MSA | 8.0 | 2.15 | 8.6 | 0.6 | 1.14 | 0.607 | 0.28 | 7.8 | –0.1 | 0.38 | 0.771 | –0.06 |
| MSA | 92.0 | 2.15 | 91.5 | –0.6 | 1.14 | 0.607 | –0.28 | 92.2 | 0.1 | 0.38 | 0.771 | 0.06 |
| Average household size | | | | | | | | | | | | |
| 2.42 or less | 32.4 | 3.81 | 30.4 | –2.0 | 0.70 | 0.012* | –0.51 | 30.5 | –1.9 | 0.74 | 0.019 | –0.49 |
| 2.43–2.80 | 29.8 | 3.87 | 31.7 | 2.0 | 0.89 | 0.040 | 0.51 | 31.9 | 2.1 | 0.92 | 0.036 | 0.54 |
| Greater than 2.80 | 37.8 | 4.82 | 37.8 | # | 1.18 | 0.988 | 0.00 | 37.6 | –0.2 | 1.10 | 0.856 | –0.04 |
| Percent with less than high school education | | | | | | | | | | | | |
| 10.4 or less | 36.9 | 1.91 | 35.4 | –1.5 | 1.07 | 0.182 | –0.78 | 36.8 | –0.1 | 0.86 | 0.903 | –0.06 |
| 10.5–20.3 | 20.7 | 2.94 | 20.3 | –0.4 | 0.88 | 0.662 | –0.13 | 20.0 | –0.7 | 0.86 | 0.456 | –0.22 |
| 20.4–32.0 | 21.9 | 2.25 | 22.3 | 0.4 | 0.90 | 0.676 | 0.17 | 22.0 | 0.1 | 0.86 | 0.883 | 0.06 |
| Greater than 32.0 | 20.5 | 2.36 | 22.0 | 1.5 | 0.91 | 0.115 | 0.64 | 21.2 | 0.6 | 0.75 | 0.405 | 0.27 |
| Percent speaking Spanish but not English | | | | | | | | | | | | |
| 0 | 41.5 | 4.97 | 43.0 | 1.5 | 1.16 | 0.202 | 0.31 | 42.4 | 1.0 | 1.05 | 0.364 | 0.20 |
| 1–28 | 32.2 | 4.90 | 29.6 | –2.6 | 1.16 | 0.039 | –0.53 | 30.4 | –1.7 | 1.15 | 0.148 | –0.36 |
| Greater than 28 | 26.4 | 3.20 | 27.4 | 1.1 | 0.91 | 0.264 | 0.33 | 27.1 | 0.8 | 0.72 | 0.300 | 0.24 |
| Percent below 150 percent of poverty | | | | | | | | | | | | |
| 10.7 or less | 40.0 | 3.57 | 38.9 | –1.1 | 0.81 | 0.186 | –0.31 | 40.0 | # | 0.89 | 0.986 | –0.01 |
| 10.8–20.0 | 21.8 | 3.62 | 19.8 | –2.0 | 0.93 | 0.045 | –0.55 | 20.0 | –1.8 | 0.98 | 0.091 | –0.49 |
| 20.1–33.3 | 18.2 | 4.56 | 19.8 | 1.6 | 0.77 | 0.049 | 0.36 | 19.0 | 0.8 | 0.60 | 0.197 | 0.18 |
| Greater than 33.3 | 20.0 | 3.68 | 21.5 | 1.5 | 0.80 | 0.082 | 0.40 | 21.0 | 1.0 | 0.75 | 0.212 | 0.26 |

See notes at end of table.

**Table 11-12. Household Study screener weighting effects for New York, by subgroup: 2003—Continued**

| Subgroup | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio | | Bias | | | Bias ratio |
| | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | $p$ value | (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | $p$ value | (4)-(1) / (2) |
| Median income (in dollars) | | | | | | | | | | | | |
| 28,400 or less | 13.1 | 2.65 | 14.6 | 1.4 | 0.59 | 0.027 | 0.54 | 14.0 | 0.9 | 0.62 | 0.156 | 0.34 |
| 28,401–37,850 | 23.4 | 3.56 | 24.3 | 0.9 | 0.88 | 0.303 | 0.26 | 23.5 | 0.2 | 0.69 | 0.817 | 0.04 |
| 37,851–52,100 | 19.3 | 2.77 | 19.3 | # | 0.86 | 0.974 | –0.01 | 18.8 | –0.5 | 0.92 | 0.593 | –0.18 |
| Greater than 52,100 | 44.2 | 3.43 | 41.9 | –2.3 | 1.12 | 0.051 | –0.68 | 43.7 | –0.6 | 0.95 | 0.547 | –0.17 |
| Percent who rent | | | | | | | | | | | | |
| 16 or less | 32.6 | 4.05 | 33.3 | 0.7 | 0.93 | 0.471 | 0.17 | 34.3 | 1.6 | 1.03 | 0.125 | 0.41 |
| 17–31 | 15.4 | 2.67 | 15.7 | 0.3 | 0.83 | 0.769 | 0.09 | 15.5 | 0.1 | 0.65 | 0.909 | 0.03 |
| 32–59 | 14.7 | 3.53 | 15.0 | 0.3 | 1.50 | 0.841 | 0.09 | 14.5 | –0.2 | 1.48 | 0.888 | –0.06 |
| Greater than 59 | 37.3 | 3.82 | 36.0 | –1.3 | 1.35 | 0.368 | –0.33 | 35.8 | –1.5 | 1.42 | 0.299 | –0.39 |

† Not applicable.
# Rounds to zero.
* Statistically significant with Bonferroni adjustment at $\alpha = 0.05$.
[1] Standard error.
[2] Metropolitan Statistical Area.
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 11-13. Household Study screener weighting effects for Oklahoma, by subgroup: 2003**

| Subgroup | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio | | Bias | | | Bias ratio |
| | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | *p* value | (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | *p* value | (4)-(1) / (2) |
| Total | 100.0 | † | 100.0 | † | † | † | † | 100.0 | † | † | † | † |
| MSA[2] status | | | | | | | | | | | | |
| Non-MSA | 38.3 | 3.18 | 38.2 | –0.1 | 0.61 | 0.937 | –0.02 | 37.7 | –0.6 | 0.22 | 0.023 | –0.17 |
| MSA | 61.7 | 3.18 | 61.8 | 0.1 | 0.61 | 0.937 | 0.02 | 62.3 | 0.6 | 0.22 | 0.023 | 0.17 |
| Average household size | | | | | | | | | | | | |
| 2.42 or less | 33.8 | 6.29 | 33.5 | –0.3 | 0.51 | 0.586 | –0.04 | 33.7 | –0.1 | 0.32 | 0.882 | –0.01 |
| 2.43–2.80 | 53.3 | 4.76 | 53.6 | 0.4 | 0.42 | 0.380 | 0.08 | 53.0 | –0.3 | 0.35 | 0.454 | –0.06 |
| Greater than 2.80 | 13.0 | 3.39 | 12.9 | –0.1 | 0.27 | 0.727 | –0.03 | 13.3 | 0.3 | 0.15 | 0.049 | 0.09 |
| Percent with less than high school education | | | | | | | | | | | | |
| 10.4 or less | 22.5 | 2.20 | 22.0 | –0.5 | 0.42 | 0.232 | –0.24 | 22.8 | 0.3 | 0.22 | 0.236 | 0.13 |
| 10.5–20.3 | 26.1 | 4.31 | 25.5 | –0.5 | 0.36 | 0.174 | –0.12 | 26.0 | –0.1 | 0.27 | 0.773 | –0.02 |
| 20.4–32.0 | 31.0 | 5.48 | 31.0 | # | 0.41 | 0.982 | 0.00 | 30.7 | –0.3 | 0.26 | 0.256 | –0.05 |
| Greater than 32.0 | 20.5 | 5.93 | 21.5 | 1.0 | 0.40 | 0.021 | 0.17 | 20.6 | 0.1 | 0.11 | 0.352 | 0.02 |
| Percent speaking Spanish but not English | | | | | | | | | | | | |
| 0 | 43.5 | 4.73 | 43.2 | –0.4 | 0.52 | 0.486 | –0.08 | 43.3 | 0.3 | 0.36 | 0.465 | –0.06 |
| 1–28 | 45.2 | 4.94 | 45.6 | 0.5 | 0.42 | 0.292 | 0.09 | 45.6 | 0.5 | 0.30 | 0.160 | 0.09 |
| Greater than 28 | 11.3 | 1.98 | 11.2 | –0.1 | 0.30 | 0.784 | –0.05 | 11.1 | –0.2 | 0.18 | 0.345 | –0.09 |
| Percent below 150 percent of poverty | | | | | | | | | | | | |
| 10.7 or less | 16.8 | 3.20 | 15.9 | –0.9 | 0.41 | 0.033 | –0.29 | 16.8 | # | 0.00 | 0.832 | 0.00 |
| 10.8–20.0 | 19.9 | 3.78 | 19.8 | –0.1 | 0.34 | 0.857 | –0.02 | 19.9 | # | 0.00 | 0.926 | 0.00 |
| 20.1–33.3 | 26.1 | 3.26 | 25.8 | –0.3 | 0.47 | 0.497 | –0.10 | 26.1 | # | 0.00 | 0.911 | 0.00 |
| Greater than 33.3 | 37.3 | 4.58 | 38.6 | 1.3 | 0.43 | 0.006 | 0.29 | 37.3 | # | 0.00 | 0.929 | 0.00 |

See notes at end of table.

## Table 11-13. Household Study screener weighting effects for Oklahoma, by subgroup: 2003—Continued

| | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio | | Bias | | | Bias ratio |
| Subgroup | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | *p* value | (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | *p* value | (4)-(1) / (2) |
| Median income (in dollars) | | | | | | | | | | | | |
| 28,400 or less | 40.2 | 4.53 | 41.2 | 1.0 | 0.45 | 0.04 | 0.22 | 40.3 | 0.1 | 0.06 | 0.361 | 0.01 |
| 28,401–37,850 | 25.9 | 2.83 | 26.1 | 0.2 | 0.40 | 0.63 | 0.07 | 25.8 | –0.1 | 0.08 | 0.561 | –0.02 |
| 37,851–52,100 | 20.3 | 3.83 | 20.1 | –0.1 | 0.41 | 0.79 | –0.03 | 20.2 | # | 0.04 | 0.796 | 0.00 |
| Greater than 52,100 | 13.7 | 3.42 | 12.6 | –1.1 | 0.25 | 0.001 | –0.31 | 13.7 | # | 0.00 | 0.616 | 0.00 |
| Percent who rent | | | | | | | | | | | | |
| 16 or less | 24.3 | 4.45 | 24.0 | –0.3 | 0.30 | 0.34 | –0.07 | 24.7 | 0.4 | 0.19 | 0.067 | 0.08 |
| 17–31 | 32.9 | 5.18 | 32.6 | –0.4 | 0.39 | 0.36 | –0.07 | 32.5 | –0.5 | 0.23 | 0.049 | –0.09 |
| 32–59 | 32.1 | 6.00 | 32.9 | 0.7 | 0.33 | 0.04 | 0.12 | 32.4 | 0.2 | 0.18 | 0.234 | 0.04 |
| Greater than 59 | 10.6 | 2.43 | 10.6 | –0.1 | 0.19 | 0.70 | –0.03 | 10.5 | –0.1 | 0.13 | 0.450 | –0.04 |

† Not applicable.
# Rounds to zero.
* Statistically significant with Bonferroni adjustment at $\alpha = 0.05$.
[1] Standard error.
[2] Metropolitan Statistical Area.
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

#### 11.1.4.2 Evaluating Bias Owing to Background Questionnaire Nonresponse

The nonresponse bias potentially resulting from background questionnaire nonresponse was evaluated in the same manner as the analysis of screener nonresponse. Additional variables were available for the analysis. Variables known for both background questionnaire respondents and nonrespondents are shown in table 11-14 and come from three sources: Census 2000 PL-94 county-level data, Census 2000 SF3A block group-level data, and the screener. The SF3A variables were categorized the same as in the screener analysis, with approximately equal sample size in each category.

The bivariate analysis, multivariate analysis, and weighting adjustment effects are presented in sections 11.1.4.2.1 through 11.1.4.2.3, respectively.

##### 11.1.4.2.1 Background Questionnaire Bivariate Analysis

The distribution of background questionnaire respondents was compared with the distribution of all eligible sample persons for each of the table 11-14 variables. Weighted percentages and standard errors were calculated using replicated composite background questionnaire base weights to reflect the complex sample design. To test the significance of the relationship between response status and each of the table 11-14 variables, a Rao-Scott chi-square (RS3) test of independence was performed. In addition, an estimate of bias was calculated for each domain. Bias was estimated as $Bias(\bar{y}_R) = (1 - W_R)(\bar{Y}_R - \bar{Y}_N)$, where $W_R$ is the weighted unit background questionnaire response rate (75.6 percent), $\bar{Y}_R$ is the weighted estimate of the domain percentage for respondents, and $\bar{Y}_N$ is the weighted estimate of the domain percentage for nonrespondents. A *t* test was performed, using a simple Bonferroni adjustment, to determine whether the bias was significantly different from 0.

**Table 11-14. Household Study variables used in background questionnaire nonresponse bias analysis, by source and values: 2003**

| Variable description | Source[1] | Values |
|---|---|---|
| Census region | PL-94 | 1: Northeast<br>2: Midwest<br>3: South<br>4: West |
| MSA[2] status | PL-94 | 1: MSA<br>2: Non-MSA |
| Average household size | SF3A | 1: 2.42 or less<br>2: 2.43–2.80<br>3: Greater than 2.80 |
| Percent aged 25+ with less than a high school education | SF3A | 1: 10.4 or less<br>2: 10.5–20.3<br>3: 20.4–32.0<br>4: Greater than 32.0 |
| Percent aged 5–64 speaking Spanish at home and English not well or not at all | SF3A | 1: 0<br>2: 1–28<br>3: Greater than 28 |
| Percent below 150 percent of poverty | SF3A | 1: 10.7 or less<br>2: 10.8–20.0<br>3: 20.1–33.3<br>4: Greater than 33.3 |
| Median income (in dollars) | SF3A | 1: 28,400 or less<br>2: 28,401–37,850<br>3: 37,851–52,100<br>4: Greater than 52,100 |
| Percent who rent | SF3A | 1: 16 or less<br>2: 17–31<br>3: 32–59<br>4: Greater than 59 |
| Age (years) | Screener | 1: 16–29<br>2: 30–49<br>3: 50–69<br>4: Greater than 70 |
| Race/ethnicity | Screener | 1: Hispanic<br>2: Non-Hispanic Black only<br>3: Other[3] |
| Gender | Screener | 1: Male<br>2: Female |

[1] The SF3A (Summary File 3A) and PL-94 (county-level Public Law 94) variables provide relevant statistics for the block group or the county of the sampled dwelling unit.
[2] Metropolitan Statistical Area.
[3] Includes non-Hispanic White, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy; U.S. Department of Commerce, U.S. Census Bureau, Decennial Census, 2000.

The results of the chi-square analysis are presented in table 11-15. At the 5 percent α-level, all analysis variables have a significant relationship to background questionnaire response status except average household size and the percentage of sample persons aged 5 to 64 who speak Spanish at home but English not well or not at all. The results of the *t* tests for bias (shown in table 11-16) are consistent with the chi-square analysis. For the same set of variables, the bias in estimating the domain percentages is significantly different from 0 for at least one domain. There is evidence of differential response rates among subgroups (table 11-1), contributing to the differential distributions of respondents and nonrespondents. For instance, sample persons aged 16 to 29 had a relatively high weighted background questionnaire response rate of 80.5 percent. Using only respondents, without weighting adjustments, would result in an overestimate of this domain percentage by 1.6 (or 6.59 percent), as shown in table 11-16. For this domain, the estimated bias is large in relation to the sampling error; the ratio of bias to the standard error is 3.86. Gender also shows a large bias ratio, with estimates of bias more than twice the standard error of the estimated percentages.

Although the relationships between response status and the table 11-14 variables are significant, the differences between the distributions of respondents and eligible sample persons are minor. The absolute bias is less than 2 for all estimated domain percentages. In addition, many of the table 11-14 variables were used in weighting adjustments (including gender and age, which showed indications of bias), and so differences between respondents and eligible sample persons were reduced through the weighting process (refer to section 11.1.4.2.3). Therefore, the bivariate analysis indicates minimal potential for bias at the background questionnaire level, and thus minimal impact of background questionnaire nonresponse on literacy scores, assuming that literacy scores are highly correlated with the variables used in the weighting adjustments.

#### 11.1.4.2.2 Background Questionnaire Multivariate Analysis

The CHAID software was used in the background questionnaire multivariate analysis to explain differential response rates. For more information on CHAID, refer to section 11.1.4.1.2.

**Table 11-15. Household Study sample distribution of background questionnaire respondents versus eligible sample persons, by analysis domain: 2003**

| Analysis domain | Respondents | | | Eligibles | | | Chi-square | |
|---|---|---|---|---|---|---|---|---|
| | Number | Domain percent | Standard error | Number | Domain percent | Standard error | Statistic | *p* value |
| Region | | | | | | | | |
| Northeast | 3,765 | 16.3 | 1.41 | 4,936 | 16.5 | 1.45 | 9.11 | 0.019 |
| Midwest | 3,612 | 25.5 | 1.37 | 4,500 | 24.7 | 1.31 | | |
| South | 8,270 | 36.3 | 1.76 | 10,510 | 36.6 | 1.71 | | |
| West | 2,894 | 21.9 | 1.02 | 3,786 | 22.2 | 1.02 | | |
| MSA[1] status | | | | | | | | |
| Non-MSA | 3,897 | 21.6 | 1.51 | 4,835 | 20.7 | 1.42 | 10.44 | 0.001 |
| MSA | 14,644 | 78.4 | 1.51 | 18,897 | 79.4 | 1.42 | | |
| Average household size | | | | | | | | |
| 2.42 or less | 5,931 | 30.5 | 1.56 | 7,474 | 29.9 | 1.51 | 3.62 | 0.149 |
| 2.43–2.80 | 6,589 | 35.0 | 1.45 | 8,465 | 35.2 | 1.39 | | |
| Greater than 2.80 | 6,021 | 34.4 | 2.08 | 7,793 | 34.8 | 1.99 | | |
| Percent with less than high school education | | | | | | | | |
| 10.4 or less | 4,727 | 30.9 | 1.77 | 6,265 | 31.8 | 1.79 | 17.89 | 0.000 |
| 10.5–20.3 | 4,656 | 27.3 | 1.41 | 6,066 | 27.6 | 1.35 | | |
| 20.4–32.0 | 4,511 | 22.8 | 1.43 | 5,731 | 22.4 | 1.34 | | |
| Greater than 32.0 | 4,647 | 19.1 | 1.54 | 5,670 | 18.2 | 1.41 | | |
| Percent speaking Spanish but not English | | | | | | | | |
| 0 | 7,782 | 41.9 | 2.16 | 9,971 | 41.9 | 2.09 | 4.96 | 0.072 |
| 1–28 | 6,107 | 35.8 | 1.60 | 7,952 | 36.3 | 1.56 | | |
| Greater than 28 | 4,652 | 22.4 | 1.85 | 5,809 | 21.8 | 1.68 | | |
| Percent below 150 percent of poverty | | | | | | | | |
| 10.7 or less | 4,961 | 32.9 | 1.53 | 6,669 | 34.5 | 1.49 | 53.72 | 0.000 |
| 10.8–20.0 | 4,379 | 26.0 | 1.70 | 5,713 | 26.1 | 1.66 | | |
| 20.1–33.3 | 4,355 | 23.3 | 1.72 | 5,502 | 22.7 | 1.63 | | |
| Greater than 33.3 | 4,846 | 17.9 | 1.24 | 5,848 | 16.7 | 1.18 | | |
| Median income (in dollars) | | | | | | | | |
| 28,400 or less | 4,873 | 19.0 | 1.48 | 5,838 | 17.6 | 1.39 | 59.47 | 0.000 |
| 28,401–37,850 | 4,475 | 24.3 | 1.38 | 5,675 | 23.6 | 1.33 | | |
| 37,851–52,100 | 4,451 | 25.6 | 1.60 | 5,800 | 25.8 | 1.55 | | |
| Greater than 52,100 | 4,742 | 31.1 | 1.82 | 6,419 | 33.0 | 1.79 | | |
| Percent who rent | | | | | | | | |
| 16 or less | 4,984 | 32.9 | 1.66 | 6,662 | 34.2 | 1.61 | 22.37 | 0.000 |
| 17–31 | 4,857 | 27.6 | 1.63 | 6,231 | 27.4 | 1.55 | | |
| 32–59 | 4,561 | 22.4 | 1.22 | 5,726 | 22.0 | 1.21 | | |
| Greater than 59 | 4,139 | 17.2 | 0.86 | 5,113 | 16.4 | 0.78 | | |
| Age (years) | | | | | | | | |
| 16–29 | 4,712 | 26.2 | 0.49 | 5,726 | 24.6 | 0.42 | 41.67 | 0.000 |
| 30–49 | 7,261 | 39.1 | 0.64 | 9,419 | 40.0 | 0.56 | | |
| 50–69 | 4,571 | 24.8 | 0.47 | 5,940 | 25.4 | 0.45 | | |
| 70+ | 1,997 | 9.9 | 0.40 | 2,647 | 10.0 | 0.34 | | |
| Gender | | | | | | | | |
| Male | 8,028 | 46.0 | 0.51 | 10,660 | 47.8 | 0.36 | 32.22 | 0.000 |
| Female | 10,513 | 54.0 | 0.51 | 13,072 | 52.2 | 0.36 | | |
| Race/ethnicity | | | | | | | | |
| Hispanic | 3,194 | 14.1 | 1.40 | 3,945 | 13.4 | 1.27 | 19.81 | 0.000 |
| Non-Hispanic Black only | 3,504 | 11.6 | 0.95 | 4,328 | 11.0 | 0.92 | | |
| Other[2] | 11,843 | 74.4 | 1.51 | 15,459 | 75.6 | 1.37 | | |

[1] Metropolitan Statistical Area.
[2] Includes non-Hispanic White, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races.
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 11-16. Household Study estimates of background questionnaire nonresponse bias, by analysis domain: 2003**

| Analysis domain | Eligibles: Domain percent (1) | Eligibles: SE[1] (2) | Respondent domain percent (3) | Nonrespondent domain percent (4) | Bias: Estimate (5) | Bias: SE[1] (6) | Bias: *p* value (7) | Relative bias (5)/(1)*100 | Bias ratio (5)/(2) |
|---|---|---|---|---|---|---|---|---|---|
| Region | | | | | | | | | |
| Northeast | 16.5 | 1.45 | 16.3 | 17.1 | –0.2 | 0.18 | 0.344 | –1.09 | –0.12 |
| Midwest | 24.7 | 1.31 | 25.5 | 22.2 | 0.8 | 0.21 | 0.000* | 3.32 | 0.63 |
| South | 36.6 | 1.71 | 36.3 | 37.7 | –0.4 | 0.34 | 0.318 | –0.96 | –0.20 |
| West | 22.2 | 1.02 | 21.9 | 23.1 | –0.3 | 0.26 | 0.254 | –1.31 | –0.28 |
| MSA[2] status | | | | | | | | | |
| Non-MSA | 20.7 | 1.42 | 21.6 | 17.6 | 1.0 | 0.31 | 0.002* | 4.75 | 0.69 |
| MSA | 79.4 | 1.42 | 78.4 | 82.4 | –1.0 | 0.31 | 0.002* | –1.24 | –0.69 |
| Average household size | | | | | | | | | |
| 2.42 or less | 29.9 | 1.51 | 30.5 | 28.1 | 0.6 | 0.32 | 0.066 | 1.97 | 0.39 |
| 2.43–2.80 | 35.2 | 1.39 | 35.0 | 35.9 | –0.2 | 0.28 | 0.473 | –0.57 | –0.14 |
| Greater than 2.80 | 34.8 | 1.99 | 34.4 | 36.0 | –0.4 | 0.35 | 0.265 | –1.12 | –0.20 |
| Percent with less than high school education | | | | | | | | | |
| 10.4 or less | 31.8 | 1.79 | 30.9 | 34.9 | -1.0 | 0.33 | 0.005* | –3.08 | –0.55 |
| 10.5–20.3 | 27.6 | 1.35 | 27.3 | 28.7 | –0.3 | 0.29 | 0.237 | –1.23 | –0.25 |
| 20.4–32.0 | 22.4 | 1.34 | 22.8 | 21.0 | 0.5 | 0.29 | 0.123 | 2.01 | 0.34 |
| Greater than 32.0 | 18.2 | 1.41 | 19.1 | 15.5 | 0.9 | 0.26 | 0.001* | 4.84 | 0.62 |
| Percent speaking Spanish but not English | | | | | | | | | |
| 0 | 41.9 | 2.09 | 41.9 | 41.8 | # | 0.28 | 0.969 | 0.02 | 0.00 |
| 1–28 | 36.3 | 1.56 | 35.8 | 38.0 | –0.6 | 0.28 | 0.053 | –1.52 | –0.35 |
| Greater than 28 | 21.8 | 1.68 | 22.4 | 20.2 | 0.5 | 0.30 | 0.075 | 2.47 | 0.32 |
| Percent below 150 percent of poverty | | | | | | | | | |
| 10.7 or less | 34.5 | 1.49 | 32.9 | 39.6 | –1.7 | 0.28 | 0.000* | –4.78 | –1.11 |
| 10.8–20.0 | 26.1 | 1.66 | 26.0 | 26.3 | –0.1 | 0.24 | 0.737 | –0.31 | –0.05 |
| 20.1–33.3 | 22.7 | 1.63 | 23.3 | 21.1 | 0.5 | 0.26 | 0.044 | 2.33 | 0.33 |
| Greater than 33.3 | 16.7 | 1.18 | 17.9 | 13.0 | 1.2 | 0.18 | 0.000* | 7.19 | 1.02 |
| Median income (in dollars) | | | | | | | | | |
| 28,400 or less | 17.6 | 1.39 | 19.0 | 13.2 | 1.4 | 0.24 | 0.000* | 8.13 | 1.03 |
| 28,401–37,850 | 23.6 | 1.33 | 24.3 | 21.4 | 0.7 | 0.25 | 0.007* | 3.01 | 0.53 |
| 37,851–52,100 | 25.8 | 1.55 | 25.6 | 26.6 | –0.3 | 0.29 | 0.395 | –0.97 | –0.16 |
| Greater than 52,100 | 33.0 | 1.79 | 31.1 | 38.9 | –1.9 | 0.32 | 0.000* | –5.73 | –1.06 |
| Percent who rent | | | | | | | | | |
| 16 or less | 34.2 | 1.61 | 32.9 | 38.2 | –1.3 | 0.34 | 0.000* | –3.80 | –0.81 |
| 17–31 | 27.4 | 1.55 | 27.6 | 27.1 | 0.1 | 0.30 | 0.687 | 0.44 | 0.08 |
| 32–59 | 22.0 | 1.21 | 22.4 | 20.8 | 0.4 | 0.24 | 0.117 | 1.73 | 0.31 |
| Greater than 59 | 16.4 | 0.78 | 17.2 | 13.9 | 0.8 | 0.18 | 0.000* | 4.88 | 1.03 |
| Age (years) | | | | | | | | | |
| 16–29 | 24.6 | 0.42 | 26.2 | 19.6 | 1.6 | 0.23 | 0.000* | 6.59 | 3.86 |
| 30–49 | 40.0 | 0.56 | 39.1 | 42.8 | –0.9 | 0.32 | 0.006* | –2.23 | –1.59 |
| 50–69 | 25.4 | 0.45 | 24.8 | 27.1 | –0.6 | 0.25 | 0.034 | –2.17 | –1.22 |
| 70+ | 10.0 | 0.34 | 9.9 | 10.6 | –0.2 | 0.16 | 0.252 | –1.79 | –0.53 |
| Gender | | | | | | | | | |
| Male | 47.8 | 0.36 | 46.0 | 53.3 | –1.8 | 0.32 | 0.000* | –3.74 | –4.97 |
| Female | 52.2 | 0.36 | 54.0 | 46.7 | 1.8 | 0.32 | 0.004* | 3.43 | 4.97 |
| Race/ethnicity | | | | | | | | | |
| Hispanic | 13.4 | 1.27 | 14.1 | 11.3 | 0.7 | 0.22 | 0.003* | 5.08 | 0.54 |
| Non-Hispanic Black only | 11.0 | 0.92 | 11.6 | 9.4 | 0.5 | 0.18 | 0.000* | 4.90 | 0.59 |
| Other[3] | 75.6 | 1.37 | 74.4 | 79.4 | –1.2 | 0.29 | 0.000* | –1.61 | –0.89 |

[#] Rounds to zero.
* Statistically significant with Bonferroni adjustment at $\alpha = 0.05$.
[1] Standard error.
[2] Metropolitan Statistical Area.
[3] Includes non-Hispanic White, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races.
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

CHAID was run with background questionnaire response status as the dependent variable and the table 11-14 variables as the independent variables. Cell sizes were limited to 230 or more sample persons (approximately 1 percent of the sample), and up to three-way interactions were allowed (three tree levels). The resulting tree is shown in figure 11-2 and summarized in table 11-17. Fifteen cells were formed, with weighted response rates ranging from 67.4 percent to 86.9 percent. The lowest response rate was for males aged 30+ in segments with high median income (greater than $52,100). The highest response rate was for persons aged 16 to 29 in segments with moderately low median income ($28,401 to $37,850) and large average household size (greater than 2.8). Median income was the dominant variable in distinguishing response rate groups, which is consistent with the results of the bivariate analysis. Gender, age, region, household size, and race/ethnicity were also significant contributors to the CHAID tree.

A logistic regression main-effect model was used to identify significant effects on background questionnaire response propensity. Only main effects were included because of limited degrees of freedom and because interactions were identified in the previous CHAID analysis. Background questionnaire response status was used as the binary dependent variable, and the table 11-14 variables were used as the predictors. The results of the logistic regression analysis are presented in table 11-18. Five variables—region, median income, age, gender, and race/ethnicity—were found to be significantly related to response propensity at the 5 percent $\alpha$-level. The regression coefficient estimates for these five variables are provided in table 11-18 to show the direction of the relationship with response propensity. For instance, the table shows that males are significantly less likely to respond than females. The results are consistent with the CHAID analysis: The five variables were included in the first three levels of the CHAID tree. All variables found to be significantly related to response propensity in the multivariate analysis, including household size from CHAID, were used in the background questionnaire weighting adjustments.[2] Thus, the potential for nonresponse bias suggested by the multivariate analysis should be reduced through the weighting adjustments, as described in the next section.

[2]A more detailed dwelling unit-level household count was used for the background questionnaire weighting adjustments in place of the segment-level variable for average household size included here.

**Figure 11-2. Household Study multivariate CHAID analysis of background questionnaire response indicators: 2003**

| Median income (in dollars) | | | | | | |
|---|---|---|---|---|---|---|
| 28,400 or less | 81.7% (5,838) | Gender | | | | |
| | | Male | 77.7% (2,483) | Region | | |
| | | | | Northeast, Midwest | 82.8% (754) | Cell 1 |
| | | | | South, West | 75.7% (1,729) | Cell 2 |
| | | Female | 85.1% (3,355) | Age (years) | | |
| | | | | 16–69 | 86.1% (2,888) | Cell 3 |
| | | | | 70 or more | 79.3% (467) | Cell 4 |
| 28,401–37,850 | 77.8% (5,675) | Age (years) | | | | |
| | | 16–29 | 82.1% (1,485) | Household size | | |
| | | | | 2.80 or less | 80.8% (1,084) | Cell 5 |
| | | | | Greater than 2.80 | 86.9% (401) | Cell 6 |
| | | 30 or more | 76.3% (4,190) | Region | | |
| | | | | Northeast, South | 73.5% (2,715) | Cell 7 |
| | | | | Midwest, West | 80.0% (1,475) | Cell 8 |
| 37,851–52,100 | 74.8% (5,800) | Age (years) | | | | |
| | | 16–29 | 81.2% (1,431) | Race/ethnicity | | |
| | | | | Hispanic, non-Hispanic Black | 85.9% (575) | Cell 9 |
| | | | | Other[1] | 79.4% (856) | Cell 10 |
| | | 30 or more | 72.6% (4,369) | Race/ethnicity | | |
| | | | | Hispanic, non-Hispanic Black | 77.7% (1,188) | Cell 11 |
| | | | | Other | 71.5% (3,181) | Cell 12 |
| Greater than 52,100 | 71.2% (6,419) | Gender | | | | |
| | | Male | 68.6% (3,011) | Age (years) | | |
| | | | | 16–29 | 73.2% (600) | Cell 13 |
| | | | | 30 or more | 67.4% (2,411) | Cell 14 |
| | | Female | 73.8% (3,408) | Cell 15 | | |

Overall weighted response rate = 75.6 percent
Total number of eligibles = 23,732

[1]Includes non-Hispanic White, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races.
NOTE: All percentages are weighted response rates and the numbers inside the parentheses are sample sizes.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 11-17. Household Study multivariate CHAID analysis of background questionnaire response indicators, by response cell: 2003**

| Response cell | Number of eligibles | Number of respondents | Unweighted response rate (percent) | Weighted response rate (percent) |
|---|---|---|---|---|
| Overall | 23,732 | 18,541 | 78.1 | 75.6 |
| 1 | 754 | 626 | 83.0 | 82.8 |
| 2 | 1,729 | 1,365 | 79.0 | 75.7 |
| 3 | 2,888 | 2,515 | 87.1 | 86.1 |
| 4 | 467 | 367 | 78.6 | 79.3 |
| 5 | 1,084 | 885 | 81.6 | 80.8 |
| 6 | 401 | 349 | 87.0 | 87.0 |
| 7 | 2,715 | 2,062 | 76.0 | 73.5 |
| 8 | 1,475 | 1,179 | 79.9 | 80.0 |
| 9 | 575 | 484 | 84.2 | 85.9 |
| 10 | 856 | 689 | 80.5 | 79.4 |
| 11 | 1,188 | 936 | 78.8 | 77.7 |
| 12 | 3,181 | 2,342 | 73.6 | 71.5 |
| 13 | 600 | 452 | 75.3 | 73.2 |
| 14 | 2,411 | 1,689 | 70.1 | 67.4 |
| 15 | 3,408 | 2,601 | 76.3 | 73.8 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 11-18. Household Study multivariate logistic regression analysis of background questionnaire response indicators, by predictor: 2003**

| | *F* test | | | | Regression coefficients | |
|---|---|---|---|---|---|---|
| Test | *F* statistic | Numerator *df*[1] | Denominator *df*[1] | *p* value | Estimate | *p* value |
| Overall fit | 6.26 | 26 | 36 | 0.000 | † | † |
| Region | 7.64 | 3 | 59 | 0.000 | † | † |
| Northeast | † | † | † | † | 0.08 | 0.305 |
| Midwest | † | † | † | † | 0.25 | 0.001 |
| South | † | † | † | † | –0.02 | 0.790 |
| MSA[2] status | 3.92 | 1 | 61 | 0.052 | † | † |
| Average household size | 0.49 | 2 | 60 | 0.615 | † | † |
| Percent with less than high school education | 0.49 | 3 | 59 | 0.691 | † | † |
| Percent speaking Spanish but not English | 0.36 | 2 | 60 | 0.703 | † | † |
| Percent below 150 percent of poverty | 0.12 | 3 | 59 | 0.945 | † | † |
| Median income (in dollars) | 4.22 | 3 | 59 | 0.009 | † | † |
| 28,400 or less | † | † | † | † | 0.58 | 0.001 |
| 28,401–37,850 | † | † | † | † | 0.33 | 0.006 |
| 37,851–52,100 | † | † | † | † | 0.16 | 0.086 |
| Percent who rent | 0.33 | 3 | 59 | 0.806 | † | † |
| Age | 12.40 | 3 | 59 | 0.000 | † | † |
| 16–29 | † | † | † | † | 0.36 | 0.000 |
| 30–49 | † | † | † | † | 0.02 | 0.788 |
| 50–69 | † | † | † | † | 0.03 | 0.697 |
| Gender | 33.99 | 1 | 61 | 0.000 | † | † |
| Male | † | † | † | † | –0.30 | 0.000 |
| Race/ethnicity | 3.41 | 2 | 60 | 0.039 | † | † |
| Hispanic | † | † | † | † | 0.22 | 0.022 |
| Non-Hispanic Black only | † | † | † | † | 0.17 | 0.075 |

† Not applicable.
# Rounds to zero.
1 Degrees of freedom.
2 Metropolitan Statistical Area.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

#### 11.1.4.2.3 Potential for Background Questionnaire Nonresponse Bias Remaining After Weighting Procedures

As described in section 12.1.4, the weighting procedures for the main NAAL background questionnaire were implemented to reduce the potential for nonresponse bias by creating nonresponse adjustment classes for which respondents' literacy-related characteristics are similar to those of nonrespondents. Other background questionnaire weighting steps were conducted to calibrate the weights to known totals from the 2003 Current Population Survey and then combine the NAAL and SAAL samples through composite weighting. Steps performed after the nonresponse adjustment are expected to have little impact on nonresponse bias and thus are not included in this analysis.

Tables 11-19 to 11-25 show and test the change in the distribution of the sample cases before and after the background questionnaire nonresponse adjustment for the national household sample and each of the six participating states. *T* tests were performed using a Bonferroni adjustment to test whether the difference in estimated percentages is significantly different from 0. To help determine whether statistically significant results are also meaningful, calculations of bias ratios are also included. The checks were performed separately for the national NAAL household sample and for each of the six SAAL states, to reflect the weighting process. Unlike the results in table 11-15, these comparisons use the actual survey weights, which were processed separately for each sample and include screener nonresponse adjustments. The following comparisons were made for each of the analysis variables in table 11-14:

- Comparison of distributions from background questionnaire base weights with those for background questionnaire respondents only, to check for differences owing to background questionnaire nonresponse, and
- Comparison of distributions from background questionnaire base weights with those from the background questionnaire nonresponse-adjusted weights, to check for differences remaining after nonresponse adjustment to the background questionnaire.

The *p* values resulting from the first set of comparisons indicate a significant difference between the eligible sample persons and respondents for most of the subgroups when background questionnaire base weights are used. A nonresponse adjustment was necessary to reduce the bias in estimates based on respondents only.

As shown in table 11-19, significant differences between eligible NAAL sample persons and respondents remain for some categories of region, MSA status, and median income after the nonresponse adjustment. However, the absolute bias in the percentage distribution of the majority of table 11-14 variables decreased and is less than 2 percentage points for all domains. In addition, the bias is less than twice the standard error of the estimated percentage for all statistically significant differences, so the differences appear to be minor. Tables 11-20 to 11-25 show similar results for the six SAAL states. Although the bias remains significantly different from 0 for some domains, after the nonresponse adjustment it is never more than twice the standard error of the estimated percentages. Therefore, the nonresponse adjustment appears to have been effective in reducing the bias owing to background questionnaire nonresponse, to the extent that table 11-14 variables are related to literacy.

### 11.1.5 Conclusion

The household sample was subject to unit nonresponse from the screener, background questionnaire, assessment, and oral module and to item nonresponse to background questionnaire items. While all background questionnaire items had response rates of more than 85 percent, two stages of data collection—the screener and the background questionnaire—had unit response rates below 85 percent and thus required an analysis of the potential for nonresponse bias.

In bivariate unit-level analyses at the screener and background questionnaire stages, estimated percentages for respondents were compared with those for the total eligible sample to identify any potential bias owing to nonresponse. Although some statistically significant differences exist, the potential for bias is small because the absolute difference between estimated percentages is less than 2 percent for all domains considered. Multivariate analyses were conducted to further explore the potential for nonresponse bias by identifying the domains with the most differential response rates. These analyses revealed that the lowest response rates for the screener were among dwelling units in segments with high median income, small average household size, and a large proportion of renters. The lowest response rates for the background questionnaire were among males aged 30 or older in segments with high median income. However, the variables used to define these areas and other pockets with low response rates were used in weighting adjustments. The analysis showed that weighting adjustments were highly effective in reducing the bias. The general conclusion is that the potential amount of nonresponse bias attributable to unit nonresponse at the screener and background questionnaire stages is likely to be negligible.

**Table 11-19. Household Study background questionnaire weighting effects for the national NAAL household sample, by subgroup: 2003**

| Subgroup | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio | | Bias | | | Bias ratio |
| | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | *p* value | (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | *p* value | (4)-(1) / (2) |
| Total | 100.0 | † | 100.0 | † | † | † | † | 100.0 | † | † | † | † |
| Region | | | | | | | | | | | | |
| Northeast | 19.7 | 0.64 | 19.7 | # | 0.23 | 0.918 | 0.03 | 19.7 | 0.1 | 0.23 | 0.791 | 0.09 |
| Midwest | 23.2 | 0.89 | 24.0 | 0.7 | 0.21 | 0.001* | 0.83 | 24.3 | 1.0 | 0.21 | 0.000* | 1.13 |
| South | 35.8 | 0.88 | 35.3 | –0.5 | 0.34 | 0.173 | –0.54 | 35.3 | –0.6 | 0.37 | 0.139 | –0.63 |
| West | 21.3 | 0.66 | 21.0 | –0.3 | 0.26 | 0.256 | –0.45 | 20.8 | –0.5 | 0.25 | 0.036 | –0.80 |
| MSA[2] status | | | | | | | | | | | | |
| Non-MSA | 20.2 | 0.67 | 21.1 | 0.9 | 0.28 | 0.002* | 1.38 | 20.9 | 0.7 | 0.28 | 0.013* | 1.05 |
| MSA | 79.8 | 0.67 | 78.9 | –0.9 | 0.28 | 0.002* | –1.38 | 79.1 | –0.7 | 0.28 | 0.013* | –1.05 |
| Average household size | | | | | | | | | | | | |
| 2.42 or less | 29.6 | 1.61 | 30.2 | 0.6 | 0.31 | 0.061 | 0.37 | 30.0 | 0.4 | 0.31 | 0.190 | 0.25 |
| 2.43–2.80 | 36.1 | 1.60 | 36.0 | –0.1 | 0.29 | 0.703 | –0.07 | 36.1 | # | 0.29 | 0.974 | 0.01 |
| Greater than 2.80 | 34.3 | 2.13 | 33.8 | –0.5 | 0.37 | 0.190 | –0.23 | 33.9 | –0.4 | 0.35 | 0.243 | –0.19 |
| Percent with less than high school education | | | | | | | | | | | | |
| 10.4 or less | 33.0 | 1.76 | 31.9 | –1.0 | 0.34 | 0.003* | –0.59 | 32.9 | # | 0.25 | 0.885 | –0.02 |
| 10.5–20.3 | 28.1 | 1.39 | 27.8 | –0.2 | 0.30 | 0.430 | –0.17 | 28.1 | 0.1 | 0.29 | 0.884 | 0.04 |
| 20.4–32.0 | 21.7 | 1.22 | 22.2 | 0.5 | 0.29 | 0.099 | 0.40 | 21.8 | 0.1 | 0.27 | 0.763 | 0.07 |
| Greater than 32.0 | 17.3 | 1.38 | 18.1 | 0.8 | 0.25 | 0.003* | 0.57 | 17.2 | –0.1 | 0.18 | 0.622 | –0.07 |
| Percent speaking Spanish but not English | | | | | | | | | | | | |
| 0 | 42.7 | 2.15 | 42.8 | # | 0.32 | 0.977 | 0.01 | 43.1 | 0.3 | 0.32 | 0.331 | 0.14 |
| 1–28 | 36.6 | 1.63 | 36.0 | –0.6 | 0.29 | 0.046 | –0.37 | 36.3 | –0.3 | 0.30 | 0.347 | –0.17 |
| Greater than 28 | 20.7 | 1.62 | 21.3 | 0.6 | 0.28 | 0.037 | 0.37 | 20.7 | # | 0.26 | 0.916 | –0.02 |

See notes at end of table.

**Table 11-19. Household Study background questionnaire weighting effects for the national NAAL household sample, by subgroup: 2003—Continued**

| | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio | | Bias | | | Bias ratio |
| Subgroup | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | *p* value | (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | *p* value | (4)-(1) / (2) |
| Percent below 150 percent of poverty | | | | | | | | | | | | |
| 10.7 or less | 36.1 | 1.46 | 34.4 | –1.7 | 0.29 | 0.000* | –1.16 | 36.1 | # | 0.00 | 0.694 | 0.00 |
| 10.8–20.0 | 26.7 | 1.57 | 26.8 | 0.1 | 0.25 | 0.793 | 0.05 | 26.5 | –0.2 | 0.21 | 0.336 | –0.13 |
| 20.1–33.3 | 21.6 | 1.56 | 22.1 | 0.5 | 0.26 | 0.049 | 0.33 | 21.9 | 0.2 | 0.21 | 0.336 | 0.14 |
| Greater than 33.3 | 15.6 | 1.11 | 16.7 | 1.1 | 0.16 | 0.000* | 1.00 | 15.6 | # | 0.00 | 0.882 | 0.00 |
| Median income (in dollars) | | | | | | | | | | | | |
| 28,400 or less | 16.7 | 1.30 | 18.0 | 1.3 | 0.22 | 0.000* | 1.01 | 17.1 | 0.4 | 0.14 | 0.005* | 0.32 |
| 28,401–37,850 | 22.8 | 1.30 | 23.5 | 0.7 | 0.25 | 0.008* | 0.52 | 23.0 | 0.2 | 0.22 | 0.403 | 0.15 |
| 37,851–52,100 | 25.6 | 1.55 | 25.5 | –0.1 | 0.31 | 0.740 | –0.07 | 25.5 | –0.1 | 0.29 | 0.709 | –0.07 |
| Greater than 52,100 | 34.9 | 1.73 | 33.0 | –1.9 | 0.35 | 0.000* | –1.10 | 34.4 | –0.5 | 0.23 | 0.035 | –0.28 |
| Percent who rent | | | | | | | | | | | | |
| 16 or less | 35.3 | 1.75 | 33.9 | –1.4 | 0.35 | 0.000* | –0.80 | 34.9 | –0.4 | 0.30 | 0.188 | –0.23 |
| 17–31 | 27.4 | 1.58 | 27.6 | 0.2 | 0.31 | 0.465 | 0.15 | 27.7 | 0.3 | 0.30 | 0.309 | 0.20 |
| 32–59 | 21.8 | 1.37 | 22.2 | 0.4 | 0.24 | 0.103 | 0.29 | 21.8 | # | 0.24 | 0.971 | 0.01 |
| Greater than 59 | 15.6 | 0.80 | 16.3 | 0.8 | 0.17 | 0.000* | 0.95 | 15.6 | 0.1 | 0.17 | 0.609 | 0.10 |
| Age (years) | | | | | | | | | | | | |
| 16–29 | 24.3 | 0.42 | 25.9 | 1.6 | 0.25 | 0.000* | 3.87 | 24.3 | # | 0.03 | 0.768 | –0.02 |
| 30–49 | 40.2 | 0.61 | 39.3 | –0.9 | 0.33 | 0.009* | –1.47 | 40.3 | 0.1 | 0.05 | 0.259 | 0.10 |
| 50–69 | 25.4 | 0.48 | 24.9 | –0.5 | 0.26 | 0.055 | –1.06 | 25.4 | # | 0.02 | 0.663 | –0.02 |
| 70+ | 10.1 | 0.37 | 9.9 | –0.2 | 0.17 | 0.273 | –0.52 | 10.0 | # | 0.04 | 0.357 | –0.11 |

See notes at end of table.

**Table 11-19. Household Study background questionnaire weighting effects for the national NAAL household sample, by subgroup: 2003—Continued**

| Subgroup | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio | | Bias | | | Bias ratio |
| | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | *p* value | (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | *p* value | (4)-(1) / (2) |
| Race/ethnicity | | | | | | | | | | | | |
| Hispanic | 12.6 | 1.18 | 13.2 | 0.7 | 0.22 | 0.004* | 0.56 | 12.9 | 0.3 | 0.19 | 0.137 | 0.25 |
| Non-Hispanic Black only | 10.3 | 0.83 | 10.8 | 0.5 | 0.18 | 0.009* | 0.58 | 10.5 | 0.2 | 0.16 | 0.277 | 0.22 |
| Other[3] | 77.1 | 1.34 | 76.0 | –1.1 | 0.29 | 0.000* | –0.85 | 76.6 | –0.5 | 0.24 | 0.059 | –0.35 |
| Gender | | | | | | | | | | | | |
| Male | 48.2 | 0.41 | 46.5 | –1.7 | 0.32 | 0.000* | –4.25 | 48.2 | # | 0.05 | 0.924 | 0.03 |
| Female | 51.8 | 0.41 | 53.6 | 1.7 | 0.32 | 0.000* | 4.25 | 51.8 | # | 0.05 | 0.924 | –0.03 |

† Not applicable.
# Rounds to zero.
* Statistically significant with Bonferroni adjustment at $\alpha = 0.05$.
[1] Standard error.
[2] Metropolitan Statistical Area.
[3] Includes non-Hispanic White, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races.
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

## Table 11-20. Household Study background questionnaire weighting effects for Kentucky, by subgroup: 2003

| Subgroup | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio | | Bias | | | Bias ratio |
| | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | *p* value | (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | *p* value | (4)-(1) / (2) |
| Total | 100.0 | † | 100.0 | † | † | † | † | 100.0 | † | † | † | † |
| MSA[2] status | | | | | | | | | | | | |
| Non-MSA | 54.0 | 1.30 | 54.0 | # | 0.76 | 0.991 | –0.01 | 53.2 | –0.8 | 0.58 | 0.182 | –0.62 |
| MSA | 46.0 | 1.30 | 46.0 | # | 0.76 | 0.991 | 0.01 | 46.8 | 0.8 | 0.58 | 0.182 | 0.62 |
| Average household size | | | | | | | | | | | | |
| 2.42 or less | 32.1 | 3.74 | 33.0 | 0.9 | 0.53 | 0.117 | 0.24 | 32.8 | 0.7 | 0.60 | 0.273 | 0.18 |
| 2.43–2.80 | 54.5 | 4.19 | 53.4 | –1.1 | 0.55 | 0.057 | –0.27 | 53.6 | –0.9 | 0.59 | 0.131 | –0.22 |
| Greater than 2.80 | 13.4 | 3.67 | 13.6 | 0.2 | 0.37 | 0.533 | 0.07 | 13.7 | 0.3 | 0.33 | 0.452 | 0.07 |
| Percent with less than high school education | | | | | | | | | | | | |
| 10.4 or less | 16.8 | 3.14 | 16.0 | –0.8 | 0.76 | 0.316 | –0.25 | 16.8 | # | 0.00 | 0.890 | 0.00 |
| 10.5–20.3 | 21.1 | 3.89 | 20.9 | –0.1 | 0.58 | 0.859 | –0.03 | 21.1 | # | 0.00 | 0.910 | 0.00 |
| 20.4–32.0 | 36.6 | 4.69 | 35.4 | –1.2 | 0.78 | 0.136 | –0.26 | 36.6 | # | 0.00 | 0.794 | 0.00 |
| Greater than 32.0 | 25.6 | 3.22 | 27.7 | 2.1 | 0.58 | 0.002* | 0.65 | 25.6 | # | 0.00 | 0.745 | 0.00 |
| Percent speaking Spanish but not English | | | | | | | | | | | | |
| 0 | 62.7 | 4.17 | 62.9 | 0.3 | 1.02 | 0.808 | 0.06 | 63.0 | 0.3 | 0.92 | 0.761 | 0.07 |
| 1–28 | 31.7 | 3.42 | 31.4 | –0.4 | 0.92 | 0.712 | –0.10 | 31.4 | –0.3 | 0.80 | 0.722 | –0.08 |
| Greater than 28 | 5.6 | 1.60 | 5.7 | 0.1 | 0.30 | 0.760 | 0.06 | 5.6 | # | 0.31 | 0.986 | 0.00 |
| Percent below 150 percent of poverty | | | | | | | | | | | | |
| 10.7 or less | 15.3 | 3.07 | 14.5 | –0.8 | 0.68 | 0.259 | –0.26 | 15.0 | –0.3 | 0.62 | 0.659 | –0.09 |
| 10.8–20.0 | 28.7 | 3.24 | 27.7 | –0.9 | 0.80 | 0.264 | –0.28 | 28.6 | –0.1 | 0.78 | 0.929 | –0.02 |
| 20.1–33.3 | 25.7 | 4.54 | 25.9 | 0.2 | 0.79 | 0.857 | 0.03 | 26.1 | 0.4 | 0.68 | 0.609 | 0.08 |
| Greater than 33.3 | 30.3 | 4.64 | 31.9 | 1.6 | 0.73 | 0.047 | 0.34 | 30.3 | # | 0.45 | 0.987 | 0.00 |

See notes at end of table.

## Table 11-20. Household Study background questionnaire weighting effects for Kentucky, by subgroup: 2003—Continued

| Subgroup | Base weight—eligible: Percent (1) | Base weight—eligible: SE[1] (2) | Base weight—respondents: Percent (3) | Base weight—respondents: Bias: Estimate (3)-(1) | Base weight—respondents: Bias: SE[1] | Base weight—respondents: Bias: p value | Base weight—respondents: Bias ratio (3)-(1) / (2) | Nonresponse-adjusted weight—respondents: Percent (4) | Nonresponse-adjusted weight—respondents: Bias: Estimate (4)-(1) | Nonresponse-adjusted weight—respondents: Bias: SE[1] | Nonresponse-adjusted weight—respondents: Bias: p value | Nonresponse-adjusted weight—respondents: Bias ratio (4)-(1) / (2) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Median income (in dollars) | | | | | | | | | | | | |
| 28,400 or less | 28.9 | 4.31 | 30.4 | 1.5 | 0.83 | 0.08 | 0.35 | 28.9 | # | 0.62 | 0.932 | –0.01 |
| 28,401–37,850 | 27.4 | 4.42 | 26.3 | –1.1 | 0.95 | 0.26 | –0.24 | 26.5 | –0.9 | 0.88 | 0.305 | –0.21 |
| 37,851–52,100 | 29.3 | 2.74 | 29.5 | 0.3 | 0.61 | 0.69 | 0.09 | 30.5 | 1.2 | 0.68 | 0.089 | 0.45 |
| Greater than 52,100 | 14.4 | 2.52 | 13.7 | –0.7 | 0.62 | 0.29 | –0.26 | 14.1 | –0.2 | 0.41 | 0.562 | –0.10 |
| Percent who rent | | | | | | | | | | | | |
| 16 or less | 30.1 | 3.00 | 28.4 | –1.7 | 1.06 | 0.13 | –0.56 | 29.0 | –1.2 | 1.02 | 0.272 | –0.38 |
| 17–31 | 40.2 | 3.43 | 40.5 | 0.3 | 0.94 | 0.78 | 0.08 | 40.1 | –0.1 | 0.88 | 0.916 | –0.03 |
| 32–59 | 20.8 | 2.34 | 22.1 | 1.3 | 0.42 | 0.005 | 0.56 | 22.0 | 1.2 | 0.42 | 0.010* | 0.51 |
| Greater than 59 | 8.9 | 1.72 | 9.0 | 0.1 | 0.37 | 0.78 | 0.06 | 8.9 | # | 0.35 | 0.910 | 0.02 |
| Age (years) | | | | | | | | | | | | |
| 16–29 | 24.2 | 0.99 | 25.7 | 1.5 | 0.42 | 0.002 | 1.53 | 24.2 | # | 0.00 | 0.909 | 0.00 |
| 30–49 | 38.8 | 1.64 | 38.8 | # | 0.63 | 0.99 | 0.01 | 38.8 | # | 0.00 | 0.928 | 0.00 |
| 50–69 | 27.5 | 1.28 | 26.5 | –1.0 | 0.67 | 0.16 | –0.76 | 27.2 | –0.3 | 0.26 | 0.347 | –0.19 |
| 70+ | 9.6 | 0.70 | 9.0 | –0.5 | 0.45 | 0.24 | –0.77 | 9.8 | 0.3 | 0.26 | 0.347 | 0.36 |
| Race/ethnicity | | | | | | | | | | | | |
| Hispanic | 1.9 | 0.53 | 2.1 | 0.2 | 0.10 | 0.14 | 0.29 | 2.0 | # | 0.15 | 0.703 | 0.10 |
| Non-Hispanic Black only | 7.4 | 1.49 | 7.3 | # | 0.30 | 0.91 | –0.03 | 7.2 | –0.1 | 0.35 | 0.709 | –0.09 |
| Other[3] | 90.7 | 1.87 | 90.6 | –0.1 | 0.33 | 0.71 | –0.06 | 90.8 | 0.1 | 0.37 | 0.843 | 0.04 |
| Gender | | | | | | | | | | | | |
| Male | 43.9 | 1.57 | 42.2 | –1.7 | 0.58 | 0.007 | -1.10 | 43.3 | –0.7 | 0.37 | 0.096 | –0.41 |
| Female | 56.1 | 1.57 | 57.8 | 1.7 | 0.58 | 0.007 | 1.10 | 56.7 | 0.7 | 0.37 | 0.096 | 0.41 |

† Not applicable.
# Rounds to zero.
* Statistically significant with Bonferroni adjustment at $\alpha = 0.05$.
[1] Standard error.
[2] Metropolitan Statistical Area.
[3] Includes non-Hispanic White, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races.
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 11-21. Household Study background questionnaire weighting effects for Maryland, by subgroup: 2003**

| Subgroup | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio | | Bias | | | Bias ratio |
| | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | p value | (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | p value | (4)-(1) / (2) |
| Total | 100.0 | † | 100.0 | † | † | † | † | 100.0 | † | † | † | † |
| MSA[2] status | | | | | | | | | | | | |
| Non-MSA | 7.1 | 1.21 | 5.6 | –1.5 | 0.63 | 0.032 | –1.21 | 7.1 | # | 0.00 | 0.942 | 0.00 |
| MSA | 92.9 | 1.21 | 94.4 | 1.5 | 0.63 | 0.032 | 1.21 | 92.9 | # | 0.00 | 1.000 | 0.00 |
| Average household size | | | | | | | | | | | | |
| 2.42 or less | 34.3 | 3.99 | 34.8 | 0.5 | 0.93 | 0.574 | 0.13 | 35.3 | 1.0 | 1.07 | 0.354 | 0.25 |
| 2.43–2.80 | 30.8 | 3.72 | 29.6 | –1.1 | 0.92 | 0.240 | –0.30 | 29.9 | –0.9 | 1.06 | 0.422 | –0.23 |
| Greater than 2.80 | 34.9 | 4.82 | 35.5 | 0.6 | 0.59 | 0.337 | 0.12 | 34.8 | –0.2 | 0.66 | 0.823 | –0.03 |
| Percent with less than high school education | | | | | | | | | | | | |
| 10.4 or less | 43.7 | 3.68 | 43.2 | –0.5 | 0.77 | 0.519 | –0.14 | 42.5 | –1.2 | 0.57 | 0.046 | –0.33 |
| 10.5–20.3 | 30.1 | 4.47 | 29.6 | –0.5 | 0.96 | 0.636 | –0.11 | 29.4 | –0.6 | 1.04 | 0.553 | –0.14 |
| 20.4–32.0 | 19.7 | 4.11 | 20.2 | 0.5 | 0.54 | 0.335 | 0.13 | 20.9 | 1.2 | 0.73 | 0.106 | 0.30 |
| Greater than 32.0 | 6.5 | 1.34 | 7.0 | 0.4 | 0.69 | 0.532 | 0.33 | 7.2 | 0.6 | 0.65 | 0.360 | 0.45 |
| Percent speaking Spanish but not English | | | | | | | | | | | | |
| 0 | 56.6 | 5.05 | 57.4 | 0.8 | 0.79 | 0.309 | 0.16 | 57.5 | 0.9 | 0.89 | 0.315 | 0.18 |
| 1–28 | 35.4 | 4.84 | 33.9 | –1.5 | 0.70 | 0.045 | –0.31 | 33.8 | –1.5 | 0.76 | 0.057 | –0.32 |
| Greater than 28 | 8.1 | 0.96 | 8.7 | 0.7 | 0.36 | 0.073 | 0.70 | 8.7 | 0.6 | 0.41 | 0.149 | 0.65 |
| Percent below 150 percent of poverty | | | | | | | | | | | | |
| 10.7 or less | 59.7 | 3.94 | 59.9 | 0.1 | 0.89 | 0.864 | 0.04 | 58.8 | –0.9 | 0.92 | 0.341 | –0.23 |
| 10.8–20.0 | 25.5 | 4.34 | 24.2 | –1.3 | 0.73 | 0.089 | –0.30 | 24.6 | –0.9 | 0.85 | 0.307 | –0.21 |
| 20.1–33.3 | 10.2 | 1.76 | 11.1 | 0.8 | 0.39 | 0.043 | 0.48 | 11.5 | 1.3 | 0.35 | 0.002 | 0.74 |
| Greater than 33.3 | 4.6 | 1.32 | 4.9 | 0.3 | 0.54 | 0.561 | 0.24 | 5.1 | 0.5 | 0.48 | 0.311 | 0.38 |

See notes at end of table.

## Table 11-21. Household Study background questionnaire weighting effects for Maryland, by subgroup: 2003—Continued

| Subgroup | Base weight—eligible: Percent (1) | Base weight—eligible: SE[1] (2) | Base weight—respondents: Percent (3) | Base weight—respondents: Bias: Estimate (3)-(1) | Base weight—respondents: Bias: SE[1] | Base weight—respondents: Bias: $p$ value | Base weight—respondents: Bias ratio (3)-(1) / (2) | Nonresponse-adjusted weight—respondents: Percent (4) | Nonresponse-adjusted weight—respondents: Bias: Estimate (4)-(1) | Nonresponse-adjusted weight—respondents: Bias: SE[1] | Nonresponse-adjusted weight—respondents: Bias: $p$ value | Nonresponse-adjusted weight—respondents: Bias ratio (4)-(1) / (2) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Median income (in dollars) | | | | | | | | | | | | |
| 28,400 or less | 7.0 | 1.48 | 8.1 | 1.2 | 0.49 | 0.024 | 0.80 | 8.3 | 1.4 | 0.38 | 0.002* | 0.91 |
| 28,401–37,850 | 12.2 | 3.55 | 12.0 | –0.1 | 0.63 | 0.856 | –0.03 | 13.1 | 1.0 | 0.40 | 0.025 | 0.27 |
| 37,851–52,100 | 22.5 | 3.55 | 21.4 | –1.0 | 0.98 | 0.309 | –0.29 | 21.3 | –1.2 | 0.82 | 0.169 | –0.33 |
| Greater than 52,100 | 58.4 | 3.33 | 58.4 | # | 0.98 | 0.967 | –0.01 | 57.3 | –1.1 | 0.75 | 0.148 | –0.34 |
| Percent who rent | | | | | | | | | | | | |
| 16 or less | 38.4 | 4.57 | 37.3 | –1.0 | 1.17 | 0.399 | –0.22 | 37.1 | –1.2 | 1.03 | 0.248 | –0.27 |
| 17–31 | 25.6 | 3.49 | 25.8 | 0.2 | 0.74 | 0.818 | 0.05 | 25.7 | 0.1 | 0.65 | 0.821 | 0.04 |
| 32–59 | 23.3 | 4.66 | 23.6 | 0.3 | 0.61 | 0.662 | 0.06 | 23.7 | 0.4 | 0.63 | 0.553 | 0.08 |
| Greater than 59 | 12.7 | 3.53 | 13.3 | 0.6 | 0.62 | 0.374 | 0.16 | 13.4 | 0.7 | 0.63 | 0.287 | 0.20 |
| Age (years) | | | | | | | | | | | | |
| 16–29 | 24.8 | 1.27 | 26.2 | 1.4 | 1.09 | 0.206 | 1.13 | 24.8 | 0.1 | 0.22 | 0.690 | 0.07 |
| 30–49 | 40.9 | 1.66 | 38.0 | –3.0 | 1.00 | 0.008* | –1.79 | 41.3 | 0.3 | 0.18 | 0.086 | 0.20 |
| 50–69 | 25.5 | 1.94 | 26.4 | 0.9 | 0.59 | 0.166 | 0.44 | 25.0 | –0.5 | 0.17 | 0.013 | –0.24 |
| 70+ | 8.8 | 1.38 | 9.5 | 0.7 | 0.47 | 0.170 | 0.48 | 8.9 | # | 0.14 | 0.776 | 0.03 |
| Race/ethnicity | | | | | | | | | | | | |
| Hispanic | 4.7 | 0.47 | 5.0 | 0.4 | 0.35 | 0.313 | 0.79 | 5.0 | 0.3 | 0.38 | 0.429 | 0.66 |
| Non-Hispanic Black only | 26.3 | 1.76 | 27.0 | 0.7 | 0.95 | 0.475 | 0.40 | 26.7 | 0.4 | 1.06 | 0.712 | 0.23 |
| Other[3] | 69.0 | 1.63 | 68.0 | –1.1 | 0.86 | 0.234 | –0.65 | 68.3 | –0.7 | 0.93 | 0.458 | –0.43 |
| Gender | | | | | | | | | | | | |
| Male | 44.5 | 1.24 | 43.4 | –1.1 | 1.02 | 0.300 | –0.87 | 43.3 | –1.2 | 0.96 | 0.230 | –0.96 |
| Female | 55.5 | 1.24 | 56.6 | 1.1 | 1.02 | 0.300 | 0.87 | 56.7 | 1.2 | 0.96 | 0.230 | 0.96 |

† Not applicable.
# Rounds to zero.
* Statistically significant with Bonferroni adjustment at α = 0.05.
[1] Standard error.
[2] Metropolitan Statistical Area.
[3] Includes non-Hispanic White, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races.
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 11-22. Household Study background questionnaire weighting effects for Massachusetts, by subgroup: 2003**

| Subgroup | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio | | Bias | | | Bias ratio |
| | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | *p* value | (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | *p* value | (4)-(1) / (2) |
| Total | 100.0 | † | 100.0 | † | † | † | † | 100.0 | † | † | † | † |
| MSA[2] status | | | | | | | | | | | | |
| Non-MSA | 1.4 | 0.64 | 1.2 | –0.2 | 0.06 | 0.008 | –0.30 | 1.1 | –0.2 | 0.21 | 0.29 | –0.36 |
| MSA | 98.6 | 0.64 | 98.8 | 0.2 | 0.06 | 0.008 | 0.30 | 98.9 | 0.2 | 0.21 | 0.29 | 0.36 |
| Average household size | | | | | | | | | | | | |
| 2.42 or less | 34.1 | 4.41 | 33.8 | –0.3 | 1.15 | 0.80 | –0.06 | 34.6 | 0.6 | 1.19 | 0.65 | 0.12 |
| 2.43–2.80 | 34.2 | 2.98 | 32.7 | –1.4 | 1.37 | 0.30 | –0.48 | 33.3 | –0.9 | 1.19 | 0.46 | –0.30 |
| Greater than 2.80 | 31.7 | 4.50 | 33.5 | 1.7 | 0.86 | 0.06 | 0.38 | 32.1 | 0.3 | 0.56 | 0.55 | 0.07 |
| Percent with less than high school education | | | | | | | | | | | | |
| 10.4 or less | 46.5 | 6.33 | 48.9 | 2.4 | 1.02 | 0.02 | 0.38 | 47.9 | 1.4 | 0.47 | 0.007 | 0.23 |
| 10.5–20.3 | 32.9 | 6.01 | 31.9 | –1.0 | 0.87 | 0.24 | –0.17 | 33.0 | 0.1 | 0.61 | 0.90 | 0.01 |
| 20.4–32.0 | 9.5 | 2.44 | 7.6 | –2.0 | 0.64 | 0.006 | –0.80 | 8.1 | –1.4 | 0.31 | 0.000 | –0.57 |
| Greater than 32.0 | 11.1 | 1.74 | 11.6 | 0.6 | 0.67 | 0.40 | 0.33 | 11.0 | –0.1 | 0.26 | 0.67 | –0.06 |
| Percent speaking Spanish but not English | | | | | | | | | | | | |
| 0 | 60.3 | 2.71 | 60.6 | 0.3 | 1.06 | 0.78 | 0.11 | 60.7 | 0.5 | 1.05 | 0.67 | 0.17 |
| 1–28 | 31.9 | 2.83 | 30.6 | –1.3 | 1.11 | 0.26 | –0.45 | 31.0 | –1.0 | 1.13 | 0.41 | –0.34 |
| Greater than 28 | 7.8 | 1.68 | 8.8 | 1.0 | 0.83 | 0.24 | 0.59 | 8.3 | 0.5 | 0.58 | 0.39 | 0.30 |
| Percent below 150 percent of poverty | | | | | | | | | | | | |
| 10.7 or less | 55.9 | 5.32 | 56.6 | 0.7 | 1.09 | 0.51 | 0.14 | 55.7 | –0.2 | 0.58 | 0.67 | –0.05 |
| 10.8–20.0 | 18.1 | 5.04 | 16.6 | –1.6 | 1.18 | 0.20 | –0.31 | 17.5 | –0.6 | 0.97 | 0.51 | –0.13 |
| 20.1–33.3 | 17.2 | 3.07 | 16.7 | –0.5 | 0.86 | 0.58 | –0.15 | 17.6 | 0.5 | 0.99 | 0.64 | 0.15 |
| Greater than 33.3 | 8.8 | 1.69 | 10.1 | 1.3 | 1.03 | 0.21 | 0.78 | 9.2 | 0.4 | 0.69 | 0.54 | 0.25 |

See notes at end of table.

**Table 11-22. Household Study background questionnaire weighting effects for Massachusetts, by subgroup: 2003—Continued**

| Subgroup | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio | | Bias | | | Bias ratio |
| | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | *p* value | (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | *p* value | (4)-(1) / (2) |
| Median income (in dollars) | | | | | | | | | | | | |
| 28,400 or less | 7.8 | 2.32 | 8.7 | 0.8 | 0.56 | 0.15( | 0.36 | 7.7 | –0.1 | 0.03 | 0.002* | –0.05 |
| 28,401–37,850 | 11.5 | 1.62 | 11.7 | 0.2 | 0.39 | 0.60: | 0.12 | 11.6 | 0.1 | 0.03 | 0.002* | 0.07 |
| 37,851–52,100 | 31.2 | 3.87 | 28.0 | –3.2 | 0.94 | 0.003' | –0.83 | 31.2 | # | 0.00 | 0.988 | 0.00 |
| Greater than 52,100 | 49.5 | 3.97 | 51.6 | 2.2 | 1.13 | 0.07: | 0.54 | 49.5 | # | 0.00 | 0.936 | 0.00 |
| Percent who rent | | | | | | | | | | | | |
| 16 or less | 34.8 | 3.93 | 34.8 | # | 0.97 | 0.96: | –0.01 | 33.9 | –1.0 | 0.97 | 0.341 | –0.24 |
| 17–31 | 19.0 | 2.17 | 19.9 | 0.9 | 1.06 | 0.40( | 0.42 | 20.0 | 1.0 | 0.98 | 0.338 | 0.45 |
| 32–59 | 21.3 | 3.55 | 19.6 | –1.7 | 0.88 | 0.07( | –0.48 | 20.7 | –0.7 | 0.79 | 0.422 | –0.18 |
| Greater than 59 | 24.9 | 2.23 | 25.7 | 0.8 | 0.83 | 0.33: | 0.37 | 25.5 | 0.6 | 0.52 | 0.235 | 0.29 |
| Age (years) | | | | | | | | | | | | |
| 16–29 | 22.5 | 1.61 | 22.6 | 0.1 | 0.75 | 0.90( | 0.06 | 22.8 | 0.3 | 0.70 | 0.675 | 0.19 |
| 30–49 | 41.6 | 1.92 | 41.6 | # | 0.83 | 0.97( | –0.02 | 41.8 | 0.2 | 0.90 | 0.832 | 0.10 |
| 50–69 | 25.5 | 2.20 | 24.9 | –0.6 | 0.94 | 0.51· | –0.29 | 24.6 | –0.9 | 0.98 | 0.350 | –0.43 |
| 70+ | 10.4 | 0.80 | 11.0 | 0.6 | 0.45 | 0.21' | 0.71 | 10.8 | 0.4 | 0.49 | 0.381 | 0.55 |
| Race/ethnicity | | | | | | | | | | | | |
| Hispanic | 7.8 | 1.21 | 8.6 | 0.8 | 0.46 | 0.10· | 0.65 | 8.0 | 0.3 | 0.47 | 0.588 | 0.22 |
| Non-Hispanic Black only | 5.7 | 0.85 | 5.9 | 0.3 | 0.30 | 0.40' | 0.29 | 6.1 | 0.4 | 0.28 | 0.154 | 0.49 |
| Other[3] | 86.6 | 1.48 | 85.5 | –1.0 | 0.48 | 0.04( | –0.71 | 85.9 | –0.7 | 0.52 | 0.210 | –0.46 |
| Gender | | | | | | | | | | | | |
| Male | 46.6 | 1.41 | 44.3 | –2.3 | 0.61 | 0.001' | –1.60 | 44.4 | –2.2 | 0.71 | 0.006* | –1.56 |
| Female | 53.4 | 1.41 | 55.7 | 2.3 | 0.61 | 0.001' | 1.60 | 55.6 | 2.2 | 0.71 | 0.006* | 1.56 |

† Not applicable.
# Rounds to zero.
* Statistically significant with Bonferroni adjustment at α = 0.05.
[1] Standard error.
[2] Metropolitan Statistical Area.
[3] Includes non-Hispanic White, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races.
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 11-23. Household Study background questionnaire weighting effects for Missouri, by subgroup: 2003**

| Subgroup | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio | | Bias | | | Bias ratio |
| | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | *p* value | (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | *p* value | (4)-(1) / (2) |
| Total | 100.0 | † | 100.0 | † | † | † | † | 100.0 | † | † | † | † |
| MSA[2] status | | | | | | | | | | | | |
| Non-MSA | 32.9 | 2.46 | 33.4 | 0.5 | 1.32 | 0.711 | 0.20 | 33.2 | 0.2 | 1.24 | 0.858 | 0.09 |
| MSA | 67.1 | 2.46 | 66.6 | –0.5 | 1.32 | 0.711 | –0.20 | 66.8 | –0.2 | 1.24 | 0.858 | –0.09 |
| Average household size | | | | | | | | | | | | |
| 2.42 or less | 37.4 | 3.52 | 38.3 | 1.0 | 1.20 | 0.432 | 0.27 | 38.0 | 0.6 | 1.03 | 0.577 | 0.17 |
| 2.43–2.80 | 46.5 | 4.22 | 46.2 | –0.3 | 1.27 | 0.808 | –0.07 | 46.5 | # | 1.10 | 0.993 | 0.00 |
| Greater than 2.80 | 16.1 | 3.31 | 15.5 | –0.7 | 0.92 | 0.487 | –0.20 | 15.5 | –0.6 | 0.71 | 0.428 | –0.18 |
| Percent with less than high school education | | | | | | | | | | | | |
| 10.4 or less | 29.1 | 6.06 | 28.1 | –1.0 | 1.11 | 0.383 | –0.16 | 28.5 | –0.6 | 0.85 | 0.516 | –0.09 |
| 10.5–20.3 | 33.5 | 3.95 | 33.9 | 0.4 | 0.91 | 0.664 | 0.10 | 34.3 | 0.8 | 0.82 | 0.334 | 0.21 |
| 20.4–32.0 | 26.0 | 3.75 | 27.0 | 1.0 | 0.88 | 0.269 | 0.27 | 26.5 | 0.5 | 0.74 | 0.510 | 0.13 |
| Greater than 32.0 | 11.5 | 3.52 | 11.1 | –0.4 | 0.67 | 0.543 | –0.12 | 10.7 | –0.8 | 0.49 | 0.142 | –0.21 |
| Percent speaking Spanish but not English | | | | | | | | | | | | |
| 0 | 56.5 | 5.53 | 57.0 | 0.5 | 1.06 | 0.632 | 0.09 | 56.7 | 0.2 | 0.96 | 0.814 | 0.04 |
| 1–28 | 39.6 | 5.42 | 39.7 | 0.1 | 1.11 | 0.956 | 0.01 | 40.0 | 0.4 | 1.01 | 0.694 | 0.07 |
| Greater than 28 | 4.0 | 2.32 | 3.4 | –0.6 | 0.42 | 0.181 | –0.25 | 3.3 | –0.6 | 0.43 | 0.162 | –0.27 |
| Percent below 150 percent of poverty | | | | | | | | | | | | |
| 10.7 or less | 29.1 | 4.92 | 28.9 | –0.2 | 1.06 | 0.835 | –0.04 | 29.4 | 0.3 | 0.66 | 0.689 | 0.05 |
| 10.8–20.0 | 26.0 | 3.96 | 25.1 | –1.0 | 1.30 | 0.473 | –0.24 | 25.5 | –0.5 | 1.04 | 0.664 | –0.12 |
| 20.1–33.3 | 29.8 | 5.05 | 30.8 | 1.0 | 0.73 | 0.191 | 0.20 | 30.4 | 0.6 | 0.58 | 0.326 | 0.11 |
| Greater than 33.3 | 15.0 | 2.69 | 15.2 | 0.2 | 0.71 | 0.797 | 0.07 | 14.7 | –0.4 | 0.45 | 0.397 | –0.15 |

See notes at end of table.

## Table 11-23. Household Study background questionnaire weighting effects for Missouri, by subgroup: 2003—Continued

| Subgroup | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio | | Bias | | | Bias ratio |
| | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | *p* value | (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | *p* value | (4)-(1) / (2) |
| Median income (in dollars) | | | | | | | | | | | | |
| 28,400 or less | 23.6 | 3.43 | 25.1 | 1.5 | 0.93 | 0.130 | 0.43 | 24.3 | 0.6 | 0.36 | 0.094 | 0.18 |
| 28,401–37,850 | 25.4 | 3.32 | 25.1 | –0.3 | 0.57 | 0.621 | –0.08 | 24.8 | –0.6 | 0.45 | 0.205 | –0.18 |
| 37,851–52,100 | 27.2 | 2.78 | 26.4 | –0.8 | 1.27 | 0.558 | –0.27 | 27.2 | # | 0.63 | 0.981 | –0.01 |
| Greater than 52,100 | 23.8 | 5.09 | 23.3 | –0.4 | 1.01 | 0.677 | –0.08 | 23.8 | # | 0.52 | 0.974 | 0.00 |
| Percent who rent | | | | | | | | | | | | |
| 16 or less | 39.7 | 4.50 | 38.2 | –1.5 | 0.92 | 0.119 | –0.33 | 38.8 | –0.9 | 0.88 | 0.312 | –0.20 |
| 17–31 | 25.0 | 4.00 | 25.2 | 0.3 | 0.60 | 0.678 | 0.06 | 25.1 | 0.1 | 0.73 | 0.839 | 0.04 |
| 32–59 | 26.6 | 3.72 | 27.2 | 0.7 | 0.92 | 0.488 | 0.17 | 27.0 | 0.5 | 0.84 | 0.591 | 0.12 |
| Greater than 59 | 8.7 | 1.57 | 9.3 | 0.6 | 0.44 | 0.190 | 0.38 | 9.1 | 0.3 | 0.35 | 0.394 | 0.20 |
| Age (years) | | | | | | | | | | | | |
| 16–29 | 23.6 | 1.27 | 24.6 | 1.0 | 0.73 | 0.193 | 0.78 | 23.6 | # | 0.00 | 0.731 | 0.00 |
| 30–49 | 40.4 | 1.60 | 40.3 | –0.1 | 0.72 | 0.894 | –0.06 | 40.4 | # | 0.00 | 0.665 | 0.00 |
| 50–69 | 25.6 | 1.38 | 24.7 | –0.9 | 0.88 | 0.344 | –0.62 | 25.6 | # | 0.00 | 0.782 | 0.00 |
| 70+ | 10.4 | 1.54 | 10.3 | # | 0.46 | 0.925 | –0.03 | 10.4 | # | 0.00 | 0.653 | 0.00 |
| Race/ethnicity | | | | | | | | | | | | |
| Hispanic | 2.4 | 0.66 | 2.6 | 0.2 | 0.22 | 0.381 | 0.30 | 2.6 | 0.1 | 0.23 | 0.554 | 0.21 |
| Non-Hispanic Black only | 10.1 | 1.10 | 10.4 | 0.3 | 1.24 | 0.841 | 0.23 | 10.1 | # | 1.10 | 0.977 | 0.03 |
| Other[3] | 87.5 | 1.28 | 87.0 | –0.5 | 1.25 | 0.726 | –0.35 | 87.3 | –0.2 | 1.09 | 0.878 | –0.13 |
| Gender | | | | | | | | | | | | |
| Male | 46.5 | 1.88 | 46.0 | –0.4 | 0.72 | 0.557 | –0.23 | 46.5 | # | 0.00 | 0.940 | 0.00 |
| Female | 53.6 | 1.88 | 54.0 | 0.4 | 0.72 | 0.557 | 0.23 | 53.6 | # | 0.00 | 0.942 | 0.00 |

† Not applicable.
# Rounds to zero.
[1] Standard error.
[2] Metropolitan Statistical Area.
[3] Includes non-Hispanic White, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races.
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 11-24. Household Study background questionnaire weighting effects for New York, by subgroup: 2003**

| Subgroup | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio | | Bias | | | Bias ratio |
| | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | *p* value | (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | *p* value | (4)-(1) / (2) |
| Total | 100.0 | † | 100.0 | † | † | † | † | 100.0 | † | † | † | † |
| MSA[2] status | | | | | | | | | | | | |
| Non-MSA | 8.3 | 2.40 | 9.8 | 1.6 | 0.74 | 0.049 | 0.65 | 10.4 | 2.2 | 0.85 | 0.019* | 0.90 |
| MSA | 91.7 | 2.40 | 90.2 | –1.6 | 0.74 | 0.049 | –0.65 | 89.6 | –2.2 | 0.85 | 0.019* | –0.90 |
| Average household size | | | | | | | | | | | | |
| 2.42 or less | 27.2 | 3.86 | 27.6 | 0.5 | 1.19 | 0.708 | 0.12 | 27.6 | 0.4 | 1.22 | 0.758 | 0.10 |
| 2.43–2.80 | 32.2 | 4.88 | 34.3 | 2.1 | 1.60 | 0.206 | 0.43 | 35.2 | 3.0 | 176 | 0.104 | 0.62 |
| Greater than 2.80 | 40.7 | 5.79 | 38.1 | –2.6 | 1.38 | 0.081 | –0.44 | 37.3 | –3.4 | 1.58 | 0.046 | –0.58 |
| Percent with less than high school education | | | | | | | | | | | | |
| 10.4 or less | 37.5 | 2.40 | 36.2 | –1.4 | 1.27 | 0.291 | –0.57 | 37.8 | 0.3 | 1.00 | 0.776 | 0.12 |
| 10.5–20.3 | 21.8 | 3.40 | 20.3 | –1.5 | 0.84 | 0.085 | –0.45 | 21.0 | –0.9 | 0.89 | 0.346 | –0.25 |
| 20.4–32.0 | 21.1 | 2.30 | 21.3 | 0.1 | 1.01 | 0.890 | 0.06 | 21.3 | 0.2 | 1.01 | 0.847 | 0.09 |
| Greater than 32.0 | 19.5 | 2.61 | 22.3 | 2.8 | 0.96 | 0.010* | 1.06 | 19.9 | 0.4 | 0.63 | 0.562 | 0.14 |
| Percent speaking Spanish but not English | | | | | | | | | | | | |
| 0 | 43.0 | 5.52 | 43.3 | 0.3 | 1.17 | 0.802 | 0.05 | 45.1 | 2.1 | 1.38 | 0.153 | 0.37 |
| 1–28 | 31.9 | 5.58 | 29.6 | –2.3 | 1.47 | 0.142 | –0.41 | 29.9 | –1.9 | 1.52 | 0.217 | –0.35 |
| Greater than 28 | 25.1 | 2.90 | 27.1 | 1.9 | 0.80 | 0.025 | 0.67 | 25.0 | –0.1 | 0.72 | 0.887 | –0.03 |
| Percent below 150 percent of poverty | | | | | | | | | | | | |
| 10.7 or less | 41.9 | 4.24 | 40.8 | –1.1 | 1.41 | 0.430 | –0.27 | 41.9 | # | 0.00 | 0.885 | 0.00 |
| 10.8–20.0 | 20.1 | 3.90 | 18.3 | –1.8 | 1.40 | 0.208 | –0.47 | 20.1 | # | 0.00 | 0.806 | 0.00 |
| 20.1–33.3 | 18.6 | 4.45 | 18.6 | 0.1 | 0.87 | 0.942 | 0.02 | 18.6 | # | 0.00 | 0.836 | 0.00 |
| Greater than 33.3 | 19.4 | 3.48 | 22.3 | 2.9 | 0.86 | 0.003* | 0.83 | 19.4 | # | 0.00 | 0.826 | 0.00 |

See notes at end of table.

**Table 11-24. Household Study background questionnaire weighting effects for New York, by subgroup: 2003—Continued**

| Subgroup | Base weight—eligible Percent (1) | Base weight—eligible SE[1] (2) | Base weight—respondents Percent (3) | Bias Estimate (3)-(1) | Bias SE[1] | Bias *p* value | Bias ratio (3)-(1) / (2) | Nonresponse-adjusted weight—respondents Percent (4) | Bias Estimate (4)-(1) | Bias SE[1] | Bias *p* value | Bias ratio (4)-(1) / (2) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Median income (in dollars) | | | | | | | | | | | | |
| 28,400 or less | 12.8 | 2.71 | 15.1 | 2.3 | 0.73 | 0.005* | 0.86 | 13.3 | 0.5 | 0.30 | 0.112 | 0.18 |
| 28,401–37,850 | 23.2 | 3.40 | 24.2 | 1.1 | 0.79 | 0.188 | 0.32 | 23.5 | 0.4 | 0.87 | 0.675 | 0.11 |
| 37,851–52,100 | 18.5 | 3.02 | 19.2 | 0.8 | 0.68 | 0.277 | 0.25 | 20.0 | 1.5 | 0.72 | 0.047 | 0.51 |
| Greater than 52,100 | 45.6 | 3.89 | 41.4 | –4.2 | 1.23 | 0.003* | –1.07 | 43.2 | –2.4 | 1.27 | 0.074 | –0.62 |
| Percent who rent | | | | | | | | | | | | |
| 16 or less | 36.5 | 4.76 | 37.2 | 0.7 | 1.27 | 0.574 | 0.15 | 38.8 | 2.3 | 1.03 | 0.040 | 0.48 |
| 17–31 | 17.4 | 3.07 | 13.4 | –4.0 | 1.45 | 0.013 | –1.30 | 14.2 | –3.2 | 0.32 | 0.026 | –1.04 |
| 32–59 | 13.2 | 2.97 | 13.7 | 0.6 | 0.89 | 0.546 | 0.19 | 13.9 | 0.8 | 0.93 | 0.424 | 0.26 |
| Greater than 59 | 32.9 | 3.27 | 35.7 | 2.7 | 0.93 | 0.009* | 0.83 | 33.1 | 0.2 | 0.79 | 0.851 | 0.05 |
| Age (years) | | | | | | | | | | | | |
| 16–29 | 24.5 | 1.67 | 26.0 | 1.5 | 0.93 | 0.116 | 0.92 | 25.7 | 1.2 | 0.90 | 0.195 | 0.73 |
| 30–49 | 37.2 | 1.73 | 36.5 | –0.7 | 1.50 | 0.665 | –0.38 | 36.4 | –0.8 | 1.32 | 0.571 | –0.44 |
| 50–69 | 27.1 | 2.16 | 26.5 | –0.6 | 1.12 | 0.597 | –0.28 | 26.9 | –0.2 | 1.14 | 0.883 | –0.08 |
| 70+ | 11.3 | 1.08 | 11.1 | –0.3 | 0.50 | 0.582 | –0.26 | 11.1 | –0.3 | 0.48 | 0.576 | –0.26 |
| Race/ethnicity | | | | | | | | | | | | |
| Hispanic | 14.6 | 2.04 | 14.9 | 0.3 | 0.80 | 0.742 | 0.13 | 13.6 | –1.0 | 0.69 | 0.150 | –0.51 |
| Non-Hispanic | | | | | | | | | | | | |
| Black only | 10.0 | 0.76 | 11.0 | 1.0 | 0.35 | 0.012* | 1.28 | 10.6 | 0.6 | 0.42 | 0.189 | 0.75 |
| Other[3] | 75.4 | 2.14 | 74.1 | –1.2 | 1.05 | 0.257 | –0.58 | 75.9 | 0.5 | 0.81 | 0.574 | 0.22 |
| Gender | | | | | | | | | | | | |
| Male | 45.2 | 1.81 | 44.5 | –0.7 | 1.29 | 0.586 | –0.40 | 44.5 | –0.7 | 1.37 | 0.615 | –0.39 |
| Female | 54.8 | 1.81 | 55.6 | 0.7 | 1.29 | 0.586 | 0.40 | 55.5 | 0.7 | 1.37 | 0.615 | 0.39 |

† Not applicable.
# Rounds to zero.
* Statistically significant with Bonferroni adjustment at $\alpha = 0.05$.
[1] Standard error.
[2] Metropolitan Statistical Area.
[3] Includes non-Hispanic White, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races.
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

## Table 11-25. Household Study background questionnaire weighting effects for Oklahoma, by subgroup: 2003

| | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio | | Bias | | | Bias ratio |
| Subgroup | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | p value | (3)-(1) / (2) | Percent (4) | Estimate (4)-(1) | SE[1] | p value | (4)-(1) / (2) |
| Total | 100.0 | † | 100.0 | † | † | † | † | 100.0 | † | † | † | † |
| MSA[2] status | | | | | | | | | | | | |
| Non-MSA | 37.7 | 3.19 | 38.9 | 1.2 | 0.78 | 0.145 | 0.37 | 37.9 | 0.2 | 0.55 | 0.756 | 0.05 |
| MSA | 62.3 | 3.19 | 61.1 | –1.2 | 0.78 | 0.145 | –0.37 | 62.1 | –0.2 | 0.55 | 0.756 | –0.05 |
| Average household size | | | | | | | | | | | | |
| 2.42 or less | 32.9 | 6.45 | 32.9 | 0.1 | 1.03 | 0.936 | 0.01 | 33.1 | 0.2 | 0.80 | 0.788 | 0.03 |
| 2.43–2.80 | 51.2 | 4.44 | 51.7 | 0.6 | 0.85 | 0.521 | 0.13 | 51.2 | # | 0.74 | 0.990 | 0.00 |
| Greater than 2.80 | 16.0 | 4.17 | 15.3 | –0.6 | 0.67 | 0.355 | –0.15 | 15.8 | –0.2 | 0.36 | 0.568 | –0.05 |
| Percent with less than high school education | | | | | | | | | | | | |
| 10.4 or less | 24.6 | 2.21 | 23.7 | –0.9 | 0.52 | 0.117 | –0.39 | 25.1 | 0.5 | 0.40 | 0.200 | 0.24 |
| 10.5–20.3 | 27.1 | 4.52 | 26.3 | –0.8 | 0.38 | 0.041 | –0.18 | 26.7 | –0.4 | 0.43 | 0.403 | –0.08 |
| 20.4–32.0 | 30.2 | 4.96 | 30.6 | 0.4 | 0.49 | 0.458 | 0.07 | 29.8 | –0.4 | 0.46 | 0.403 | –0.08 |
| Greater than 32.0 | 18.1 | 5.06 | 19.4 | 1.3 | 0.46 | 0.010* | 0.26 | 18.4 | 0.2 | 0.28 | 0.418 | 0.05 |
| Percent speaking Spanish but not English | | | | | | | | | | | | |
| 0 | 43.7 | 4.45 | 43.9 | 0.2 | 0.71 | 0.822 | 0.04 | 44.1 | 0.4 | 0.65 | 0.534 | 0.09 |
| 1–28 | 46.0 | 4.86 | 45.4 | –0.6 | 0.70 | 0.442 | –0.11 | 45.5 | –0.5 | 0.64 | 0.453 | –0.10 |
| Greater than 28 | 10.3 | 1.95 | 10.7 | 0.4 | 0.37 | 0.314 | 1.19 | 10.4 | 0.1 | 0.32 | 0.816 | 0.04 |
| Percent below 150 percent of poverty | | | | | | | | | | | | |
| 10.7 or less | 18.8 | 3.68 | 18.1 | –0.7 | 0.81 | 0.432 | –0.18 | 19.4 | 0.6 | 0.67 | 0.354 | 0.17 |
| 10.8–20.0 | 20.3 | 4.06 | 18.8 | –1.5 | 0.84 | 0.101 | –0.36 | 19.6 | –0.7 | 0.70 | 0.350 | –0.17 |
| 20.1–33.3 | 26.6 | 3.89 | 26.9 | 0.3 | 0.70 | 0.695 | 0.07 | 26.4 | –0.2 | 0.54 | 0.726 | –0.05 |
| Greater than 33.3 | 34.3 | 4.56 | 36.2 | 1.8 | 0.70 | 0.017 | 0.40 | 34.6 | 0.2 | 0.47 | 0.627 | 0.05 |

See notes at end of table.

## Table 11-25. Household Study background questionnaire weighting effects for Oklahoma, by subgroup: 2003—Continued

| Subgroup | Base weight—eligible | | Base weight—respondents | | | | | Nonresponse-adjusted weight—respondents | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bias | | | Bias ratio (3)-(1) / (2) | | Bias | | | Bias ratio (4)-(1) / (2) |
| | Percent (1) | SE[1] (2) | Percent (3) | Estimate (3)-(1) | SE[1] | *p* value | | Percent (4) | Estimate (4)-(1) | SE[1] | *p* value | |
| Median income (in dollars) | | | | | | | | | | | | |
| 28,400 or less | 36.7 | 4.08 | 38.4 | 1.7 | 0.54 | 0.006* | 0.41 | 36.9 | 0.2 | 0.36 | 0.559 | 0.05 |
| 28,401–37,850 | 26.3 | 2.84 | 26.8 | 0.5 | 0.45 | 0.257 | 0.19 | 26.1 | –0.2 | 0.36 | 0.559 | –0.07 |
| 37,851–52,100 | 21.5 | 4.10 | 20.1 | –1.3 | 0.60 | 0.039 | –0.32 | 21.2 | –0.2 | 0.33 | 0.504 | –0.05 |
| Greater than 52,100 | 15.6 | 3.96 | 14.7 | –0.9 | 0.51 | 0.097 | –0.22 | 15.8 | 0.2 | 0.33 | 0.504 | 0.06 |
| Percent who rent | | | | | | | | | | | | |
| 16 or less | 27.4 | 4.93 | 27.0 | –0.4 | 0.59 | 0.454 | –0.09 | 27.6 | 0.2 | 0.35 | 0.542 | 0.04 |
| 17–31 | 33.7 | 5.08 | 33.5 | –0.2 | 0.85 | 0.803 | –0.04 | 33.4 | –0.3 | 0.83 | 0.770 | –0.05 |
| 32–59 | 28.7 | 5.39 | 29.8 | 1.1 | 0.69 | 0.138 | 0.20 | 29.2 | 0.5 | 0.51 | 0.339 | 0.09 |
| Greater than 59 | 10.2 | 2.75 | 9.9 | –0.4 | 0.84 | 0.639 | –0.15 | 9.8 | –0.5 | 0.72 | 0.516 | –0.17 |
| Age (years) | | | | | | | | | | | | |
| 16–29 | 27.8 | 1.33 | 29.3 | 1.5 | 0.68 | 0.040 | 1.14 | 29.2 | 1.4 | 0.65 | 0.041 | 1.07 |
| 30–49 | 35.2 | 1.66 | 34.3 | –0.9 | 0.55 | 0.122 | –0.54 | 34.2 | –1.0 | 0.55 | 0.094 | –0.59 |
| 50–69 | 25.6 | 1.58 | 25.6 | # | 0.60 | 0.987 | 0.01 | 25.7 | 0.1 | 0.51 | 0.933 | 0.03 |
| 70+ | 11.4 | 0.88 | 10.8 | –0.6 | 0.31 | 0.056 | –0.70 | 10.9 | –0.5 | 0.34 | 0.166 | –0.54 |
| Race/ethnicity | | | | | | | | | | | | |
| Hispanic | 5.6 | 0.85 | 5.7 | 0.2 | 0.36 | 0.693 | 0.18 | 5.6 | # | 0.36 | 0.977 | –0.01 |
| Non-Hispanic Black only | 6.1 | 0.77 | 5.9 | –0.2 | 0.39 | 0.621 | –0.26 | 5.9 | –0.1 | 0.37 | 0.726 | –0.17 |
| Other[3] | 88.4 | 0.96 | 88.4 | 0.1 | 0.64 | 0.937 | 0.05 | 88.5 | 0.1 | 0.59 | 0.815 | 0.15 |
| Gender | | | | | | | | | | | | |
| Male | 46.3 | 1.65 | 45.8 | –0.4 | 0.75 | 0.580 | –0.26 | 45.9 | –0.4 | 0.84 | 0.660 | –0.23 |
| Female | 53.8 | 1.65 | 54.2 | 1.4 | 0.75 | 0.580 | 0.26 | 54.1 | 0.4 | 0.84 | 0.660 | 0.23 |

† Not applicable.
# Rounds to zero.
* Statistically significant with Bonferroni adjustment at $\alpha = 0.05$.
[1] Standard error.
[2] Metropolitan Statistical Area.
[3] Includes non-Hispanic White, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races.
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

## 11.2 PRISON LITERACY STUDY

As in the household study, the NAAL prison study estimates are subject to potential bias owing to nonresponse at various stages of data collection. Data were collected using a background questionnaire, an assessment, and an oral module, which included multiple components. This section provides a systematic analysis of bias arising from nonparticipating prisons, nonresponding inmates, and item nonresponse. Section 11.2.1 provides unweighted and weighted response rates at the unit level. Section 11.2.2 summarizes the response rates at the item level. Section 11.2.3 provides the nonresponse bias analysis for the prison study.

### 11.2.1 Unit Nonresponse

The Prison Study had four stages of data collection where unit nonresponse occurred: (1) prison participation, (2) inmate background questionnaire, (3) inmate assessment, and (4) inmate oral module.

Variables known for both respondents and nonrespondents were used as selected analysis domain variables. The variables came from the following sources:

- Prison-level data from the 2000 Census of State and Federal Adult Correctional Facilities (referred to as the Census in this report);
- Prison-level data from the 2003 American Correctional Association (ACA) Directory;
- Prison-level updates obtained from websites and through telephone calls; and
- Inmate data from the background questionnaire.

Prison-level base weights were used in the weighted prison-level response rate calculations, and background questionnaire base weights were used for the background questionnaire, assessment, and oral module calculations. Response rates were calculated in the same manner as presented in the formula in section 7.1.1. The overall assessment weighted response rate is the product of the weighted response rates from the prison-level, background questionnaire and assessment.

Table 11-26 displays the variables and their sources. Table 11-27 shows the weighted unit response rates (refer to tables 7-13 and 7-14 for unweighted response rates for the prison study). There was a 97.5 percent weighted response rate at the prison level. Among types of prisons, state prisons had the lowest response rate (97.1 percent). Among security levels, prisons with supermaximum and maximum security levels combined had the lowest response rate (96.6 percent). Looking at gender composition, the only nonparticipating prisons were male-only facilities, with a response rate of 97.3 percent. Among census regions, the Midwest was the only region that experienced nonresponse, with a response rate of 87.9 percent.

**Table 11-26. Prison Study variables used to calculate response rates, by source and values: 2003**

| Variable description | Source[1] | Values |
|---|---|---|
| Prison level | | |
| Prison type | Census, ACA | Federal, state, private |
| Security level | Census, ACA, updates | Supermaximum, maximum, medium, minimum, administrative, other |
| Gender | ACA, updates | Male, female, mixed |
| Census region | Census, ACA | Northeast, Midwest, South, West |
| Inmate level | | |
| Age | Background questionnaire | 16-29, 30-49, 50-69, 70+ years |
| Gender | Background questionnaire | Male, female |
| Race | Background questionnaire | Hispanic, non-Hispanic Black only, other[2] |
| Education | Background questionnaire | Less than high school, high school, more than high school |

[1] 2000 Census of State and Federal Adult Correctional Facilities (Census), American Correctional Association Directory (ACA), and 2003 NAAL background questionnaire.
[2] Includes non-Hispanic White, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 11-27. NAAL Prison Study weighted unit response rates, by domain, in percent: 2003**

| Domain | Prison level | Background questionnaire | Assessment | Oral module | Overall[1] |
|---|---|---|---|---|---|
| Total | 97.5 | 90.6 | 98.8 | 96.9 | 87.2 |
| | | | | | |
| Prison type | | | | | |
| Federal | 100.0 | 93.4 | 97.9 | 95.1 | 91.4 |
| State | 97.1 | 89.9 | 98.9 | 97.1 | 86.4 |
| Private | 100.0 | 97.3 | 100.0 | 100.0 | 97.3 |
| | | | | | |
| Prison security level | | | | | |
| Supermaximum/maximum | 96.6 | 84.5 | 98.5 | 96.1 | 80.4 |
| Medium | 98.0 | 92.3 | 99.4 | 98.1 | 90.0 |
| Minimum | 97.2 | 92.5 | 98.5 | 95.9 | 88.6 |
| Administrative | 100.0 | 91.7 | 90.9 | 90.9 | 83.3 |
| Other | 100.0 | 95.8 | 95.7 | 95.7 | 91.7 |
| | | | | | |
| Prison gender composition | | | | | |
| Male | 97.3 | 90.4 | 98.7 | 96.7 | 86.8 |
| Female | 100.0 | 92.3 | 100.0 | 100.0 | 92.3 |
| Mixed | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| | | | | | |
| Prison census region | | | | | |
| Northeast | 100.0 | 86.9 | 98.7 | 90.9 | 85.8 |
| Midwest | 87.9 | 93.3 | 99.0 | 98.6 | 81.2 |
| South | 100.0 | 94.4 | 98.9 | 98.0 | 93.4 |
| West | 100.0 | 82.8 | 98.3 | 96.6 | 81.4 |
| | | | | | |
| Inmate age (years) | | | | | |
| 16–29 | — | — | 99.5 | 97.7 | — |
| 30–49 | — | — | 98.2 | 96.2 | — |
| 50–69 | — | — | 100.0 | 98.2 | — |
| 70+ | — | — | 100.0 | 100.0 | — |
| | | | | | |
| Inmate gender | | | | | |
| Male | — | — | 98.7 | 96.7 | — |
| Female | — | — | 100.0 | 100.0 | — |
| | | | | | |
| Inmate race | | | | | |
| Hispanic | — | — | 99.6 | 93.8 | — |
| Non-Hispanic Black only | — | — | 98.2 | 97.0 | — |
| Other[2] | — | — | 99.1 | 98.4 | — |
| | | | | | |
| Inmate education | | | | | |
| Less than high school | — | — | 98.8 | 96.6 | — |
| High school | — | — | 98.5 | 97.1 | — |
| More than high school | — | — | 99.4 | 97.7 | — |

—Not available.

[1] The overall response rate is the product of the response rates from the prison level, background questionnaire, and assessment.

[2] Includes non-Hispanic White, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

At the background questionnaire level, the weighted response rate was 90.6 percent. Across the domains, the West region had the lowest response rate (82.8 percent). For background questionnaire respondents, the weighted response rate was 98.8 percent for the assessment and 96.9 percent for the oral module. Response rates for the assessment and oral module were all above 94 percent, with the exception of relatively low oral module response rates in the Northeast (90.9 percent) and for Hispanic adults (93.8 percent) and relatively low assessment and oral module response rates (90.9 percent) for administrative prisons.

The overall weighted response rate for the prison study was 87.2 percent. The lowest response rate among types of prisons was for state prisons (86.4 percent). Among security levels, prisons classified as either supermaximum or maximum had the lowest response rate (80.4 percent). Looking at gender composition, male-only prisons had the lowest response rate (86.8 percent). Among census regions, the Midwest and West regions had the lowest response rates (81.2 percent and 81.4 percent, respectively).

### 11.2.2 Item Nonresponse Rates

Response rates were calculated for 280 items on the prison study background questionnaire. Two hundred items had a 100 percent response rate, 78 items had a response rate between 99.0 and 99.9 percent, 1 item had an 80.5 percent response rate, and 1 item had a 71.0 percent response rate. The last two items are shown in table 11-28.

**Table 11-28. Prison Study item response rates below 85 percent, by variable: 2003**

| Variable | Description | Number of eligibles | Unweighted item response rate (percent) | Weighted item response rate (percent) |
|---|---|---|---|---|
| BQ1830 | What was the highest level of education your mother (stepmother, or female guardian) completed? | 1,161 | 80.3 | 80.5 |
| BQ1840 | What was the highest level of education your father (stepfather, or male guardian) completed? | 1,161 | 70.7 | 71.0 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

The numerator in the item response rate consisted of the item respondents; the denominator contained all unit respondents, excluding those with a valid skip for the item. Unweighted item response rates ranged from 70.7 percent to 100 percent, with a median of 100 percent. Weighted item response rates ranged from 71.0 percent to 100 percent, with a median of 100 percent.

### 11.2.3 Nonresponse Bias Analysis

As noted in section 11.1.4, NCES Standard 4-4-1 requires an evaluation of the potential magnitude of nonresponse bias for any survey stage with a unit or item response rate of less than 85 percent. The final prison and inmate (for the background questionnaire, assessment, and oral module) weighted response rates were all above 85 percent. Because of the high response rate, the nonresponse bias analysis was not necessary to examine unit nonresponse.

For item nonresponse, only two items on the background questionnaire fell below the 85 percent weighted response rate threshold. These two items are “What was the highest level of education your mother (stepmother, or female guardian) completed?” with an 80.5 percent weighted response rate, and “What was the highest level of education your father (stepfather, or male guardian) completed?” with a 71.0 percent weighted response rate. Because these two items do not meet the NCES nonresponse standard, a bivariate analysis (section 11.2.3.1), an item nonresponse bias estimation (section 11.2.3.2), and a multivariate analysis (section 11.2.3.3) were performed for each item to examine the potential bias owing to nonresponse.

#### 11.2.3.1 Bivariate Analysis

For the two items with a low response rate in table 11-28, the distributions of item respondents and eligibles (i.e., unit respondents) were compared on survey domains such as key characteristics and background questionnaire items with response rates of 100 percent. To test for independence between the response indicators and survey domains, a Rao-Scott (RS3) chi-square test was computed, using WesVar. Final weights were used in the analysis. The results are shown in table 11-29. Replicate weights were used to adequately reflect the effect of two-stage cluster sampling.

**Table 11-29. Prison Study sample distribution for item respondents versus eligibles, by survey domain: 2003**

| Survey domain | Respondents: Sample size | Respondents: Domain percent | Respondents: Standard error | Eligibles: Sample size | Eligibles: Domain percent | Eligibles: Standard error | Chi-square: Statistic | Chi-square: *p* value |
|---|---|---|---|---|---|---|---|---|
| **A. BQ1830: What was the highest level of education your mother (stepmother, or female guardian) completed?** | | | | | | | | |
| Prison security level | | | | | | | | |
| 1: Supermax, maximum | 303 | 34.5 | 0.67 | 368 | 33.5 | 0.18 | 4.97 | 0.077 |
| 2: Medium | 416 | 46.9 | 0.79 | 536 | 48.5 | 0.14 | | |
| 3: Minimum/other | 213 | 18.5 | 0.51 | 257 | 18.0 | 0.09 | | |
| Region/prison type | | | | | | | | |
| 1: Northeast | 102 | 12.0 | 0.73 | 132 | 12.3 | 0.18 | 5.09 | 0.159 |
| 2: Midwest | 177 | 18.9 | 0.39 | 208 | 17.7 | 0.07 | | |
| 3: South | 387 | 38.7 | 0.71 | 480 | 38.4 | 0.16 | | |
| 4: West | 161 | 19.4 | 0.44 | 205 | 19.9 | 0.10 | | |
| 5: Federal | 105 | 11.0 | 0.40 | 136 | 11.7 | 0.03 | | |
| Inmate gender | | | | | | | | |
| 1: Male | 868 | 93.1 | 0.22 | 1,086 | 93.6 | 0.07 | 6.58 | 0.010 |
| 2: Female | 64 | 6.9 | 0.22 | 75 | 6.4 | 0.07 | | |
| Inmate race/ethnicity | | | | | | | | |
| 1: Hispanic | 158 | 15.8 | 0.63 | 223 | 17.9 | 0.13 | 12.82 | 0.002 |
| 2: Non-Hispanic Black only | 403 | 46.3 | 0.87 | 491 | 45.8 | 0.14 | | |
| 3: Other | 371 | 37.9 | 0.75 | 447 | 36.3 | 0.17 | | |
| Inmate age category | | | | | | | | |
| 1: 16–29 | 310 | 36.9 | 0.67 | 388 | 37.3 | 0.13 | 1.39 | 0.477 |
| 2: 30–49 | 526 | 55.4 | 0.64 | 659 | 55.4 | 0.16 | | |
| 3: 50+ | 96 | 7.6 | 0.31 | 114 | 7.3 | 0.10 | | |
| Inmate highest level of education | | | | | | | | |
| 1: Less than high school | 325 | 33.3 | 1.03 | 458 | 37.5 | 0.17 | 19.36 | 0.000 |
| 2: High school or higher | 607 | 66.7 | 1.03 | 703 | 62.5 | 0.17 | | |
| Inmate marital status | | | | | | | | |
| 1: Never married | 424 | 44.1 | 0.78 | 530 | 44.1 | 0.15 | 0.00 | 0.995 |
| 0: Other | 508 | 55.9 | 0.78 | 631 | 55.9 | 0.15 | | |
| Inmate country of birth | | | | | | | | |
| 0: Born elsewhere | 82 | 9.0 | 0.49 | 110 | 9.5 | 0.25 | 1.56 | 0.212 |
| 1: Born in the U.S. | 850 | 91.0 | 0.49 | 1,051 | 90.5 | 0.25 | | |
| BQ1100: Which language do you usually speak now? | | | | | | | | |
| 0: Other | 57 | 6.0 | 0.56 | 78 | 6.4 | 0.53 | 1.07 | 0.300 |
| 1: English | 875 | 94.0 | 0.56 | 1,083 | 93.6 | 0.53 | | |
| BQ1140: With regard to the English language, how well do you... understand it when it is spoken to you? | | | | | | | | |
| 1: Very well | 717 | 77.5 | 1.68 | 882 | 76.8 | 1.37 | 3.75 | 0.143 |
| 2: Well | 183 | 19.1 | 1.72 | 230 | 19.2 | 1.50 | | |
| 3: Not well, not at all | 32 | 3.3 | 0.59 | 49 | 3.9 | 0.60 | | |

See notes at end of table.

**Table 11-29. Prison Study sample distribution for item respondents versus eligibles, by survey domain: 2003—Continued**

| Survey domain | Respondents | | | Eligibles | | | Chi-square | |
|---|---|---|---|---|---|---|---|---|
| | Sample size | Domain percent | Standard error | Sample size | Domain percent | Standard error | Statistic | *p* value |
| **A. BQ1830: What was the highest level of education your mother (stepmother, or female guardian) completed?** | | | | | | | | |
| BQ1155: With regard to the English language, how well do you... write it? | | | | | | | | |
| 1: Very well | 549 | 59.3 | 2.00 | 658 | 57.4 | 1.75 | 6.19 | 0.100 |
| 2: Well | 271 | 29.2 | 2.04 | 348 | 30.0 | 1.79 | | |
| 3: Not well | 84 | 8.7 | 1.00 | 114 | 9.4 | 0.91 | | |
| 4: Not at all | 28 | 2.9 | 0.54 | 41 | 3.3 | 0.54 | | |
| BQ1208: Since your most recent admission to prison, have you completed any additional education? | | | | | | | | |
| 1: Yes | 331 | 35.6 | 1.72 | 403 | 35.0 | 1.53 | 0.77 | 0.381 |
| 2: No | 601 | 64.4 | 1.72 | 758 | 65.0 | 1.53 | | |
| BQ1245: Have you ever taken part in a program other than in regular school in order to improve your basic skills, that is, basic reading, writing, and arithmetic skills? The program may have been in prison or it may have been outside of prison. | | | | | | | | |
| 1: Yes | 224 | 23.9 | 1.72 | 287 | 24.7 | 1.64 | 2.53 | 0.112 |
| 2: No | 708 | 76.1 | 1.72 | 874 | 75.3 | 1.64 | | |
| BQ1500: Have you ever been placed on probation, either as a juvenile or as an adult? | | | | | | | | |
| 1: Yes | 693 | 74.8 | 1.50 | 859 | 74.8 | 1.40 | 0.00 | 0.956 |
| 2: No | 239 | 25.2 | 1.50 | 299 | 25.2 | 1.40 | | |
| BQ1560: In the year before your incarceration on {BGQ1490}, did you receive income from... unemployment insurance compensation and/or workman's compensation? | | | | | | | | |
| 1: Yes | 63 | 6.5 | 0.73 | 77 | 6.4 | 0.65 | 0.06 | 0.808 |
| 2: No | 869 | 93.5 | 0.73 | 1,083 | 93.6 | 0.65 | | |
| BQ1785: Do you ever use a computer? | | | | | | | | |
| 1: Yes | 248 | 26.3 | 1.84 | 292 | 25.0 | 1.73 | 4.92 | 0.027 |
| 2: No | 684 | 73.7 | 1.84 | 869 | 75.0 | 1.73 | | |
| BQ1855: Have you ever received… food stamps? | | | | | | | | |
| 1: Yes | 266 | 28.6 | 1.55 | 336 | 28.7 | 1.48 | 0.01 | 0.919 |
| 2: No | 666 | 71.4 | 1.55 | 823 | 71.3 | 1.48 | | |
| BQ1945: In general, how would you rate your overall health? Would you say it is… | | | | | | | | |
| 1: Excellent | 265 | 28.9 | 1.86 | 319 | 28.0 | 1.66 | 14.36 | 0.005 |
| 2: Very Good | 335 | 36.4 | 2.03 | 399 | 35.0 | 1.79 | | |
| 3: Good | 205 | 21.3 | 1.31 | 259 | 21.7 | 1.18 | | |
| 4: Fair | 91 | 9.7 | 0.88 | 132 | 11.1 | 0.88 | | |
| 5: Poor | 36 | 3.6 | 0.57 | 51 | 4.2 | 0.51 | | |
| **B. BQ1840: What was the highest level of education your father (stepfather, or male guardian) completed?** | | | | | | | | |
| Prison security level | | | | | | | | |
| 1:Supermax, maximum | 258 | 33.4 | 1.06 | 368 | 33.5 | 0.18 | 0.83 | 0.642 |
| 2: Medium | 375 | 47.9 | 1.12 | 536 | 48.5 | 0.14 | | |
| 3: Minimum/other | 188 | 18.7 | 0.72 | 257 | 18.0 | 0.09 | | |
| Region/prison type | | | | | | | | |
| 1: Northeast | 83 | 10.9 | 0.91 | 132 | 12.3 | 0.18 | 6.87 | 0.104 |
| 2: Midwest | 159 | 19.4 | 0.65 | 208 | 17.7 | 0.07 | | |
| 3: South | 341 | 38.8 | 1.01 | 480 | 38.4 | 0.16 | | |
| 4: West | 147 | 20.0 | 0.78 | 205 | 19.9 | 0.10 | | |
| 5: Federal | 91 | 10.9 | 0.88 | 136 | 11.7 | 0.03 | | |

See notes at end of table.

**Table 11-29. Prison Study sample distribution for item respondents versus eligibles, by survey domain: 2003—Continued**

| Survey domain | Respondents: Sampl e size | Respondents: Domain percent | Respondents: Standard error | Eligibles: Sample size | Eligibles: Domain percent | Eligibles: Standard error | Chi-square: Statistic | Chi-square: *p* value |
|---|---|---|---|---|---|---|---|---|
| **B. BQ1840: What was the highest level of education your father (stepfather, or male guardian) completed?** | | | | | | | | |
| Inmate gender | | | | | | | | |
| 1: Male | 773 | 94.1 | 0.36 | 1,086 | 93.6 | 0.07 | 1.30 | 0.254 |
| 2: Female | 48 | 5.9 | 0.36 | 75 | 6.4 | 0.07 | | |
| Inmate race/ethnicity | | | | | | | | |
| 1: Hispanic | 147 | 16.6 | 0.72 | 223 | 17.9 | 0.13 | 14.43 | 0.001 |
| 2: Non-Hispanic Black only | 332 | 43.7 | 1.02 | 491 | 45.8 | 0.14 | | |
| 3: Other[1] | 342 | 39.8 | 0.95 | 447 | 36.3 | 0.17 | | |
| Inmate age category | | | | | | | | |
| 1: 16–29 | 274 | 37.1 | 0.77 | 388 | 37.3 | 0.13 | 0.15 | 0.928 |
| 2: 30–49 | 467 | 55.7 | 0.81 | 659 | 55.4 | 0.16 | | |
| 3: 50+ | 80 | 7.3 | 0.41 | 114 | 7.3 | 0.10 | | |
| Inmate highest level of education | | | | | | | | |
| 1: Less than high school | 297 | 34.6 | 0.99 | 458 | 37.5 | 0.17 | 9.37 | 0.002 |
| 2: High school or higher | 524 | 65.4 | 0.99 | 703 | 62.5 | 0.17 | | |
| Inmate marital status | | | | | | | | |
| 1: Never married | 380 | 44.8 | 1.10 | 530 | 44.1 | 0.15 | 0.31 | 0.576 |
| 0: Other | 441 | 55.2 | 1.10 | 631 | 55.9 | 0.15 | | |
| Inmate country of birth | | | | | | | | |
| 0: Born elsewhere | 75 | 9.3 | 0.58 | 110 | 9.5 | 0.25 | 0.14 | 0.710 |
| 1: Born in the U.S. | 746 | 90.7 | 0.58 | 1,051 | 90.5 | 0.25 | | |
| BQ1100: Which language do you usually speak now? | | | | | | | | |
| 0: Other | 52 | 6.1 | 0.65 | 78 | 6.4 | 0.53 | 0.30 | 0.584 |
| 1: English | 769 | 93.9 | 0.65 | 1,083 | 93.6 | 0.53 | | |
| BQ1140: With regard to the English language, how well do you... understand it when it is spoken to you? | | | | | | | | |
| 1: Very well | 623 | 76.6 | 1.75 | 882 | 76.8 | 1.37 | 3.89 | 0.140 |
| 2: Well | 171 | 20.2 | 1.79 | 230 | 19.2 | 1.50 | | |
| 3: Not well, not at all | 27 | 3.2 | 0.69 | 49 | 3.9 | 0.60 | | |
| BQ1155: With regard to the English language, how well do you... write it? | | | | | | | | |
| 1: Very well | 481 | 59.3 | 2.11 | 658 | 57.4 | 1.75 | 3.81 | 0.238 |
| 2: Well | 241 | 29.1 | 2.20 | 348 | 30.0 | 1.79 | | |
| 3: Not well | 76 | 8.9 | 1.02 | 114 | 9.4 | 0.91 | | |
| 4: Not at all | 23 | 2.7 | 0.61 | 41 | 3.3 | 0.54 | | |
| BQ1208: Since your most recent admission to prison, have you completed any additional education? | | | | | | | | |
| 1: Yes | 287 | 35.2 | 1.81 | 403 | 35.0 | 1.53 | 0.02 | 0.889 |
| 2: No | 534 | 64.8 | 1.81 | 758 | 65.0 | 1.53 | | |
| BQ1245: Have you ever taken part in a program other than in regular school in order to improve your basic skills, that is, basic reading, writing, and arithmetic skills? The program may have been in prison or it may have been outside of prison. | | | | | | | | |
| 1: Yes | 203 | 24.7 | 1.84 | 287 | 24.7 | 1.64 | 0.00 | 0.962 |
| 2: No | 618 | 75.3 | 1.84 | 874 | 75.3 | 1.64 | | |

See notes at end of table.

**Table 11-29. Prison Study sample distribution for item respondents versus eligibles, by survey domain: 2003—Continued**

| Survey domain | Respondents Sample size | Respondents Domain percent | Respondents Standard error | Eligibles Sample size | Eligibles Domain percent | Eligibles Standard error | Chi-square Statistic | Chi-square *p* value |
|---|---|---|---|---|---|---|---|---|
| **B. BQ1840: What was the highest level of education your father (stepfather, or male guardian) completed?** | | | | | | | | |
| BQ1500: Have you ever been placed on probation, either as a juvenile or as an adult? | | | | | | | | |
| 1: Yes | 606 | 74.5 | 1.59 | 859 | 74.8 | 1.40 | 0.10 | 0.756 |
| 2: No | 215 | 25.5 | 1.59 | 299 | 25.2 | 1.40 | | |
| BQ1560: In the year before your incarceration on {BGQ1490}, did you receive income from... unemployment insurance compensation and/or workman's compensation? | | | | | | | | |
| 1: Yes | 61 | 7.1 | 0.84 | 77 | 6.4 | 0.65 | 2.72 | 0.099 |
| 2: No | 760 | 92.9 | 0.84 | 1083 | 93.6 | 0.65 | | |
| BQ1785: Do you ever use a computer? | | | | | | | | |
| 1: Yes | 224 | 27.1 | 2.07 | 292 | 25.0 | 1.73 | 7.01 | 0.008 |
| 2: No | 597 | 72.9 | 2.07 | 869 | 75.0 | 1.73 | | |
| BQ1855: Have you ever received… food stamps? | | | | | | | | |
| 1: Yes | 220 | 27.0 | 1.65 | 336 | 28.7 | 1.48 | 4.25 | 0.039 |
| 2: No | 601 | 73.0 | 1.65 | 823 | 71.3 | 1.48 | | |
| BQ1945: In general, how would you rate your overall health? Would you say it is… | | | | | | | | |
| 1: Excellent | 238 | 29.5 | 1.96 | 319 | 28.0 | 1.66 | 6.92 | 0.088 |
| 2: Very Good | 287 | 35.5 | 2.06 | 399 | 35.0 | 1.79 | | |
| 3: Good | 182 | 21.4 | 1.17 | 259 | 21.7 | 1.18 | | |
| 4: Fair | 84 | 10.0 | 0.87 | 132 | 11.1 | 0.88 | | |
| 5: Poor | 30 | 3.6 | 0.59 | 51 | 4.2 | 0.51 | | |

[1] Includes non-Hispanic White, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races.
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

The chi-square tests show that there was a significant relationship between the response indicator for "What was the highest level of education your mother (stepmother, or female guardian) completed?" and 5 of the 20 survey domains (gender, race/ethnicity, education, ever used or not used a computer, and health).

The chi-square tests also show a significant relationship between the response indicator for "What was the highest level of education your father (stepfather, or male guardian) completed?" and 4 of the 20 survey domains (race/ethnicity, education, ever used or not used a computer, and ever received food stamps).

### 11.2.3.2 Item Nonresponse Bias Estimates

To understand the magnitude of the potential bias, estimates of bias were computed (as shown in table 11-30). *T* tests were performed to determine whether the bias was significantly different from 0. The *t* tests used a simple Bonferroni adjustment so that the overall α-level equaled 0.05 for each domain variable. The adjustment was computed as $\alpha'=\alpha/g$, where $g$ is the number of comparisons. For example, for prison security level, three *t* tests were conducted. The Bonferroni adjustment is $\alpha'=0.05/3=0.0166$. The results are fairly consistent with the chi-square analysis described in section 11.2.3.1.

With regard to mother's education, the survey domains with statistically significant bias estimates are the Midwest, males, females, Hispanics, levels of education, and having fair health. With regard to father's education, the survey domains with statistically significant bias estimates are the Midwest, race/ethnicity other than Black or Hispanic, levels of education, and ever used or not used a computer.

Even though there are some statistically significant relationships between the response indicators and the survey domains, there do not seem to be important indications of potential bias. Almost all absolute bias estimates are within 2 percentage points. However, among the statistically significant bias estimates, there are exceptions for mother's education, namely, –2.1 percent (–11.73 percent relative bias) for Hispanic inmates, –4.3 percent for inmates with low educational attainment (–11.47 percent relative bias), and 4.3 percent for inmates with high educational attainment (6.88 percent relative bias).

## Table 11-30. Prison Study estimate of item nonresponse bias, by survey domain: 2003

| Survey domain | Eligibles Domain percent | Eligibles SE[1] | Respondent domain percent | Nonrespondent domain percent | Bias Estimate | Bias SE[1] | Bias *p* value | Relative bias | Bias ratio |
|---|---|---|---|---|---|---|---|---|---|
| **A. BQ1830: What was the highest level of education your mother (stepmother, or female guardian) completed?** | | | | | | | | | |
| Prison security level | | | | | | | | | |
| 1:Supermax, maximum | 33.5 | 0.18 | 34.5 | 29.3 | 1.0 | 0.64 | 0.118 | 2.99 | 5.56 |
| 2: Medium | 48.5 | 0.14 | 46.9 | 55.1 | –1.6 | 0.75 | 0.040 | –3.30 | –11.43 |
| 3: Minimum/ other | 18.0 | 0.09 | 18.5 | 15.7 | 0.6 | 0.50 | 0.266 | 3.33 | 6.67 |
| Region/prison type | | | | | | | | | |
| 1: Northeast | 12.3 | 0.18 | 12.0 | 13.8 | –0.4 | 0.65 | 0.588 | –3.25 | –2.22 |
| 2: Midwest | 17.7 | 0.07 | 18.9 | 12.9 | 1.2 | 0.38 | 0.003* | 6.78 | 17.14 |
| 3: South | 38.4 | 0.16 | 38.7 | 37.0 | 0.3 | 0.68 | 0.616 | 0.78 | 1.88 |
| 4: West | 19.9 | 0.10 | 19.4 | 22.0 | –0.5 | 0.47 | 0.296 | –2.51 | –5.00 |
| 5: Federal | 11.7 | 0.03 | 11.0 | 14.4 | –0.7 | 0.40 | 0.106 | –5.98 | –23.33 |
| Inmate gender | | | | | | | | | |
| 1: Male | 93.6 | 0.07 | 93.1 | 95.7 | –0.5 | 0.19 | 0.010* | –0.53 | –7.14 |
| 2: Female | 6.4 | 0.07 | 6.9 | 4.3 | 0.5 | 0.19 | 0.010* | 7.81 | 7.14 |
| Inmate race/ethnicity | | | | | | | | | |
| 1: Hispanic | 17.9 | 0.13 | 15.8 | 26.8 | –2.1 | 0.59 | 0.001* | –11.73 | –16.15 |
| 2: Non-Hispanic Black only | 45.8 | 0.14 | 46.3 | 43.3 | 0.6 | 0.84 | 0.485 | 1.31 | 4.29 |
| 3: Other2 | 36.3 | 0.17 | 37.9 | 29.9 | 1.5 | 0.72 | 0.036 | 4.13 | 8.82 |
| Inmate age category | | | | | | | | | |
| 1: 16–29 | 37.3 | 0.13 | 36.9 | 39.0 | –0.4 | 0.64 | 0.538 | –1.07 | –3.08 |
| 2: 30–49 | 55.4 | 0.16 | 55.4 | 55.2 | 0.0 | 0.62 | 0.943 | 0.00 | 0.00 |
| 3: 50+ | 7.3 | 0.10 | 7.6 | 5.8 | 0.4 | 0.29 | 0.226 | 5.48 | 4.00 |
| Inmate highest level of education | | | | | | | | | |
| 1: Less than high school | 37.5 | 0.17 | 33.3 | 55.2 | –4.3 | 0.88 | 0.000* | –11.47 | –25.29 |
| 2: High school or higher | 62.5 | 0.17 | 66.7 | 44.8 | 4.3 | 0.88 | 0.000* | 6.88 | 25.29 |
| Inmate marital status | | | | | | | | | |
| 1: Never married | 44.1 | 0.15 | 44.1 | 44.2 | 0.0 | 0.80 | 0.995 | 0.00 | 0.00 |
| 0: Other | 55.9 | 0.15 | 55.9 | 55.8 | 0.0 | 0.80 | 0.995 | 0.00 | 0.00 |
| Inmate country of birth | | | | | | | | | |
| 0: Born elsewhere | 9.5 | 0.25 | 9.0 | 11.8 | –0.5 | 0.44 | 0.217 | –5.26 | –2.00 |
| 1: Born in the U.S. | 90.5 | 0.25 | 91.0 | 88.2 | 0.5 | 0.44 | 0.217 | 0.55 | 2.00 |

See notes at end of table.

## Table 11-30. Prison Study estimate of item nonresponse bias, by survey domain: 2003—Continued

| Survey domain | Eligibles Domain percent | Eligibles SE[1] | Respondent domain percent | Nonrespondent domain percent | Bias Estimate | Bias SE[1] | Bias *p* value | Relative bias | Bias ratio |
|---|---|---|---|---|---|---|---|---|---|
| **A. BQ1830: What was the highest level of education your mother (stepmother, or female guardian) completed?** | | | | | | | | | |
| BQ1100: Which language do you usually speak now? | | | | | | | | | |
| 0: Other | 6.4 | 0.53 | 6.0 | 8.1 | –0.4 | 0.40 | 0.301 | –6.25 | –0.75 |
| 1: English | 93.6 | 0.53 | 94.0 | 91.9 | 0.4 | 0.40 | 0.301 | 0.43 | 0.75 |
| BQ1140: With regard to the English language, how well do you... understand it when it is spoken to you? | | | | | | | | | |
| 1: Very well | 76.8 | 1.37 | 77.5 | 73.9 | 0.7 | 0.74 | 0.346 | 0.91 | 0.51 |
| 2: Well | 19.2 | 1.50 | 19.1 | 19.6 | –0.1 | 0.65 | 0.887 | –0.52 | –0.07 |
| 3:Not well, not at all | 3.9 | 0.60 | 3.3 | 6.5 | –0.6 | 0.29 | 0.041 | –15.38 | –1.00 |
| BQ1155: With regard to the English language, how well do you... write it? | | | | | | | | | |
| 1: Very well | 57.4 | 1.75 | 59.3 | 49.5 | 1.9 | 0.83 | 0.026 | 3.31 | 1.09 |
| 2: Well | 30.0 | 1.79 | 29.2 | 33.2 | –0.8 | 0.77 | 0.315 | –2.67 | –0.45 |
| 3: Not well | 9.4 | 0.91 | 8.7 | 12.2 | –0.7 | 0.52 | 0.197 | –7.45 | –0.77 |
| 4: Not at all | 3.3 | 0.54 | 2.9 | 5.1 | –0.4 | 0.31 | 0.165 | –12.12 | –0.74 |
| BQ1208: Since your most recent admission to prison, have you completed any additional education? | | | | | | | | | |
| 1: Yes | 35.0 | 1.53 | 35.6 | 32.6 | 0.6 | 0.65 | 0.383 | 1.71 | 0.39 |
| 2: No | 65.0 | 1.53 | 64.4 | 67.4 | –0.6 | 0.65 | 0.383 | –0.92 | –0.39 |
| BQ1245: Have you ever taken part in a program other than in regular school in order to improve your basic skills, that is, basic reading, writing, and arithmetic skills? The program may have been in prison or it may have been outside of prison. | | | | | | | | | |
| 1: Yes | 24.7 | 1.64 | 23.9 | 28.0 | –0.8 | 0.49 | 0.117 | –3.24 | –0.49 |
| 2: No | 75.3 | 1.64 | 76.1 | 72.0 | 0.8 | 0.49 | 0.117 | 1.06 | 0.49 |
| BQ1500: Have you ever been placed on probation, either as a juvenile or as an adult? | | | | | | | | | |
| 1: Yes | 74.8 | 1.40 | 74.8 | 74.6 | 0.0 | 0.68 | 0.957 | 0.00 | 0.00 |
| 2: No | 25.2 | 1.40 | 25.2 | 25.4 | 0.0 | 0.68 | 0.957 | 0.00 | 0.00 |
| BQ1560: In the year before your incarceration on {BGQ1490}, did you receive income from... unemployment insurance compensation and/or workman's compensation? | | | | | | | | | |
| 1: Yes | 6.4 | 0.65 | 6.5 | 6.0 | 0.1 | 0.35 | 0.810 | 1.56 | 0.15 |
| 2: No | 93.6 | 0.65 | 93.5 | 94.0 | –0.1 | 0.35 | 0.810 | –0.11 | –0.15 |
| BQ1785: Do you ever use a computer? | | | | | | | | | |
| 1: Yes | 25.0 | 1.73 | 26.3 | 19.4 | 1.3 | 0.61 | 0.032 | 5.20 | 0.75 |
| 2: No | 75.0 | 1.73 | 73.7 | 80.6 | –1.3 | 0.61 | 0.032 | –1.73 | –0.75 |
| BQ1855: Have you ever received… food stamps? | | | | | | | | | |
| 1: Yes | 28.7 | 1.48 | 28.6 | 28.9 | –0.1 | 0.59 | 0.919 | –0.35 | –0.07 |
| 2: No | 71.3 | 1.48 | 71.4 | 71.1 | 0.1 | 0.59 | 0.919 | 0.14 | 0.07 |
| BQ1945: In general, how would you rate your overall health? Would you say it is… | | | | | | | | | |
| 1: Excellent | 28.0 | 1.66 | 28.9 | 24.3 | 0.9 | 0.59 | 0.135 | 3.21 | 0.54 |
| 2: Very Good | 35.0 | 1.79 | 36.4 | 29.1 | 1.4 | 0.73 | 0.057 | 4.00 | 0.78 |
| 3: Good | 21.7 | 1.18 | 21.3 | 23.2 | –0.4 | 0.61 | 0.543 | –1.84 | –0.34 |
| 4: Fair | 11.1 | 0.88 | 9.7 | 17.0 | –1.4 | 0.50 | 0.007* | –12.61 | –1.59 |
| 5: Poor | 4.2 | 0.51 | 3.6 | 6.4 | –0.5 | 0.29 | 0.064 | –11.90 | –0.98 |

See notes at end of table.

**Table 11-30. Prison Study estimate of item nonresponse bias, by survey domain: 2003—Continued**

| Survey domain | Eligibles: Domain percent | Eligibles: SE[1] | Respondent domain percent | Nonrespondent domain percent | Bias: Estimate | Bias: SE[1] | Bias: *p* value | Relative bias | Bias ratio |
|---|---|---|---|---|---|---|---|---|---|
| **B. BQ1840: What was the highest level of education your father (stepfather, or male guardian) completed?** | | | | | | | | | |
| Prison security level | | | | | | | | | |
| 1: Supermax, maximum | 33.5 | 0.18 | 33.4 | 33.8 | –0.1 | 1.01 | 0.905 | –0.30 | –0.56 |
| 2: Medium | 48.5 | 0.14 | 47.9 | 49.9 | –0.6 | 1.09 | 0.591 | –1.24 | –4.29 |
| 3: Minimum/ other | 18.0 | 0.09 | 18.7 | 16.3 | 0.7 | 0.70 | 0.313 | 3.89 | 7.78 |
| Region/prison type | | | | | | | | | |
| 1: Northeast | 12.3 | 0.18 | 10.9 | 15.7 | –1.4 | 0.78 | 0.075 | –11.38 | –7.78 |
| 2: Midwest | 17.7 | 0.07 | 19.4 | 13.6 | 1.7 | 0.61 | 0.008* | 9.60 | 24.29 |
| 3: South | 38.4 | 0.16 | 38.8 | 37.4 | 0.4 | 0.98 | 0.672 | 1.04 | 2.50 |
| 4: West | 19.9 | 0.10 | 20.0 | 19.8 | 0.1 | 0.77 | 0.927 | 0.50 | 1.00 |
| 5: Federal | 11.7 | 0.03 | 10.9 | 13.5 | –0.8 | 0.88 | 0.380 | –6.84 | –26.67 |
| Inmate gender | | | | | | | | | |
| 1: Male | 93.6 | 0.07 | 94.1 | 92.5 | 0.5 | 0.40 | 0.260 | 0.53 | 7.14 |
| 2: Female | 6.4 | 0.07 | 5.9 | 7.5 | –0.5 | 0.40 | 0.260 | –7.81 | –7.14 |
| Inmate race/ethnicity | | | | | | | | | |
| 1: Hispanic | 17.9 | 0.13 | 16.6 | 21.2 | –1.4 | 0.71 | 0.058 | –7.82 | –10.77 |
| 2: Non-Hispanic Black only | 45.8 | 0.14 | 43.7 | 50.8 | –2.1 | 0.98 | 0.037 | –4.59 | –15.00 |
| 3: Other | 36.3 | 0.17 | 39.8 | 28.0 | 3.5 | 0.88 | 0.000* | 9.64 | 20.59 |
| Inmate age category | | | | | | | | | |
| 1: 16–29 | 37.3 | 0.13 | 37.1 | 38.0 | –0.3 | 0.76 | 0.714 | –0.80 | –2.31 |
| 2: 30–49 | 55.4 | 0.16 | 55.7 | 54.7 | 0.3 | 0.81 | 0.719 | 0.54 | 1.88 |
| 3: 50+ | 7.3 | 0.10 | 7.3 | 7.3 | 0.0 | 0.41 | 0.979 | 0.00 | 0.00 |
| Inmate highest level of education | | | | | | | | | |
| 1: Less than high school | 37.5 | 0.17 | 34.6 | 44.7 | –2.9 | 0.95 | 0.003* | –7.73 | –17.06 |
| 2: High school or higher | 62.5 | 0.17 | 65.4 | 55.3 | 2.9 | 0.95 | 0.003* | 4.64 | 17.06 |
| Inmate marital status | | | | | | | | | |
| 1: Never married | 44.1 | 0.15 | 44.8 | 42.6 | 0.6 | 1.12 | 0.580 | 1.36 | 4.00 |
| 0: Other | 55.9 | 0.15 | 55.2 | 57.4 | –0.6 | 1.12 | 0.580 | –1.07 | –4.00 |
| Inmate country of birth | | | | | | | | | |
| 0: Born elsewhere | 9.5 | 0.25 | 9.3 | 10.0 | –0.2 | 0.54 | 0.712 | –2.11 | –0.80 |
| 1: Born in the U.S. | 90.5 | 0.25 | 90.7 | 90.0 | 0.2 | 0.54 | 0.712 | 0.22 | 0.80 |

See notes at end of table.

**Table 11-30. Prison Study estimate of item nonresponse bias, by survey domain: 2003—Continued**

| Survey domain | Eligibles Domain percent | Eligibles SE[1] | Respondent domain percent | Nonrespondent domain percent | Bias Estimate | Bias SE[1] | Bias p value | Relative bias | Bias ratio |
|---|---|---|---|---|---|---|---|---|---|
| **B. BQ1840: What was the highest level of education your father (stepfather, or male guardian) completed?** | | | | | | | | | |
| BQ1100: Which language do you usually speak now? | | | | | | | | | |
| 0: Other | 6.4 | 0.53 | 6.1 | 7.0 | –0.3 | 0.47 | 0.586 | –4.69 | –0.57 |
| 1: English | 93.6 | 0.53 | 93.9 | 93.0 | 0.3 | 0.47 | 0.586 | 0.32 | 0.57 |
| BQ1140: With regard to the English language, how well do you... understand it when it is spoken to you? | | | | | | | | | |
| 1: Very well | 76.8 | 1.37 | 76.6 | 77.4 | –0.2 | 0.98 | 0.805 | –0.26 | –0.15 |
| 2: Well | 19.2 | 1.50 | 20.2 | 16.8 | 1.0 | 0.95 | 0.287 | 5.21 | 0.67 |
| 3: Not well, not at all | 3.9 | 0.60 | 3.2 | 5.8 | –0.8 | 0.42 | 0.074 | –20.51 | –1.33 |
| BQ1155: With regard to the English language, how well do you... write it? | | | | | | | | | |
| 1: Very well | 57.4 | 1.75 | 59.3 | 52.9 | 1.9 | 1.21 | 0.129 | 3.31 | 1.09 |
| 2: Well | 30.0 | 1.79 | 29.1 | 32.0 | –0.9 | 1.14 | 0.454 | –3.00 | –0.50 |
| 3: Not well | 9.4 | 0.91 | 8.9 | 10.4 | –0.4 | 0.53 | 0.411 | –4.26 | –0.44 |
| 4: Not at all | 3.3 | 0.54 | 2.7 | 4.7 | –0.6 | 0.38 | 0.145 | –18.18 | –1.11 |
| BQ1208: Since your most recent admission to prison, have you completed any additional education? | | | | | | | | | |
| 1: Yes | 35.0 | 1.53 | 35.2 | 34.7 | 0.1 | 0.92 | 0.890 | 0.29 | 0.07 |
| 2: No | 65.0 | 1.53 | 64.8 | 65.3 | –0.1 | 0.92 | 0.890 | –0.15 | –0.07 |
| BQ1245: Have you ever taken part in a program other than in regular school in order to improve your basic skills, that is, basic reading, writing, and arithmetic skills? The program may have been in prison or it may have been outside of prison. | | | | | | | | | |
| 1: Yes | 24.7 | 1.64 | 24.7 | 24.8 | 0.0 | 0.67 | 0.962 | 0.00 | 0.00 |
| 2: No | 75.3 | 1.64 | 75.3 | 75.2 | 0.0 | 0.67 | 0.962 | 0.00 | 0.00 |
| BQ1500: Have you ever been placed on probation, either as a juvenile or as an adult? | | | | | | | | | |
| 1: Yes | 74.8 | 1.40 | 74.5 | 75.4 | –0.3 | 0.81 | 0.758 | –0.40 | –0.21 |
| 2: No | 25.2 | 1.40 | 25.5 | 24.6 | 0.3 | 0.81 | 0.758 | 1.19 | 0.21 |
| BQ1560: In the year before your incarceration on {BGQ1490}, did you receive income from... unemployment insurance compensation and/ or workman's compensation? | | | | | | | | | |
| 1: Yes | 6.4 | 0.65 | 7.1 | 4.6 | 0.7 | 0.45 | 0.106 | 10.94 | 1.08 |
| 2: No | 93.6 | 0.65 | 92.9 | 95.4 | –0.7 | 0.45 | 0.106 | –0.75 | –1.08 |
| BQ1785: Do you ever use a computer? | | | | | | | | | |
| 1: Yes | 25.0 | 1.73 | 27.1 | 19.9 | 2.1 | 0.80 | 0.011* | 8.40 | 1.21 |
| 2: No | 75.0 | 1.73 | 72.9 | 80.1 | –2.1 | 0.80 | 0.011* | –2.80 | –1.21 |
| BQ1855: Have you ever received… food stamps? | | | | | | | | | |
| 1: Yes | 28.7 | 1.48 | 27.0 | 32.7 | –1.7 | 0.81 | 0.044 | –5.92 | –1.15 |
| 2: No | 71.3 | 1.48 | 73.0 | 67.3 | 1.7 | 0.81 | 0.044 | 2.38 | 1.15 |
| BQ1945: In general, how would you rate your overall health? Would you say it is… | | | | | | | | | |
| 1: Excellent | 28.0 | 1.66 | 29.5 | 24.5 | 1.5 | 0.77 | 0.063 | 5.36 | 0.90 |
| 2: Very Good | 35.0 | 1.79 | 35.5 | 33.8 | 0.5 | 1.06 | 0.648 | 1.43 | 0.28 |
| 3: Good | 21.7 | 1.18 | 21.4 | 22.4 | –0.3 | 0.80 | 0.729 | –1.38 | –0.25 |
| 4: Fair | 11.1 | 0.88 | 10.0 | 13.6 | –1.1 | 0.43 | 0.016 | –9.91 | –1.25 |
| 5: Poor | 4.2 | 0.51 | 3.6 | 5.6 | –0.6 | 0.35 | 0.087 | –14.29 | –1.18 |

* Statistically significant with Bonferroni adjustment at α = 0.05.

[1] Standard error.

[2] Includes non-Hispanic White, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races.

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

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

For father's education, there are also exceptions among the statistically significant bias estimates, namely, 3.5 percent absolute bias for race/ethnicity other than Black or Hispanic (9.64 percent relative bias), –2.9 percent for inmates with low educational attainment (–7.73 percent relative bias), and 2.9 percent for inmates with high educational attainment (4.64 percent relative bias), as well as for inmates who either use (2.1 percent, 8.40 percent relative bias) or do not ever use (–2.1 percent, –2.8 percent relative bias) a computer.

### 11.2.3.3 Multivariate Analysis

The bivariate analysis described in section 11.2.3.2 is useful for explaining each domain variable individually. A multivariate analysis is useful in showing relationships among a number of domain variables for the item (BQ1830, BQ1840) response indicator.

For the multivariate analysis of item nonresponse, CHAID was used to divide the sample into subgroups that best explain differential response rates. The resulting classification trees reveal the survey domains, as defined by combinations of variables with the most differential response rates, thereby leading to survey domains with the highest potential for nonresponse bias. Item response status was used as the dependent variable, and the survey domains indicated in table 11-29 were the predictors.

The trees for the low response rate items are displayed in figures 11-3 and 11-4. For mother's education, the variables used to form the CHAID cells were education, region, "In general, how would you rate your overall health?" and race/ethnicity indicators (figure 11-3). Inmates with low levels of education in the Northeast, South, and West show the greatest potential for nonresponse bias for this item because those inmates have the lowest response rate (69.5 percent). Inmates who had high levels of education, considered themselves to be in excellent or very good health, and were non-Hispanic had the lowest potential for nonresponse bias for this item, with the highest response rate (91.0 percent).

For father's education, the variables race/ethnicity, "Have you ever received food stamps?" "How well do you understand English when it is spoken to you?" "In general, how would you rate your overall health?" "How well do you write English?" and marital status were used to form the response cells (figure 11-4). The last three variables were not significant when considered individually (as shown in table 11-29). However, through interactions with other explanatory variables, they become important

**Figure 11-3. Prison Study multivariate analysis of item nonresponse to BQ1830: What was the highest level of education your mother (stepmother, or female guardian) completed?: 2003**

| Education | | Region / BQ1945: Overall health | | Race/ethnicity | | |
|---|---|---|---|---|---|---|
| Less than high school | 71.5% (458) | Region: Northeast, South, West | 69.5% (395) | | | Cell 1 |
| | | Region: Midwest | 84.7% (63) | | | Cell 2 |
| High school or higher | 86.1% (703) | BQ1945: Overall health: Excellent, very good | 88.8% (451) | Hispanic | 76.5% (72) | Cell 3 |
| | | | | Non-Hispanic Black only, other[1] | 91.0% (379) | Cell 4 |
| | | Good, fair, poor | 81.2% (252) | | | Cell 5 |

Overall weighted response rate = 80.6 percent
Total number of eligibles = 1,161

[1]Includes non-Hispanic White, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races.
NOTE: All percentages are weighted response rates and the numbers inside the parentheses are sample sizes.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Figure 11-4. Prison Study multivariate analysis of item nonresponse to BQ1840: What was the highest level of education your father (stepfather, or male guardian) completed?: 2003**

Overall weighted response rate = 70.7 percent
Total number of eligibles = 1,161

| Race/ethnicity | | BQ1140: Understand English / BQ1945: Overall health | | BQ1155: Write English / Marital status / BQ1855: Received food stamps | | |
|---|---|---|---|---|---|---|
| Hispanic, non-Hispanic Black only | 66.8% (714) | BQ1140: Understand English | | BQ1155: Write English | | |
| | | Very well | 64.7% (520) | Very well | 69.0% (376) | Cell 1 |
| | | | | Well, not well, not at all | 53.5% (144) | Cell 2 |
| | | Well | 77.0% (150) | Marital status | | |
| | | | | Other | 67.4% (53) | Cell 3 |
| | | | | Never married | 82.3% (97) | Cell 4 |
| | | Not well, not at all | 58.4% (44) | Cell 5 | | |
| Other | 77.4% (447) | BQ1945: Overall health | | | | |
| | | Excellent, very good, good | 80.8% (361) | Cell 6 | | |
| | | Fair, poor | 62.4% (86) | BQ1855: Received food stamps | | |
| | | | | Yes | 48.6% (40) | Cell 7 |
| | | | | No | 74.5% (46) | Cell 8 |

NOTE: All percentages are weighted response rates and the numbers inside the parentheses are sample sizes.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

variables. Inmates who are of other than Black or Hispanic race/ethnicity, are in fair or poor health, and have received food stamps showed the greatest potential for nonresponse bias for this item, with the lowest response rate (48.6 percent). In contrast, Hispanic and non-Hispanic Black inmates who understand English well and have never been married had the lowest potential for nonresponse bias for this item, with the highest response rate (82.3 percent).

### 11.2.4 Conclusion

Through the systematic analysis described above, several prison study variables that were known for both respondents and nonrespondents were analyzed to see whether there was an important impact of nonresponse bias. At the unit level, concern over nonresponse is minimal because the response rates were acceptably high for each of the data collection stages. Out of the hundreds of background questionnaire items, only two were below NCES standards for response rates.

For the background questionnaire item asking for mother's education, the CHAID analysis shows that the domain with the lowest response rate (69.5 percent) includes inmates with less than a high school education who live in the Northeast, South, or West. Other important domains that may affect the potential for bias, as given in the bias estimates, are Hispanic race/ethnicity and inmate education level. Some other domains showed some significant relationship with the response indicator for mother's education but did not show important levels of bias based on the bias ratio.

For the background questionnaire item asking for father's education, the CHAID analysis shows that the domain with the lowest response rate (48.6 percent) was inmates of non-Black, non-Hispanic race/ethnicity who rated their health as fair or poor and who had received food stamps. Other important domains that may affect the potential for bias as given in the bias estimates are race/ethnicity other than Black or Hispanic, inmate education levels, and levels of computer use. Some domains given in the CHAID tree include not being able to understand English well or not at all, not being able to write English well, and having a marital status other than never married; however, the impact of these CHAID domains is dependent on other variables, as shown in figure 11-4. Some other domains showed some significant relationship with the response indicator for father's education but did not show important levels of bias based on the bias ratio.

The results of the prison study nonresponse bias analysis show very little potential for nonresponse bias. In fact, there is minimal concern for unit-level bias, and there is concern for only two background questionnaire items. Some caution should be used when analyzing parent's education with the variables that showed importance in the CHAID analysis and in the bivariate analysis.

*This page is intentionally left blank.*

# CHAPTER 12

# WEIGHTING AND VARIANCE ESTIMATION

Thomas Krenzke and Leyla Mohadjer, Westat

The National Assessment of Adult Literacy (NAAL) sample includes both a household component and a prison component. The household component includes two sets of household samples: (1) a national NAAL household sample and (2) household samples from six states, used to administer the State Assessment of Adult Literacy (SAAL). A prison component, involving a sample of adult inmates in federal and state prisons, was conducted to improve the representation of the target population. The complex sample design involved variable sampling rates, stratification, and several stages of selection. To make valid inferences from the responding adults to the target population, the sample must be weighted to account for the special sample design features as well as other complexities arising from nonresponse. In addition, simple formulas (that assume simple random sampling) for variance estimation are not appropriate. Even if sampling weights are used to construct the survey estimates, inferences will not be valid unless the corresponding variance estimator appropriately reflects all the complex features of the sample design. The complex weighting procedures were used to combine the national and state household samples and the prison samples, account for oversampling, and reduce the bias due to nonresponse.

This chapter is divided into two major subsections. The first, section 12.1, discusses the weighting and variance estimation procedures for the NAAL and SAAL household samples. The second, section 12.2, describes the weighting and variance estimation procedures for the correctional institution sample, referred to here as the prison study sample.

## 12.1 HOUSEHOLD SAMPLES

Differential probabilities of selection were adjusted by computing base weights for all adults selected into the household samples. The base weight was calculated as the reciprocal of a respondent's final probability of selection. Further, to combine the NAAL and SAAL household samples, composite weights were calculated for the respondents in the six participating states and the respondents in the national NAAL household sample located within the six SAAL states. Finally, to adjust for nonresponse, weights were adjusted through poststratification and raking to match the 2003 Current Population Survey (CPS) data. The remainder of this section provides detailed information on the weighting and variance estimation procedures used for the household samples.

This section begins by describing the preliminary steps in weighting the household samples (section 12.1.1). The first steps involved computing base weights and nonresponse adjustments for the dwelling units selected for screening (section 12.1.2). Once the screener weighting steps were processed for the NAAL sample and each of the six SAAL states, weighting steps began for the sample persons selected for the background questionnaire. The background questionnaire weighting steps were done separately for each NAAL and SAAL household sample and involved computing base weights, making nonresponse adjustments, and trimming the weights (section 12.1.3). Before compositing, the household sample weights were calibrated to known population estimates (section 12.1.4). After calibration, in order to combine the NAAL and SAAL household samples into one sample, the weights were composited (section 12.1.5). The composited weights were adjusted using a raking procedure, as described in section 12.1.6. Finally, replicate weights were created using the stratified jackknife method, as described in section 12.1.7.

Sample weights were produced for sample persons who either completed the background questionnaire or could not complete the background questionnaire owing to language problems or mental disabilities. The purpose of calculating sample weights was to permit inferences from sample persons to the populations from which they were drawn and to have the tabulations reflect estimates of the population totals. The sample weighting process was designed to accomplish the following objectives:

1. Permit unbiased estimates, taking into account the fact that all persons in the population did not have the same probability of selection;

2. Minimize the potential bias arising from differences between respondents and nonrespondents;

3. Combine the state and national samples in an efficient manner;

4. Use auxiliary data on known population characteristics in such a way as to reduce sampling errors and to bring data up to the dimensions of the population totals;

5. Reduce the variation of the weights and prevent a small number of observations from dominating domain estimates; and

6. Facilitate sampling error estimation under complex sample designs.

Objective 1 was accomplished by computing base weights for the households selected for screening and, subsequently, for persons selected for the background questionnaire and assessment from the eligible participating households. The details of the base weight calculations for the screener and the background questionnaire are presented in sections 12.1.2.1 and 12.1.3.1, respectively.

Objective 2 was accomplished through nonresponse weighting adjustments that accounted for screener nonresponse and background questionnaire nonresponse. Sections 12.1.2.2 and 12.1.3.2 discuss the nonresponse adjustments for the screener and background questionnaire, respectively. Some reduction in potential bias was also achieved while meeting Objective 4 by calibrating the weights. This was accomplished by using weighting variables that were not used for nonresponse adjustment because data were available only for respondents.

Objective 3 was addressed through the composite weighting procedure. Composite weights were computed for the respondents in the six state samples and the respondents in the national sample primary sampling units (PSUs) of those six states. Area sampling procedures that included stratification, PSU formation, sample design, and selection at the various stages of sampling were applied to the national and state components. Further, the same instruments were used to screen households and to collect background information and literacy assessment data in the state and national surveys. To take full advantage of this comparability, the samples were combined to produce both state- and national-level statistics. Section 12.1.5 describes the composite estimation procedures.

To meet Objective 4, the weights were calibrated to known totals from the 2003 March Supplement of the CPS.[1] The weights were raked so that numerous totals calculated with the resulting full-sample weights would agree with the CPS totals. Calibration procedures were implemented for both the national sample within each state and the state sample prior to compositing the weights. After the weights had been composited, another raking process was conducted to rescale the weights. The calibration procedures are described in sections 12.1.4 and 12.1.6.2.

Objective 5 was addressed by trimming the weights. A small number of weights were reduced using a type of inspection approach (referred to as the k x median rule) within prespecified sampling and analytical domains. The trimming procedure was implemented twice during the weighting process, once before compositing the weights and once after compositing. For more discussion of the trimming procedure, refer to sections 12.1.3.3 and 12.1.6.2.2.

Finally, Objective 6 was accomplished by creating 61 replicate weights using the stratified jackknife method. The NCES standards ask for the number of replicates to be greater than 29 and less than 101. Full-sample and replicate weights were calculated for each record to facilitate the computation

[1] The March CPS supplement is an annual survey, conducted by the Bureau of Labor Statistics and the U.S. Census Bureau, to collect detailed information on demographics, income, and work experience.

of unbiased estimates and their standard errors. The weighting procedures were repeated for 61 strategically constructed subsets of the sample to create a set of replicate weights for variance estimation using the jackknife method. The replication scheme was designed to produce stable estimates of standard errors for the national and six individual state estimates. The replication design and the significance of the number of replicates is discussed further in section 12.1.7. The variance strata and variance units created for the replication process can also be used in estimating sampling error using Taylor series approximation (Wolter 1985).

Prior to the weighting process, it was necessary to resolve any issues related to the data used in weighting. The next section discusses this preliminary data cleaning procedure.

### 12.1.1 Preliminary Steps in Weighting

The data used in the weighting process underwent consistency checks to prevent any errors in the sample weights. The checks were performed only on variables required for weighting and were limited to records that required weights.

The consistency checks also helped identify any unusual values. Westat prepared listings of records with missing values in any of the weighting variables. The listings showed the following variables: the respondent’s case identification (ID) number, age, date of birth, gender, race/ethnicity, country of birth, and level of education; race of the head of household; and the number of age-eligible members and respondents in the household. The printed listings were used to review the extent of missing data, identify the pattern of missing data, and prepare for imputation. The age, gender, and race/ethnicity data from the screener and the background questionnaire were also compared for consistency. Inconsistency in survey data is due to individuals reporting data for others in the screener. In all, less than 1 percent had missing data or inconsistent data between the screener and background questionnaire for these items.

The weighting variables that were at a finer level of detail than was necessary for the later steps of weighting (age, gender, race/ethnicity, country of birth, and level of education) were recoded (i.e., collapsed to the required levels). Age, race/ethnicity, and gender were collected in both the screener and the background questionnaire, thereby providing two measures of the same item. The background questionnaire measure was preferred for all items. For the few cases in which the background questionnaire measure was missing, the screener value was used as a direct substitute.

For level of education and country of birth, which were not collected through the screener, a limited amount of imputation was performed to fill in the data for respondents so that the variables could be used in the calibration and raking processes. To the extent possible, missing values were filled in using information from other items in the background questionnaire. For example, if the data on country of birth were missing, the questions regarding length of residence in the United States, education obtained before coming to the United States, and language the respondent first learned to read or write were consulted to determine whether the respondent was born in the United States. Similarly, several education variables were used to create a three-level education measure (less than high school, high school or equivalent, or more than high school) if this variable had missing data.

If no other background questionnaire data were available for imputing these items, and since there were a small number of missing values remaining, a simple imputation procedure was performed as follows.[2] Two cases still had missing values for the "born in the U.S." item. To obtain values for these cases, cells were formed by PSU and segment. Then the most frequent value in the cell was given to the missing case (i.e., modal within cell hotdeck[3]). For the seven remaining cases with missing education data, cells were formed by PSU, age (16–19, 20–29, 30–69, 70+), and race/ethnicity. Again, the most frequent value for education in the cell was given to the missing case.

Some additional dwelling units came into the sample as a result of the missed structure and hidden dwelling unit procedures (refer to section 7.1.3.5 for more information), which allowed units that were missed in the segment listing activities to be included in the sample with a known probability of selection. All newly discovered dwelling units within a segment were included unless the total number was unusually large, in which case a sample of newly discovered dwelling units was taken. Whenever a sample of missed units was selected, detailed records indicated the PSU, segment, number of new dwelling units selected, and total number of newly discovered dwelling units. This information was attached to each of these records prior to the calculation of base weights.

A few final checks were run (refer to section 12.1.6.3 for further discussion) before the screener base weights were calculated to ensure the availability and validity of all fields required by the base weights program (fields created for the special cases mentioned above and fields for the total number of age-eligible household members and the number of sample persons for each dwelling unit). A detailed description of the screener base weight computation is provided in the next section.

[2] For the 355 nonrespondents who did not complete the survey because of language problems or mental disability, the imputation method was more complex. Details are provided in section 12.1.3.2.2.

[3] Hotdeck is an imputation procedure that uses data from the same sample survey.

### 12.1.2 Screener Base Weights and Nonresponse Adjustments

To produce unbiased estimates, differential weights must be used for various subsets of the population whenever subsets have been sampled at different rates. Weighting was required to account for the oversampling of Blacks and Hispanics in high-minority segments of the national sample, as discussed in section 7.1.3.3. The screener data helped determine the probabilities of selection for the screener. Section 12.1.2.1 summarizes the base weight computation for the household samples.

If every selected household had agreed to complete the screener and every selected person had agreed to complete the background questionnaire and the assessment booklet, weighted estimates based on the data would be approximately unbiased (from a sampling point of view). However, nonresponse occurs in any survey operation, even when participation is mandatory, and adjustments are always necessary to avoid potential nonresponse bias. The weighting adjustments for screener nonresponse are discussed in section 12.1.2.2.

#### 12.1.2.1 Screener Base Weights

The probability of a dwelling unit k being selected into NAAL or SAAL, denoted as Pijk(mdu) (as given in table 7-8), is the product of the conditional probabilities at the PSU, segment, and dwelling unit levels. Other factors entering into the probability of selection were due to chunking (refer to section 7.1.2.3), dwelling unit selection from segments selected for both NAAL and SAAL (refer to section 7.1.3.3), missed dwelling units identified through the missed structure process (refer to section 7.1.3.5), and subsampling of nonminority dwelling units in oversampled high-minority segments (refer to section 7.1.3.3). The screener base weights were computed as the reciprocal of the probability of selection of dwelling unit k of PSU i and segment j, after accounting for subsampling due to the missed dwelling units (mdu) procedure, as shown in the following formula:

$$W_{ijk}^{base,SCR} = \frac{1}{P_{ijk(mdu)}}.$$

Table 12-1 shows the distribution of the screener base weights for the NAAL sample and for each of the SAAL samples. The variation—as seen by the minimum, maximum, and coefficient of variation—can be explained by several factors. These factors include oversampling of Blacks and Hispanics, sampling of missed dwelling units in segments where a large number of dwelling units were found by the interviewers as they canvassed the listing area, and a small number of unique sampling situations. The table also indicates that the coefficient of variation is much lower for the SAAL states than for the national NAAL sample due to an equal probability design for households.

**Table 12-1. Screener base weight distribution for the household samples, by sample: 2003**

| Sample | Screener base weights | | | | |
|---|---|---|---|---|---|
| | Sample cases | Median | Minimum | Maximum | Coefficient of variation (percent)[1] |
| NAAL | 25,450 | 7,240 | 1,207 | 21,719 | 45 |
| SAAL | | | | | |
| Kentucky | 2,306 | 771 | 386 | 1,542 | 6 |
| Maryland | 1,493 | 1,528 | 764 | 2,292 | 11 |
| Massachusetts | 1,509 | 1,750 | 875 | 3,499 | 8 |
| Missouri | 1,499 | 1,658 | 829 | 2,487 | 5 |
| New York | 1,499 | 5,151 | 2,575 | 5,151 | 2 |
| Oklahoma | 1,609 | 992 | 496 | 2,975 | 27 |

[1] The coefficient of variation is the standard deviation of the weights divided by the mean weight.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 12.1.2.2 Screener Nonresponse Adjustment

For the screener nonresponse adjustment, the nonrespondents were divided into two categories. The first category consisted of cases involving nonliteracy-related nonresponse, such as refusals and nonresponse because of illness. Nonliteracy-related nonrespondents were likely to be similar to respondents with respect to English literacy scores. The second category was literacy-related nonresponse. Language problems are the only type of literacy-related nonresponse at the screener level, with only 160 such cases in the NAAL and SAAL household samples. Households with this type of nonresponse were presumed to differ from responding households with respect to literacy. Therefore, the weighting procedures adjusted the weights of the respondents to represent the nonliteracy-related nonresponse only. The weights of the language problem cases were not adjusted during the screener-level nonresponse adjustment because their literacy status was expected to differ from that of respondents. The contribution of the screener-level literacy-related nonresponse to the total population was accounted for by literacy-related nonresponse adjustment carried out for the background questionnaire sample (refer to section 12.1.3.2.2).

Little was known about the nonresponding households, including their eligibility.[4] Before any nonresponse adjustment was processed, an adjustment for unknown eligibility was performed. In this step, the weights of the households with unknown eligibility status, such as those with maximum callbacks, were distributed among the cases with known eligibility status. The second step distributed the weights of the eligible nonrespondents among the eligible respondents.

All adjustments were made within weighting classes. Because very little was known about the households that did not respond to the screener, information used to form weighting classes had to come from a different source. The frame contained only aggregate demographic information, such as region and Metropolitan Statistical Area (MSA) status. However, because the sampling was performed using census geography, the sampled segments were merged to the Census 2000 Summary File 3 (SF3)[5] files to create segment-level weighting variables by extracting segment-level census data.

Prior to the weight adjustments, classification software was used to help identify weighting classes for the adjustments for unknown eligibility status and nonresponse. A Chi-squared Automatic Interaction Detector (CHAID) (Kass 1980) was used to help identify important variables to be used in forming weighting classes that were homogeneous in terms of response propensity. CHAID is a classification algorithm that divides a population into homogeneous subgroups with respect to a target characteristic (the dependent variable). Once the weighting variables were identified through CHAID, the weighting classes were formed through a hierarchical ordering of the weighting variables. The hierarchical ordering was formed using the general order as they were selected for the CHAID tree classification. Table 12-2 shows the variables selected to form the weighting classes for the NAAL and SAAL household screener samples based on the CHAID results. Weighting classes were combined if the cell size was less than 30 or the adjustment factor was greater than 1.50. The criteria for cell size and maximum adjustment factor is a guideline and can vary from survey to survey, and by weighting stage within a survey (Kalton and Kasprzyk 1986).

---

[4] Households were ineligible only if they were vacant or were not a residential dwelling unit.

[5] The SF3 files contain data from the 52-item census long form that was issued to about 19 million households. The files contain data on demographics, education, income, commuting, and other characteristics.

For each weighting cell α, the screener unknown eligibility adjustment factor is computed as follows:

$$F_{\alpha}^{unk,SCR} = \frac{\sum_{k \in S(\alpha)} W_{ijk}^{base,SCR}}{\sum_{k \in SK(\alpha)} W_{ijk}^{base,SCR}},$$

where

*S(α)* = the set of sampled cases (i.e., STATUS = 0, 1, 2, 3, or 4) in weighting cell *α* and

*SK(α)* = the set of sampled cases with known eligibility status (i.e., STATUS = 0, 1, 2, or 3) in weighting cell *α*

and where

STATUS = 0 literacy-related nonrespondents (language problems only);

1 respondents;

2 nonrespondents known to be eligible, including respondents who refused and those unavailable due to illness;

3 ineligibles, including households subsampled out (those with nonminority reference persons in high-minority segments), vacancies, and sampled cases that were not dwelling units; and

4 cases for which the eligibility status was not known.

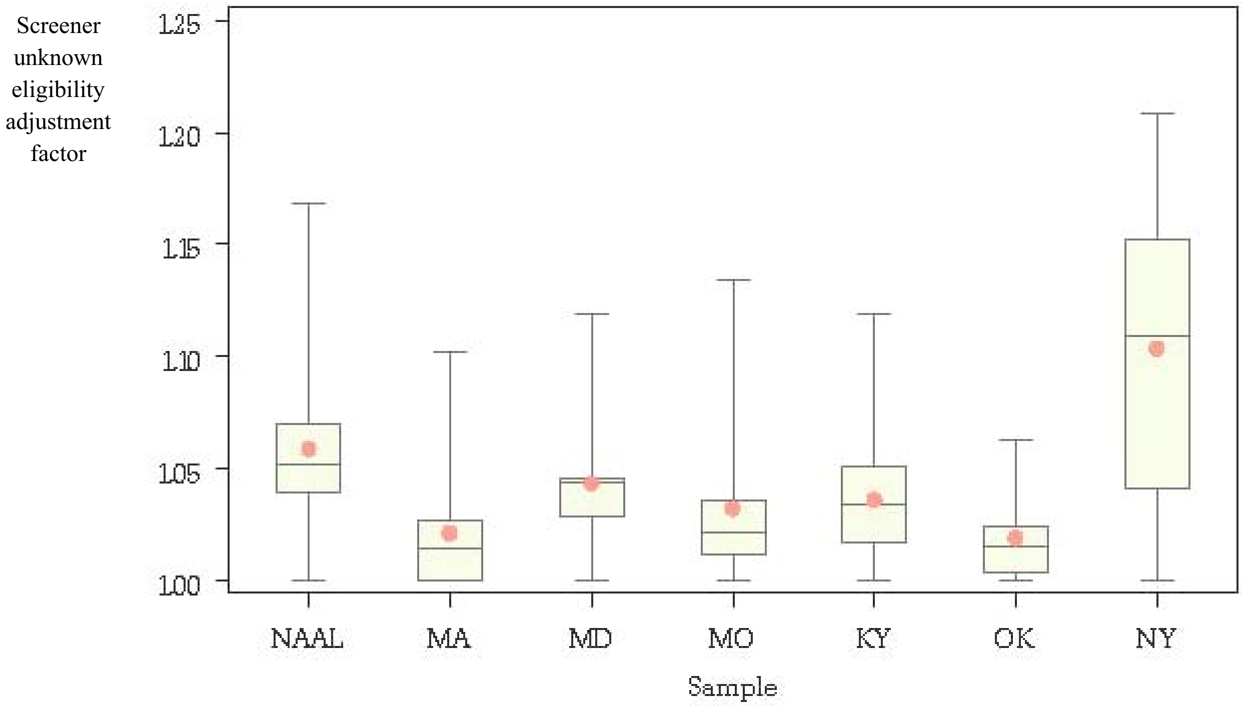
**Table 12-2. Variables used in forming weighting classes for the screener nonresponse adjustment, by sample: 2003**

| Sample | Variables |
|---|---|
| NAAL | Indicator that percentage of Black or Hispanic population in segment exceeds 12.5 percent<br>Percentage of segment population who do not speak English at home but speak English well<br>Percentage of segment population below 150 percent of poverty<br>Census region |
| SAAL | |
| Kentucky | Median household income in segment<br>Percentage of segment population who speak Spanish at home but do not speak English well or at all<br>Percentage of segment population with less than a high school education |
| Maryland | Percentage of segment population with a high school education<br>Percentage of segment population who speak a language other than Spanish or English at home but speak English well<br>MSA status of PSU<br>Percentage of segment population below 150 percent of poverty |
| Massachusetts | Indicator that percentage of Black or Hispanic population in segment exceeds 12.5 percent<br>Percentage of segment population with more than a high school education but less than a bachelor's degree<br>Median household income in segment |
| Missouri | Whether the segment is in an urban area<br>Median household income in segment<br>Percentage of segment population with a bachelor's degree |
| New York | Percentage of segment population who speak English only<br>Percentage of segment population with a bachelor's degree or higher |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

The distribution of the adjustment factors for cases with known eligibility in the NAAL and SAAL samples is shown in figure 12-1. The figure displays a box-and-whisker plot that shows the median (the horizontal line inside the box), the mean (the dot inside the box), the 25th and 75th percentiles (bottom and top of box, repectively), and the minimum and maximum values (the end of the line below the box, and the end of the line above the box, respectively). The figure shows that the adjustment factors for the national NAAL sample range from 1.0 to about 1.7. New York's SAAL sample has the largest range (from 1.0 to 1.2). The other SAAL states' average adjustment factors are small (less than 1.05).

**Figure 12-1. Distribution of the unknown eligibility adjustment factors for the household samples, by sample: 2003**




SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy

Subsequently, the screener nonresponse adjustment factor was computed in the following way:

$$F_{\alpha}^{nr,SCR} = \frac{\sum_{k \in SE(\alpha)} W_{ijk}^{base,SCR} F_{\alpha}^{unk,SCR}}{\sum_{k \in SC(\alpha)} W_{ijk}^{base,SCR} F_{\alpha}^{unk,SCR}},$$

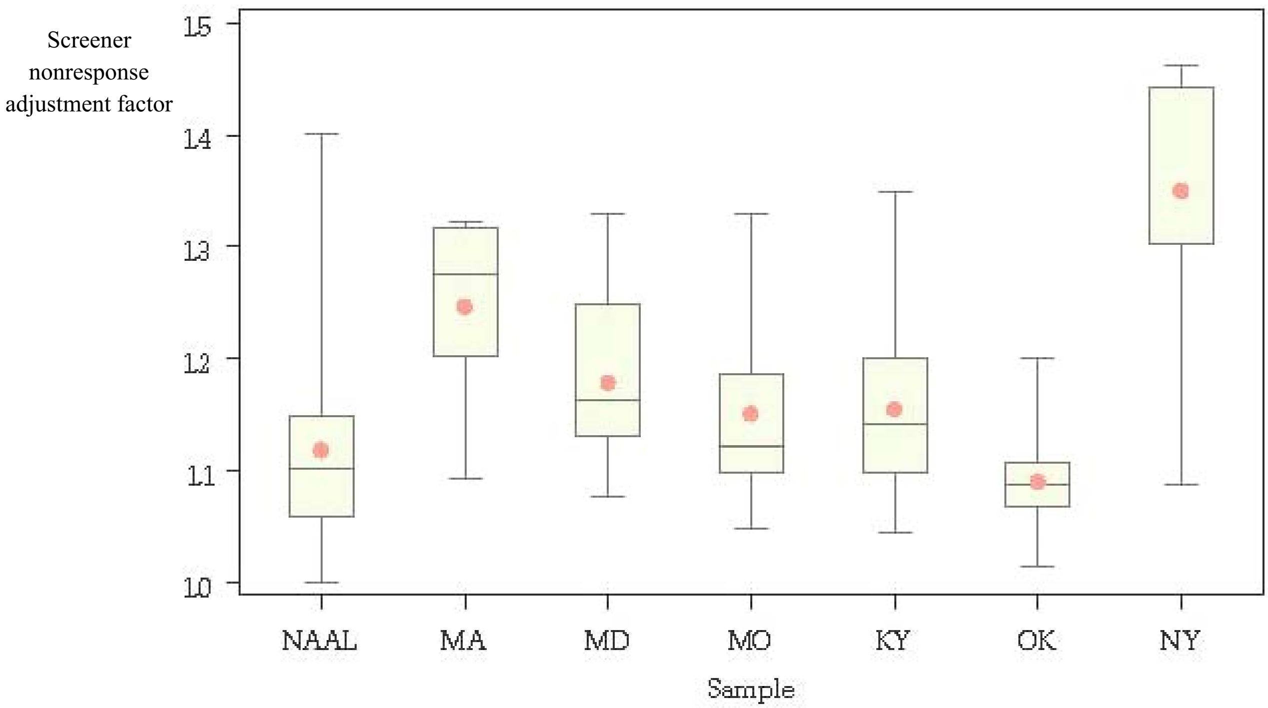
where

*SE(α)* = *the set of eligible sampled cases (i.e., STATUS category 1 or 2) in weighting cell α and*

*SC(α)* = *the set of completed cases (i.e., STATUS category 1) in weighting cell α.*

For simplicity, the notation assumes the same cells as used in the unknown eligibility adjustment, when in fact the cell definitions changed as a result of cell collapsing when the number of respondents was less than 30 or adjustment factors were greater than 1.50. The distribution of the screener nonresponse adjustment factors for screener respondents in the household samples is shown in figure 12-2. The figure shows that New York's SAAL sample had relatively high adjustment factors on average due to its relatively low screener response rates. The national NAAL sample's adjustment factors range from 1.0 to about 1.4.

**Figure 12-2. Distribution of the screener nonresponse adjustment factors for the household samples, by sample: 2003**




SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy

The adjustment $F_{\alpha}^{nr,SCR}$ was applied only to the unknown eligibility-adjusted weights of the screener completes (i.e., STATUS category 1). That is, the nonresponse-adjusted weight, $W_{ijk}^{nr,SCR}$, was computed as follows:

$W_{ijk}^{nr,SCR} = W_{ijk}^{base,SCR} F_{\alpha}^{unk,SCR}$, if dwelling unit k was a literacy-related nonrespondent (STATUS category 0);

= $W_{ijk}^{base,SCR} F_{\alpha}^{unk,SCR} F_{\alpha}^{nr,SCR}$, if dwelling unit k was a respondent (STATUS category 1); and

= 0, if dwelling unit k was a nonrespondent, ineligible, or of unknown eligibility status (STATUS category 2, 3, or 4).

### 12.1.3 Background Questionnaire Base Weights, Nonresponse Adjustments, and Trimming

The derivation of base weights was necessary to prevent potentially serious biases in the outcome statistics. The study specifications called for the selection of one person in households with fewer than four eligible members and two persons in households with four or more eligible members. Members of households with only one eligible member had twice the chance of selection as those in households with two (or four) eligible members. To produce unbiased estimates, different weights had to be used to account for the within-household selection rate. Section 12.1.3.1 summarizes the base weight computation for the background questionnaire sample, section 12.1.3.2 presents the background questionnaire nonresponse adjustment procedures, and section 12.1.3.3 describes the trimming procedure used to reduce the impact of extreme weights.

#### 12.1.3.1 Background Questionnaire Base Weights

The background questionnaire base weights were computed as the product of the screener nonresponse-adjusted weight and the reciprocal of the within-household probability of selection for person l within household k of PSU i and segment j, as shown in the following formula:

$$W_{ijkl}^{base,BQ} = W_{ijkl}^{nr,SCR} \frac{1}{CP_{ijkl}},$$

where

$CP_{ijkl}$ = the within-household probability of person l being selected into NAAL or SAAL, which is the ratio of the number of persons selected in household k to the number of eligible persons in household k.

Table 12-3 shows the distribution of the background questionnaire base weights for NAAL and for each of the SAAL household samples. The oversampling of Blacks and Hispanics resulted in a larger coefficient of variation for the national NAAL household sample, as expected. Other major reasons for the variation in the sampling weights include sampling of missed dwelling units in segments where a large number of dwelling units were discovered by the interviewers as they canvassed the listing area, the number of eligible persons in the household, and the screener nonresponse adjustment factors.

**Table 12-3. Distribution of the background questionnaire base weights for household samples, by sample: 2003**

| Sample | Number of sample persons | Median | Minimum | Maximum | Coefficient of variation (percent) |
|---|---|---|---|---|---|
| NAAL | 16,409 | 8,797.33 | 1,283.36 | 87,353.59 | 63.09 |
| | | | | | |
| SAAL | | | | | |
| Kentucky | 1,694 | 1,766.13 | 815.15 | 3,301.97 | 35.58 |
| Maryland | 1,290 | 3,719.97 | 1,652.13 | 8,369.14 | 37.26 |
| Massachusetts | 1,116 | 4,495.24 | 1,070.26 | 13,714.16 | 33.69 |
| Missouri | 1,368 | 3,703.51 | 1,739.19 | 7,826.35 | 38.35 |
| New York | 956 | 14,851.10 | 3,824.44 | 26,789.43 | 37.71 |
| Oklahoma | 1,293 | 2,189.19 | 1,016.34 | 6,622.13 | 38.66 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 12.1.3.2 Background Questionnaire Nonresponse Adjustment

#### 12.1.3.2.1 Nonliteracy-Related Nonresponse

At the background questionnaire level, separate adjustments were made for the literacy-related nonrespondents and the other nonrespondents. This section discusses weighting adjustments for nonliteracy-related nonresponse. For the household samples, the variables available for nonresponse adjustments for the background questionnaire included variables from the Census 2000 SF3 file and screener variables (region, age, race/ethnicity, and gender). The weighting variables used in the screener nonresponse adjustment were also considered for the background questionnaire adjustment.

The sample persons were classified into the following STATUS groups:

STATUS = 0 literacy-related nonrespondents (language problems only);

1 respondents; and

2 nonliteracy-related nonrespondents.

The classification software, CHAID, was used to identify the key variables to use to form the weighting classes. Table 12-4 shows the variables selected to form the weighting classes for the household samples based on the CHAID analysis. More discussion on this approach is given in section 12.1.2.2. After the variables were identified, the weighting classes were formed through a hierarchical ordering of the weighting variables.

**Table 12-4. Variables forming background questionnaire nonresponse adjustment weighting classes, by sample: 2003**

| Sample | Variables | Source |
|---|---|---|
| NAAL | Percentage of segment population below 150 percent of poverty | SF3[1] |
| | Percentage of segment population who speak Spanish and speak English well | SF3 |
| | Gender | Screener |
| | Age category | Screener |
| | Household size | Screener |
| SAAL | | |
| Kentucky | Percentage of segment population with less than a high school education | SF3 |
| | Age category | Screener |
| | Gender | Screener |
| Maryland | MSA status of PSU | Sampling files |
| | Age category | Screener |
| | Percentage of segment population who speak another language at home and speak English well | Screener |
| Massachusetts | Median household income in segment | SF3 |
| | Race/ethnicity | Screener |
| | Percentage of segment population with more than a high school education | SF3 |
| Missouri | Age category | Screener |
| | Gender | Screener |
| | Median household income in segment | SF3 |
| New York | Percentage of segment population below 150 percent of poverty | SF3 |
| | Race/ethnicity | Screener |
| | Percentage of segment population with more than a college education | SF3 |
| Oklahoma | Median household income in segment | SF3 |
| | Household size | Screener |
| | Percentage of segment population with more than a high school education | SF3 |

[1] Census 2000 Summary File 3.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

Once the weighting classes had been identified, the nonresponse adjustment factors were computed. Weighting classes were combined if the cell size was less than 30 or the adjustment factor was greater than 1.75. The maximum adjustment of 1.75 is larger than that for the screener adjustment. There is not a fixed rule for the maximum, although the statistician attempts to balance an increase in variance due to large adjustments, with decrease in bias due to nonresponse. Refer to Kalton and Kasprzyk (1986) for more discussion. The corresponding sample-based nonresponse adjustment is defined to be the ratio of sums:

$$F_{\beta}^{nr,BQ} = \frac{\sum_{l \in SN(\beta)} W_{ijkl}^{base,BQ}}{\sum_{l \in SC(\beta)} W_{ijkl}^{base,BQ}},$$

where

*SN(β)* = *the set of completed background questionnaires and nonliteracy-related nonrespondents (STATUS category 1 or 2) in weighting class β and*

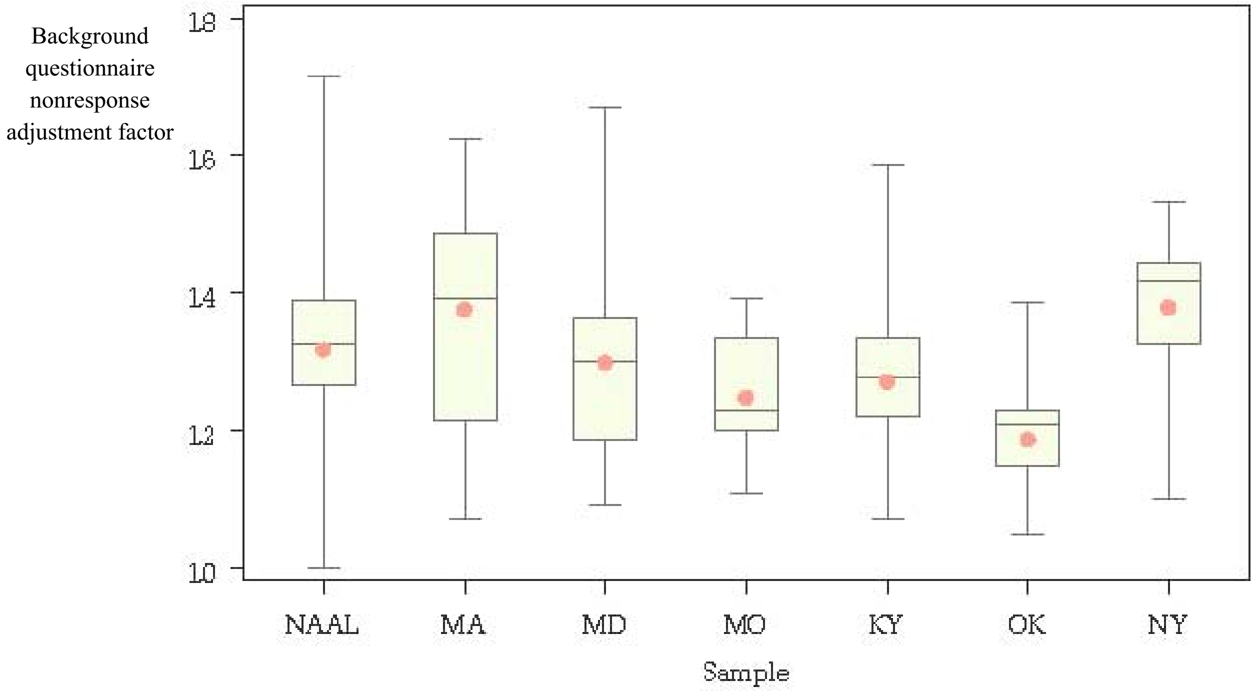
*SC(β)* = *the set of completed background questionnaires (STATUS category 1) in weighting class β.*

The distribution of the background questionnaire nonresponse adjustment factors for screener respondents in the NAAL and SAAL samples is shown in figure 12-3. The figure shows that the national NAAL sample's adjustment factors ranged from 1.0 to about 1.7. Oklahoma had a low adjustment factor on average, at about 1.2.

#### 12.1.3.2.2 Literacy-Related Nonresponse

Of the 355 sample persons who did not complete the background questionnaire for literacy-related reasons, 211 sample persons had language problems and 144 sample persons had mental disabilities as determined by the interviewers and documented in the noninterview reports (refer to chapter 8 for more discussion). These cases were included in the background questionnaire data file along with their age, race, and gender information from the screener. Educational attainment and country of birth, two variables needed for calibrating the weights (section 12.1.4), were imputed using logistic regression models that included segment-level education and poverty data from the Census 2000 SF3 data.

**Figure 12-3. Distribution of the background questionnaire nonresponse adjustment factors for the household samples, by sample: 2003**




SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

Through the logistic regression models, the predicted values (response propensities or probabilities) were generated for each classification of education (or country of birth). The imputed value was assigned through a random draw from the probability distribution, as predicted by the model. This approach is discussed in Thibaudeau et al. (1997).

Before the background questionnaire weights were calibrated, the background questionnaire literacy-related respondent weights were adjusted to account for the 160 literacy-related screener nonrespondents. This adjustment was necessary primarily to allow the literacy-related background questionnaire respondents to represent the literacy-related screener nonrespondents in the calibration procedure. This adjustment assumed that the literacy-related nonrespondents to the screener and the background questionnaire are similar in literacy. The weighting class, β, was simply the national NAAL household sample and each of the six SAAL states. The corresponding sample-based nonresponse adjustment is defined to be the ratio of sums:

$$F_{\beta}^{nr,BQ} = \frac{\sum_{k \in SL(\beta)} W_{ijk}^{base,SCR}}{\sum_{k \in S0(\beta)} W_{ijk}^{base,SCR}},$$

where

$SL(\beta)$= *the set of sample dwelling units with either a literacy-related screener nonresponse or a literacy-related background questionnaire nonrespondent in weighting class β and*

$S0(\beta)$= *the set of literacy-related background questionnaire nonrespondents (STATUS category 0) in weighting class β.*

#### 12.1.3.3 Background Questionnaire Trimming Adjustment

A trimming algorithm was used to reduce the variation in the background questionnaire nonresponse-adjusted weights. Reasons for the variation in the NAAL and SAAL sampling weights include subsampling of newly discovered dwelling units, number of eligible persons in the household, and screener and background questionnaire nonresponse adjustment factors.

In general, trimming procedures introduce some bias into the sampling weights (Lee 1995). However, as Lee discusses, the trimming adjustment in most cases will reduce the sampling error component of the overall mean square error more than it increases the bias when the adjustment is applied to only a very small number of weights. Trimming cells were formed by crossing the high/low-minority segment indicator (defining sampling domains) with a three-category race variable (defining analytical domains). Within each cell, cases that had weights greater than three times the median were considered for having their weights reduced. (This approach is hereinafter referred to as the 3× median rule.) This type of inspection approach, which is very common in survey weighting practices, is discussed in Potter (1990). The trimming factor, denoted by $F_{ijkl}^{trim,BQ}$, was the ratio of the cutoff value to the background questionnaire nonresponse adjustment weight. The trimming factor for the full-sample weights was then applied to the replicate weights (refer to section 12.1.7 for a discussion of replicate weights as they pertain to variance estimation).

For the NAAL household sample, 52 full-sample weights were trimmed. Table 12-5 shows the number of weights trimmed in each cell and the distribution of trimmed weights and trimming factors. All 17 cases requiring trimming in the “other” race category were due to the subsampling of dwelling units found during the missed structure procedure. For the Hispanic and Black trimmed cases, the large background questionnaire nonresponse-adjusted weights were due mainly to large nonresponse adjustment factors. For the SAAL states, no trimming was needed for Kentucky, Maryland, Missouri, New York, or Oklahoma. One case involving missed dwelling unit subsampling was trimmed for Massachusetts.

**Table 12-5. Distribution of trimmed weights and trimming factors, by minority status and race: 2003**

| Minority status of segment | Race | No. of weights trimmed | Trimmed weight | | | | Trimming factors | |
|---|---|---|---|---|---|---|---|---|
| | | | *N* | Mean | Coefficient of variation | Max | Mean | Coefficient of variation |
| Total | | 52 | 12,753 | | | | | |
| Low | Hispanic | 0 | 282 | 22,632.22 | 34.01 | 46,178.90 | 1.0000 | 0.00 |
| Low | Non-Hispanic Black | 0 | 197 | 19,881.10 | 48.39 | 47,995.49 | 1.0000 | 0.00 |
| Low | Other | 15 | 5,915 | 21,971.97 | 39.59 | 68,246.93 | 0.9996 | 0.91 |
| High | Hispanic | 20 | 2,587 | 7,217.94 | 41.25 | 20,578.82 | 0.9987 | 1.71 |
| High | Non-Hispanic Black | 15 | 2,640 | 6,231.19 | 47.07 | 18,582.41 | 0.9986 | 2.27 |
| High | Other | 2 | 1,132 | 17,112.82 | 50.89 | 52,981.88 | 0.9997 | 0.83 |

NOTE: All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as other include non-Hispanics of all other races including multiracial. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 12.1.4 Calibration Adjustments Prior to Compositing

Undercoverage of the target population is a common problem in surveys. Undercoverage occurs when some population units are not included in the sampling frame and have no chance of being selected into the sample. Almost all surveys are subject to some amount of undercoverage, and NAAL and SAAL are no exception. A calibration adjustment to the weights accounted for any undercoverage and balanced the samples within each SAAL state prior to the compositing process. For this step, the entire sample was divided into the NAAL and SAAL sample in the six SAAL states. The NAAL sample in the remaining 44 states and the District of Columbia was excluded from this step because there was no SAAL sample in those states. After compositing, the combined NAAL and SAAL household sample weights were calibrated through a raking adjustment process (refer to section 12.1.6.2).

The creation of the control totals used for the calibration adjustment is discussed in section 12.1.4.1. The calibration adjustments are discussed in section 12.1.4.2.

#### 12.1.4.1 Control Totals

Control totals were computed for the purpose of calibrating the sample weights within the six SAAL states prior to compositing. The totals were computed from the 2003 CPS March Supplement. For each sample, control totals were computed for the following variables: MSA status, age, gender, education, country of birth, race/ethnicity, and national certainty status of the PSU. The number of variables to use was limited because the external source of the control totals needed to have the exact same wording of questions as the NAAL and no missing NAAL responses. Furthermore, not all of these variables were used for each state sample because of small sample sizes in certain domains. However, the variables used for the calibration step were defined to the finest classification that the data allow. Also the effectiveness of calibration methods (raking in particular) depends on the relationship between the auxiliary variables used in calibration and the survey estimates (Brick et al. 2003). Table 12-6 displays the resulting variables involved in the calibration process.

#### 12.1.4.2 Calibration

Calibration is commonly used in sample surveys to reduce the mean square error of estimates and to create consistency with statistics from other studies. However, the primary reason for calibration in the setting of NAAL is to provide a common base for the NAAL and SAAL samples in each of the six SAAL states before applying the composite weighting factors. The trimmed background questionnaire weights for the six states were calibrated to the 2003 CPS March Supplement control totals. Respondents who completed the background questionnaire were included in the calibration. Literacy-related nonrespondents were also included because they are part of the target population from which the control totals were derived. Variables critical to the weighting were recoded and imputed, as necessary, before the calculation of base weights.

**Table 12-6. Variables involved in the calibration process prior to compositing, by sample: 2003**

| Sample[1] | MSA status | Age | Gender[2] | Education[3] | Country of birth | Race/ethnicity | NAAL certainty PSU |
|---|---|---|---|---|---|---|---|
| NAAL Kentucky | | 16–49, 50+ | M, F | <HS, HS, + | | Hispanic/Black, other | |
| NAAL Maryland | | 16–49, 50+ | M, F | <HS/HS, + | | Hispanic/Black, other | |
| NAAL Massachusetts | | 16–49, 50+ | M, F | <HS/HS, + | | Hispanic/Black, other | |
| NAAL Missouri | | 16–49, 50+ | M, F | <HS/HS, + | | | |
| NAAL New York | | | | | | Hispanic/Black, other | Certainty, noncertainty |
| NAAL Oklahoma | | 16–49, 50+ | M, F | <HS, HS, + | | | |
| SAAL Kentucky | MSA, non-MSA | 16–29, 30–49, 50–69, 70+ | M, F | <HS, HS, + | | Hispanic/Black, other | |
| SAAL Maryland | | 16–29, 30–49, 50–69, 70+ | M, F | <HS, HS, + | U.S., other | Hispanic/Black, other | |
| SAAL Massachusetts | | 16–29, 30–49, 50–69, 70+ | M, F | <HS, HS, + | U.S., other | Hispanic/Black, other | |
| SAAL Missouri | MSA, non-MSA | 16–29, 30–49, 50–69, 70+ | M, F | <HS, HS, + | | Hispanic/Black, other | |
| SAAL New York | | | | | | Hispanic/Black, other | |
| SAAL Oklahoma | MSA, non-MSA | 16–29, 30–49, 50–69, 70+ | M, F | <HS, HS, + | | Hispanic/Black, other | |

[1] NAAL XX means the national NAAL sample cases in state XX.
[2] M: males; F: females.
[3] <HS: less than high school; HS: high school diploma or equivalent; +: more than high school; <HS/HS: either less than high school or high school diploma or equivalent.
NOTE: All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as other include non-Hispanics of all other races including multiracial. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

A raking procedure (i.e., iterative poststratification) was used for the calibration. In raking, categories are formed from certain variables and the weights are calibrated to control totals for each category. In some instances, such cross-tabulations may contain sparse cells, or population distributions may be known for the marginal but not the joint distributions for variables used to define the weighting classes. Typically, raking is conducted when the control totals for interior cells of a cross-tabulation are unknown or sample sizes in some cells are too small for efficient estimation. Raking is related to poststratification in that it poststratifies (or calibrates) to marginal population totals of several variables (or raking dimensions) in an iterative manner. Oh and Scheuren (1987) provide a concise description of the raking procedure and its properties.

A raked weight was calculated for each respondent as follows. Let $N_\gamma$ denote the population count in the raking dimension category γ as obtained from the 2003 CPS March Supplement, as discussed in section 12.1.4.1. Let $\hat{N}_\gamma$ be the corresponding survey estimate obtained by using the survey weights prior to raking (as calculated below):

$$\hat{N}_\gamma = \sum_{i=SPL(\gamma)} W_{ijkl}^{base,BQ} F_{\beta}^{nr,BQ} F_{ijkl}^{trim,BQ},$$

where

$W_{ijkl}^{base,BQ} F_{\beta}^{nr,BQ} F_{ijkl}^{trim,BQ}$ = the sample weight for person l, reflecting all weighting adjustments prior to raking, and

SPL(γ) = the set of background questionnaire respondents and literacy-related background questionnaire nonrespondents in raking dimension category γ.

The adjustment factor for raking dimension category γ is given by $F_\gamma = N_\gamma / \hat{N}_\gamma$. The same process is applied for each raking dimension, each time using the adjusted weights from the previous dimension. This is done iteratively until the sums of the adjusted weights equal all control totals. The raking processes all converged in less than 15 iterations.

For simplicity, the raking factor can be denoted as $F_\gamma^{Cal,BQ}$, where γ can denote each of the interior cells defined by the raking dimensions shown in table 12-6.

At this stage of the weighting process, the calibration is done only for cases in the SAAL states to provide a common base for the NAAL and SAAL samples, prior to compositing the weights. Therefore, the calibration factor was set equal to 1 for all persons outside the six SAAL states. The calibration factor is then applied to the sample weights to create the weights used in the composite weighting process:

$$W_{ijkl}^{cal,BQ} = W_{ijkl}^{base,BQ} F_{\beta}^{nr,BQ} F_{ijkl}^{trim,BQ} F_{\gamma}^{Cal,BQ}$$

### 12.1.5 Compositing Data from the National and State Household Components

The original plan for the 1992 National Adult Literacy Survey (NALS) was to consider the national and state samples as two separate surveys so that national statistics would be prepared from the national sample only and state data would be prepared from the state samples only. An evaluation of the 1992 NALS data showed that the increased sample size resulting from the combination of the two samples improved precision for both state and national estimates (Burke et al. 1994). The combined sample had the additional advantage of producing a single database for state and national statistics. Therefore, the NAAL and SAAL samples were combined for the 2003 NAAL as well. The method of combining data from the state and national samples is referred to as composite estimation.

The standard theoretical foundation of composite estimation requires a knowledge of variances of the statistics of interest, in this case, the literacy scores. This information is necessary to produce the parameters used to combine data from various surveys in a way that minimizes the variances of the composite estimates. After the literacy data became available from the 1992 NALS, new compositing factors were computed for a selected set of statistics. (Refer to section 11.2.4 of the Technical Report and Data File User's Manual for the 1992 National Adult Literacy Survey [NCES 2001-457] [Kirsch et al. 2001]). Also, at that time an approach was developed for creating efficient compositing factors for the next national adult literacy study.

Section 12.1.5.1 describes the composite estimation procedure used for the 2003 NAAL. The calculation of the compositing factors is discussed in section 12.1.5.2.

#### 12.1.5.1 Composite Estimation Procedure

In general, the composite estimator for a combined state sample is given by

$$\hat{Y} = \beta\hat{Y}_{st} + (1-\beta)\hat{Y}_{nt},$$

where

$\hat{Y}$ = the composite estimate for variable Y;

$\beta$ = the composite factor (0 <   < 1);

$\hat{Y}_{st}$ = the estimate of Y coming from the state sample; and

$\hat{Y}_{nt}$ = the estimate of Y coming from the national sample.

The variance of a composite estimator will be smaller than the variance of both the national and state estimates if appropriate composite factors are used. Optimal factors can be found when unbiased estimators exist for the two components and approximate estimates of their variances are available. It should be noted that a composite estimator produces unbiased estimates for any value of ˜ The optimum value of . is the one that results in the lowest variance.

As stated above, the national and state samples were selected independently and each could, thus, produce unbiased estimates of subdomain statistics for persons 16 years and older. Therefore, factors could be derived to produce composite estimators with variances that were smaller than those of either of the two estimates. For statistic Y, the optimal composite factor for a particular state is

$$\beta = \frac{V(\hat{Y}_{nt})}{V(\hat{Y}_{nt}) + V(\hat{Y}_{st})}, \qquad (1)$$

where

$V(\hat{Y}_{nt})$ = the variance of the estimate of Y coming from the national sample and

$V(\hat{Y}_{st})$ = the variance of the estimate of Y coming from the state sample.

A different optimal value of   might be found for each statistic of interest. However, data analyses would be complicated if item-specific values of   were used because items would not add up to totals, or totals derived by summing different items would not agree. Consequently, the goal for NAAL was to associate with each sample person a single compositing factor that although not precisely optimal for any particular statistic would be robust enough to enhance the precision of virtually all composited statistics. This objective was accomplished by focusing on aspects of the sample design that were likely to affect the variance, regardless of the choice of statistic.

### 12.1.5.2 Estimating the Compositing Factors

Two aspects of the design should be reflected in the compositing factors. One is the distinction between cases coming from national certainty or noncertainty PSUs. The next design aspect is the oversampling of Blacks and Hispanics in the national sample. The oversampling introduced variability in the weights and increased the design effect for cases coming from the national sample. To best reflect these design features, separate compositing factors (denoted by δ) were created from the combinations of state, certainty status of national PSUs, and race/ethnicity.

The compositing factor in equation (1) can be rewritten as follows:

$$\bar{\beta}_{\delta} = \frac{Ratio_{\delta}(\text{var})}{Ratio_{\delta}(\text{var})+1},$$

where $Ratio_{\delta}(\text{var})$ is the ratio of the variances from subgroup $\delta$ coming from the state and national samples. This ratio is calculated differently for PSUs that are certainties in the national sample and those that are not certainties in the national sample:

$$Ratio_{\delta}(\text{var}) = \bar{R}_{qg}\frac{n_{(st)\delta}}{n_{(nt)\delta}} \quad \text{for national certainty PSU and}$$

$$Ratio_{\delta}(\text{var}) = \bar{R}_{qg} \times \left\{ \frac{\dfrac{\bar{P}_{(nt)g}}{m_{(nt)\delta}} + \dfrac{1-\bar{P}_{(nt)g}}{n_{(nt)\delta}}}{\dfrac{\bar{P}_{(st)g}}{m_{(st)\delta}} + \dfrac{1-\bar{P}_{(st)g}}{n_{(st)\delta}}} \right\} \quad \text{otherwise,}$$

where

*$n(nt)\delta$* = *the number of respondents in subgroup $\delta$ of the national sample;*

*$n(st)\delta$* = *the number of respondents in subgroup $\delta$ of the state sample;*

*$m(nt)\delta$* = *the number of PSUs in subgroup $\delta$ of the national sample;*

*m(st)δ* = *the number of PSUs in subgroup δ of the state sample;*

$\overline{R}_{qg}$ = *the average value of the ratio of the unit variances for sample cases in race/ethnicity category g in national PSUs with certainty status q;*

$\overline{P}_{(nt)g}$ = *the average proportion of the national unit variance for subgroup g coming from the between-PSU component; and*

$\overline{P}_{(st)g}$ = *the average proportion of the state unit variance for subgroup g coming from the between-PSU component.*

The values of $\overline{R}_{qg}$, $\overline{P}_{(nt)g}$, and $\overline{P}_{(st)g}$ are parameters computed from the postweighting 1992 NALS analysis. These values, along with the calculations of Ratioδ (var) and $\overline{\beta}_{\delta}$, are shown in table 12-7.

**Table 12-7. Calculations of Ratioδ (var) and $\overline{\beta}_{\delta}$, by certainty in national sample, race/ethnicity, and state: 2003**

| Certainty in national sample | Race/ethnicity | State | $n$(nt)δ | m(nt)δ | $\overline{R}_{qg}$ | $\overline{P}_{(nt)g}$ | $n$(st)δ | m(st)δ | $\overline{P}_{(st)g}$ | Ratioδ(var) | $\overline{\beta}_{\delta}$ | $1-\overline{\beta}_{\delta}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| No | Hispanic | KY | 3 | 2 | 1.1322 | 0.0032 | 26 | 18 | 0.0004 | 9.8264 | 0.9076 | 0.0924 |
| No | Black | KY | 53 | 2 | 0.7365 | 0.0054 | 113 | 18 | 0.0041 | 1.7487 | 0.6362 | 0.3638 |
| No | Other | KY | 138 | 2 | 0.8115 | 0.0008 | 1,212 | 18 | 0.0006 | 7.2272 | 0.8785 | 0.1215 |
| No | Hispanic | MD | 2 | 1 | 1.1322 | 0.0032 | 38 | 13 | 0.0004 | 21.5641 | 0.9557 | 0.0443 |
| No | Black | MD | 75 | 1 | 0.7365 | 0.0054 | 247 | 13 | 0.0041 | 3.1615 | 0.7597 | 0.2403 |
| No | Other | MD | 53 | 1 | 0.8115 | 0.0008 | 601 | 13 | 0.0006 | 9.3317 | 0.9032 | 0.0968 |
| Yes | Hispanic | MA | 18 | 1 | 1.1275 | — | 28 | 1 | — | 1.7539 | 0.6369 | 0.3631 |
| Yes | Black | MA | 6 | 1 | 0.7363 | — | 21 | 1 | — | 2.5771 | 0.7204 | 0.2796 |
| Yes | Other | MA | 81 | 1 | 0.8118 | — | 331 | 1 | — | 3.317 | 0.7684 | 0.2316 |
| No | Hispanic | MA | 33 | 1 | 1.1322 | 0.0032 | 43 | 5 | 0.0004 | 1.621 | 0.6185 | 0.3815 |
| No | Black | MA | 17 | 1 | 0.7365 | 0.0054 | 29 | 5 | 0.0041 | 1.3386 | 0.5724 | 0.4276 |
| No | Other | MA | 89 | 1 | 0.8115 | 0.0008 | 378 | 5 | 0.0006 | 3.5312 | 0.7793 | 0.2207 |
| No | Hispanic | MO | 1 | 1 | 1.1322 | 0.0032 | 25 | 12 | 0.0004 | 28.2927 | 0.9659 | 0.0341 |
| No | Black | MO | 0 | 0 | 0.7365 | 0.0054 | 99 | 12 | 0.0041 | † | 1.0000 | — |
| No | Other | MO | 84 | 1 | 0.8115 | 0.0008 | 800 | 12 | 0.0006 | 7.9293 | 0.8880 | 0.1120 |
| Yes | Hispanic | NY | 181 | 4 | 1.1275 | — | 90 | 4 | — | 0.5606 | 0.3592 | 0.6408 |
| Yes | Black | NY | 187 | 4 | 0.7363 | — | 75 | 4 | — | 0.2953 | 0.2280 | 0.7720 |
| Yes | Other | NY | 150 | 4 | 0.8118 | — | 203 | 4 | — | 1.0986 | 0.5235 | 0.4765 |
| No | Hispanic | NY | 27 | 4 | 1.1322 | 0.0032 | 15 | 8 | 0.0004 | 0.6403 | 0.3904 | 0.6096 |
| No | Black | NY | 92 | 4 | 0.7365 | 0.0054 | 13 | 8 | 0.0041 | 0.1161 | 0.1041 | 0.8959 |
| No | Other | NY | 386 | 4 | 0.8115 | 0.0008 | 311 | 8 | 0.0006 | 0.6881 | 0.4076 | 0.5924 |
| No | Hispanic | OK | 8 | 2 | 1.1322 | 0.0032 | 60 | 12 | 0.0004 | 8.5593 | 0.8954 | 0.1046 |
| No | Black | OK | 5 | 2 | 0.7365 | 0.0054 | 70 | 12 | 0.0041 | 10.1925 | 0.9107 | 0.0893 |
| No | Other | OK | 184 | 2 | 0.8115 | 0.0008 | 960 | 12 | 0.0006 | 4.3366 | 0.8126 | 0.1874 |

— Not available.
† Not applicable.
NOTE: All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as other include non-Hispanics of all other races including multiracial. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 12.1.6 Computing Final Weights

The final weights were created by applying the composite factors to the calibrated weights (section 12.1.6.1) and then raking the weights to control totals. The raking process included the following sequence of subtasks: raking, trimming, and reraking (refer to section 12.1.6.2).

#### 12.1.6.1 Compositing the NAAL and SAAL Samples

After calculating the compositing factor, $\overline{\beta}_{\delta}$, the composited weight, $W_{ijkl}^{comp,BQ}$, was computed as follows:

$W_{ijkl}^{comp,BQ}$ = $\overline{\beta}_{\delta} W_{ijkl}^{cal,BQ}$ for person l in subgroup δ, and associated with the SAAL sample;

= $(1-\overline{\beta}_{\delta})W_{ijkl}^{cal,BQ}$ for person l in subgroup δ, and associated with the national NAAL household sample in SAAL states; and

= $W_{ijkl}^{cal,BQ}$ for person l in non-SAAL states.

#### 12.1.6.2 Raking Composite Weights

The final step in weighting was to rake the composited weights to control totals. The raking process was completed for the entire sample in a manner similar to the calibration performed before compositing for the six SAAL states. The process included the following steps: creating control totals, raking, trimming, and reraking. The creation of the control totals for the calibration adjustment is discussed in section 12.1.6.2.1. The weighting steps of raking, trimming, and reraking are discussed in section 12.1.6.2.2.

##### 12.1.6.2.1 Control Totals

Control totals were computed for the purpose of creating the final weights for the entire combined NAAL and SAAL household sample. The totals were computed from the 2003 CPS March Supplement for each SAAL state and the remainder of the United States. For each sample, control totals were computed for the following variables: MSA status, age, gender, education, country of birth, and race/ethnicity. Census region was used for the sample containing the remainder of the United States. Not all of these variables were used for each state sample because of small sample sizes in certain domains. However, the variables used for the raking step were defined to the finest classification that the data allow. Further discussion as to the selection of raking variables is in section 12.1.4.1. Table 12-8 displays the variables involved in the raking process.

#### 12.1.6.2.2 Final Adjustments

The final steps in the weighting process were raking, trimming, and raking again to recalibrate the weights to the control totals. A general overview of trimming is provided in section 12.1.3.3, and a general overview of raking is provided in section 12.1.4.2.

Respondents who completed the background questionnaire were included in the raking process. Literacy-related nonrespondents were also included because (1) the reasons for nonparticipation have been found to be highly related to literacy results (NAAL 2001) and (2) they are part of the target population from which the control totals were derived. Table 12-9 summarizes the raking factors for the first round of raking. In general, the mean raking factors are near 1.0 for SAAL states and 1.13 for the rest of the nation. By domain, the means are more variable as one would expect because the sample sizes are smaller for domains in general. The table also shows the range of the raking factors by state and by raking dimension. Convergence in raking is generally achieved within the first few iterations. A maximum number of 15 iterations was preset for the reason that any further processing would be for naught because convergence would be unlikely. All raking processes converged in less than 15 iterations.

After raking, the trimming process was repeated to adjust extreme weights created after raking. In this step, fewer than 1 percent of the weights were reduced.

The last step was a second round of raking. Table 12-10 summarizes the raking factors for the second round of raking after the compositing procedure was applied. As shown in the table, this raking step had little effect on the weights as the adjustment factors are near 1.0.

**Table 12-8. Variables involved in the raking process after compositing, by sample: 2003**

| Sample | MSA status | Age | Gender[1] | Education[2] | Country of birth | Race/ethnicity | Census region |
|---|---|---|---|---|---|---|---|
| Kentucky | MSA, non-MSA | 16–29, 30–49, 50+ | M, F | <HS, HS, + | | Hispanic/Black, other | |
| Maryland | | 16–29, 30–49, 50–69, 70+ | M, F | <HS, HS, + | U.S., other | Hispanic/Black, other | |
| Massachusetts | | 16–29, 30–49, 50–69, 70+ | M, F | <HS, HS, + | U.S., other | Hispanic/Black, other | |
| Missouri | MSA, non-MSA | 16–29, 30–49, 50–69, 70+ | M, F | <HS, HS, + | | Hispanic/Black, other | |
| New York | MSA, non-MSA | 16–29, 30–49, 50–69, 70+ | M, F | <HS, HS, + | U.S., other | Hispanic/Black, other | |
| Oklahoma | MSA, non-MSA | 16–29, 30–49, 50–69, 70+ | M, F | <HS, HS, + | | Hispanic/Black, other | |
| Remainder of U.S. | MSA, non-MSA | 16–29, 30–49, 50–69, 70+ | M, F | <HS, HS, + | U.S., other | Hispanic/Black, other | Northeast, Midwest, South, West |

[1] M: Males, F: Females.
[2] <HS: less than high school; HS: high school diploma or equivalent; +: more than high school.
NOTE: All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as other include non-Hispanics of all other races including multiracial. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 12-9. Raking factors for the first round of raking after compositing, by state and domain: 2003**

| Domain | Kentucky | | | | Maryland | | | | Massachusetts | | | | Missouri | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | Min | Max | Mean | *n* | Min | Max | Mean | *n* | Min | Max | Mean | *n* | Min | Max | Mean |
| Overall | 1,545 | 0.90 | 1.05 | 0.99 | 1,016 | 0.85 | 1.24 | 0.99 | 1,074 | 0.76 | 1.20 | 0.99 | 1,009 | 0.66 | 1.19 | 0.96 |
| Region | | | | | | | | | | | | | | | | |
| Northeast | † | † | † | † | † | † | † | † | † | † | † | † | † | † | † | † |
| Midwest | † | † | † | † | † | † | † | † | † | † | † | † | † | † | † | † |
| South | † | † | † | † | † | † | † | † | † | † | † | † | † | † | † | † |
| West | † | † | † | † | † | † | † | † | † | † | † | † | † | † | † | † |
| MSA status | | | | | | | | | | | | | | | | |
| Non-MSA | 807 | 0.93 | 1.05 | 1.01 | † | † | † | † | † | † | † | † | 400 | 0.66 | 0.79 | 0.75 |
| MSA | 738 | 0.90 | 1.02 | 0.98 | † | † | † | † | † | † | † | † | 609 | 0.99 | 1.19 | 1.10 |
| Age | | | | | | | | | | | | | | | | |
| 16–29 | 361 | 0.95 | 1.05 | 1.01 | 222 | 0.95 | 1.24 | 1.06 | 224 | 0.82 | 1.20 | 1.04 | 233 | 0.66 | 1.14 | 0.95 |
| 30–49 | 591 | 0.90 | 1.00 | 0.98 | 406 | 0.85 | 1.11 | 0.95 | 465 | 0.76 | 1.11 | 0.97 | 385 | 0.67 | 1.16 | 0.97 |
| 50–69 | 593 | 0.93 | 1.03 | 1.00 | 281 | 0.89 | 1.15 | 0.99 | 257 | 0.78 | 1.12 | 0.98 | 267 | 0.67 | 1.16 | 0.96 |
| 70+ | | | | | 107 | 0.90 | 1.18 | 1.00 | 128 | 0.83 | 1.19 | 1.04 | 124 | 0.76 | 1.19 | 0.98 |
| Gender | | | | | | | | | | | | | | | | |
| Male | 637 | 0.90 | 1.05 | 0.99 | 413 | 0.87 | 1.24 | 1.00 | 486 | 0.78 | 1.20 | 1.00 | 458 | 0.68 | 1.19 | 0.96 |
| Female | 908 | 0.91 | 1.05 | 1.00 | 603 | 0.85 | 1.21 | 0.98 | 588 | 0.76 | 1.19 | 0.99 | 551 | 0.66 | 1.18 | 0.96 |
| Education | | | | | | | | | | | | | | | | |
| Less than HS[1] | 363 | 0.90 | 1.02 | 0.98 | 151 | 0.85 | 1.16 | 0.95 | 176 | 0.76 | 1.11 | 0.93 | 213 | 0.69 | 1.19 | 0.95 |
| HS | 518 | 0.93 | 1.04 | 1.00 | 256 | 0.91 | 1.20 | 0.99 | 209 | 0.83 | 1.20 | 1.02 | 248 | 0.68 | 1.17 | 0.96 |
| More than HS | 664 | 0.94 | 1.05 | 1.00 | 609 | 0.90 | 1.24 | 1.00 | 689 | 0.82 | 1.19 | 1.00 | 548 | 0.66 | 1.15 | 0.97 |
| Born in U.S. | | | | | | | | | | | | | | | | |
| No | † | † | † | † | 130 | 0.97 | 1.24 | 1.10 | 244 | 0.76 | 1.00 | 0.87 | † | † | † | † |
| Yes | † | † | † | † | 886 | 0.85 | 1.09 | 0.97 | 830 | 0.92 | 1.20 | 1.03 | † | † | † | † |
| Race/ethnicity | | | | | | | | | | | | | | | | |
| Hispanic | 195 | 0.904 | 1.01 | 0.95 | 362 | 0.85 | 1.20 | 0.97 | 195 | 0.84 | 1.20 | 0.98 | 125 | 0.66 | 1.09 | 0.99 |
| Non-Hispanic | | | | | | | | | | | | | | | | |
| Other | 1,350 | 0.94 | 1.05 | 1.00 | 654 | 0.88 | 1.24 | 1.00 | 879 | 0.76 | 1.10 | 1.00 | 884 | 0.72 | 1.19 | 0.96 |

See notes at end of table.

**Table 12-9. Raking factors for the first round of raking after compositing, by state and domain: 2003—Continued**

| Domain | New York | | | | Oklahoma | | | | Remaining sample | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | Min | Max | Mean | *n* | Min | Max | Mean | *n* | Min | Max | Mean |
| Overall | 1,730 | 0.61 | 4.30 | 1.03 | 1,287 | 0.95 | 1.09 | 0.99 | 10,880 | 0.65 | 1.77 | 1.13 |
| Region | | | | | | | | | | | | |
| Northeast | † | † | † | † | † | † | † | † | 961 | 0.86 | 1.77 | 1.25 |
| Midwest | † | † | † | † | † | † | † | † | 2,603 | 0.65 | 1.41 | 1.00 |
| South | † | † | † | † | † | † | † | † | 4,422 | 0.75 | 1.65 | 1.14 |
| West | † | † | † | † | † | † | † | † | 2,894 | 0.80 | 1.76 | 1.19 |
| MSA status | | | | | | | | | | | | |
| Non-MSA | 73 | 1.84 | 4.30 | 2.73 | 515 | 0.97 | 1.09 | 1.00 | 2,043 | 0.65 | 1.40 | 0.98 |
| MSA | 1,657 | 0.61 | 1.90 | 0.95 | 772 | 0.95 | 1.08 | 0.99 | 8,837 | 0.81 | 1.77 | 1.16 |
| Age | | | | | | | | | | | | |
| 16–29 | 424 | 0.64 | 3.88 | 0.99 | 355 | 0.99 | 1.09 | 1.01 | 2,893 | 0.67 | 1.67 | 1.14 |
| 30–49 | 669 | 0.65 | 3.92 | 1.04 | 429 | 0.95 | 1.06 | 0.98 | 4,316 | 0.66 | 1.61 | 1.10 |
| 50–69 | 437 | 0.61 | 3.70 | 1.00 | 337 | 0.97 | 1.08 | 1.00 | 2,573 | 0.65 | 1.63 | 1.12 |
| 70+ | 200 | 0.71 | 4.30 | 1.12 | 166 | 0.97 | 1.08 | 0.99 | 1,098 | 0.80 | 1.77 | 1.23 |
| Gender | | | | | | | | | | | | |
| Male | 706 | 0.70 | 4.30 | 1.10 | 557 | 0.96 | 1.09 | 1.00 | 4,771 | 0.68 | 1.77 | 1.14 |
| Female | 1,024 | 0.61 | 3.83 | 0.97 | 730 | 0.95 | 1.09 | 0.99 | 6,109 | 0.65 | 1.74 | 1.12 |
| Education | | | | | | | | | | | | |
| Less than HS[1] | 441 | 0.61 | 2.44 | 0.83 | 314 | 0.95 | 1.09 | 1.00 | 2,928 | 0.69 | 1.37 | 1.06 |
| HS | 389 | 1.08 | 4.30 | 1.44 | 355 | 0.96 | 1.08 | 1.00 | 2,734 | 0.86 | 1.77 | 1.36 |
| More than HS | 900 | 0.75 | 3.00 | 0.94 | 618 | 0.96 | 1.09 | 0.99 | 5,218 | 0.65 | 1.33 | 1.05 |
| Born in U.S. | | | | | | | | | | | | |
| No | 470 | 0.65 | 3.83 | 0.96 | † | † | † | † | 2,247 | 0.65 | 1.72 | 1.08 |
| Yes | 1,260 | 0.61 | 4.30 | 1.05 | † | † | † | † | 8,633 | 0.66 | 1.77 | 1.14 |
| Race/ethnicity | | | | | | | | | | | | |
| Hispanic | 313 | 0.64 | 3.83 | 0.91 | 143 | 1.03 | 1.09 | 1.06 | 2,596 | 0.65 | 1.57 | 1.08 |
| Non-Hispanic Black Only | 367 | 0.76 | 1.90 | 1.10 | † | † | † | † | 2,402 | 0.72 | 1.67 | 1.15 |
| Other | 1,050 | 0.61 | 4.30 | 1.03 | 1,144 | 0.95 | 1.02 | 0.99 | 5,882 | 0.72 | 1.77 | 1.14 |

† Not applicable.
[1] High school.
NOTE: All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as other include non-Hispanics of all other races including multiracial. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 12-10. Raking factors for the second round of raking after compositing, by state and domain: 2003**

| Domain | Kentucky | | | | Maryland | | | | Massachusetts | | | | Missouri | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | Min | Max | Mean | *n* | Min | Max | Mean | *n* | Min | Max | Mean | *n* | Min | Max | Mean |
| Overall | 1,545 | 1.00 | 1.02 | 1.00 | 1,016 | 0.99 | 1.01 | 1.00 | 1,074 | 0.98 | 1.07 | 1.01 | 1,009 | 0.98 | 1.04 | 1.01 |
| Region | | | | | | | | | | | | | | | | |
| Northeast | † | † | † | † | † | † | † | † | † | † | † | † | † | † | † | † |
| Midwest | † | † | † | † | † | † | † | † | † | † | † | † | † | † | † | † |
| South | † | † | † | † | † | † | † | † | † | † | † | † | † | † | † | † |
| West | † | † | † | † | † | † | † | † | † | † | † | † | † | † | † | † |
| MSA status | | | | | | | | | | | | | | | | |
| Non-MSA | 807 | 1.00 | 1.02 | 1.00 | † | † | † | † | † | † | † | † | 400 | 0.98 | 1.03 | 1.00 |
| MSA | 738 | 1.00 | 1.02 | 1.00 | † | † | † | † | † | † | † | † | 609 | 1.00 | 1.04 | 1.01 |
| Age | | | | | | | | | | | | | | | | |
| 16–29 | 361 | 1.00 | 1.02 | 1.00 | 222 | 0.99 | 1.01 | 1.00 | 224 | 0.99 | 1.07 | 1.02 | 233 | 0.99 | 1.04 | 1.01 |
| 30–49 | 591 | 1.00 | 1.02 | 1.00 | 406 | 0.99 | 1.01 | 1.01 | 465 | 0.98 | 1.06 | 1.01 | 385 | 0.99 | 1.04 | 1.00 |
| 50–69 | 593 | 1.00 | 1.01 | 1.01 | 281 | 0.99 | 1.01 | 1.00 | 257 | 0.98 | 1.06 | 1.00 | 267 | 0.98 | 1.03 | 1.00 |
| 70+ | † | † | † | † | 107 | 0.99 | 1.01 | 1.00 | 128 | 0.98 | 1.05 | 1.00 | 124 | 0.98 | 1.03 | 1.00 |
| Gender | | | | | | | | | | | | | | | | |
| Male | 637 | 1.00 | 1.02 | 1.00 | 413 | 0.99 | 1.01 | 1.00 | 486 | 0.98 | 1.07 | 1.01 | 458 | 0.98 | 1.04 | 1.00 |
| Female | 908 | 1.00 | 1.02 | 1.00 | 603 | 0.99 | 1.01 | 1.00 | 588 | 0.98 | 1.06 | 1.01 | 551 | 0.99 | 1.04 | 1.01 |
| Education | | | | | | | | | | | | | | | | |
| Less than HS | 363 | 1.00 | 1.02 | 1.00 | 151 | 0.99 | 1.01 | 1.00 | 176 | 0.98 | 1.04 | 1.01 | 213 | 0.98 | 1.01 | 1.00 |
| HS | 518 | 1.00 | 1.01 | 1.00 | 256 | 0.99 | 1.01 | 1.00 | 209 | 1.00 | 1.07 | 1.02 | 248 | 1.01 | 1.04 | 1.03 |
| More than HS | 664 | 1.00 | 1.02 | 1.00 | 609 | 0.99 | 1.01 | 1.00 | 689 | 0.98 | 1.05 | 1.00 | 548 | 0.98 | 1.01 | 1.00 |
| Born in U.S. | | | | | | | | | | | | | | | | |
| No | † | † | † | † | 130 | 0.99 | 1.01 | 1.00 | 244 | 0.98 | 1.06 | 1.01 | † | † | † | † |
| Yes | † | † | † | † | 886 | 0.99 | 1.01 | 1.00 | 830 | 0.99 | 1.07 | 1.01 | † | † | † | † |
| Race/ethnicity | | | | | | | | | | | | | | | | |
| Hispanic | 195 | 1.01 | 1.02 | 1.01 | 362 | 1.00 | 1.01 | 1.01 | 195 | 1.02 | 1.07 | 1.04 | 125 | 0.98 | 1.04 | 1.01 |
| Non-Hispanic | † | † | † | † | † | † | † | † | † | † | † | † | † | † | † | † |
| Other | 1,350 | 1.00 | 1.01 | 1.00 | 654 | 0.99 | 1.01 | 1.00 | 879 | 0.98 | 1.03 | 1.00 | 884 | 0.98 | 1.04 | 1.00 |

See notes at end of table.

**Table 12-10. Raking factors for the second round of raking after compositing, by state and domain: 2003—Continued**

| Domain | New York | | | | Oklahoma | | | | Remaining sample | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | Min | Max | Mean | *n* | Min | Max | Mean | *n* | Min | Max | Mean |
| Overall | 1,730 | 0.97 | 1.12 | 1.02 | 1,287 | 1.00 | 1.00 | 1.00 | 10,880 | 0.994 | 1.01 | 1.00 |
| Region | | | | | | | | | | | | |
| Northeast | † | † | † | † | † | † | † | † | 961 | 0.996 | 1.01 | 1.00 |
| Midwest | † | † | † | † | † | † | † | † | 2,603 | 0.994 | 1.01 | 1.00 |
| South | † | † | † | † | † | † | † | † | 4,422 | 0.997 | 1.01 | 1.00 |
| West | † | † | † | † | † | † | † | † | 2,894 | 0.994 | 1.01 | 1.00 |
| MSA status | | | | | | | | | | | | |
| Non-MSA | 73 | 1.00 | 1.12 | 1.05 | 515 | 1.00 | 1.00 | 1.00 | 2,043 | 0.994 | 1.01 | 1.00 |
| MSA | 1,657 | 0.97 | 1.12 | 1.02 | 772 | 1.00 | 1.00 | 1.00 | 8,837 | 0.996 | 1.01 | 1.00 |
| Age | | | | | | | | | | | | |
| 16–29 | 424 | 0.98 | 1.12 | 1.02 | 355 | 1.00 | 1.00 | 1.00 | 2,893 | 0.995 | 1.01 | 1.00 |
| 30–49 | 669 | 0.97 | 1.11 | 1.02 | 429 | 1.00 | 1.00 | 1.00 | 4,316 | 0.994 | 1.00 | 1.00 |
| 50–69 | 437 | 0.99 | 1.12 | 1.03 | 337 | 1.00 | 1.00 | 1.00 | 2,573 | 0.995 | 1.01 | 1.00 |
| 70+ | 200 | 0.98 | 1.11 | 1.02 | 166 | 1.00 | 1.00 | 1.00 | 1,098 | 0.994 | 1.01 | 1.00 |
| Gender | | | | | | | | | | | | |
| Male | 706 | 0.99 | 1.12 | 1.04 | 557 | 1.00 | 1.00 | 1.00 | 4,771 | 0.995 | 1.01 | 1.00 |
| Female | 1,024 | 0.97 | 1.11 | 1.01 | 730 | 1.00 | 1.00 | 1.00 | 6,109 | 0.994 | 1.01 | 1.00 |
| Education | | | | | | | | | | | | |
| Less than HS | 441 | 0.97 | 1.03 | 1.00 | 314 | 1.00 | 1.00 | 1.00 | 2,928 | 0.994 | 1.01 | 1.00 |
| HS | 389 | 1.06 | 1.12 | 1.08 | 355 | 1.00 | 1.00 | 1.00 | 2,734 | 0.999 | 1.01 | 1.00 |
| More than HS | 900 | 0.99 | 1.05 | 1.01 | 618 | 1.00 | 1.00 | 1.00 | 5,218 | 0.994 | 1.01 | 1.00 |
| Born in U.S. | | | | | | | | | | | | |
| No | 470 | 0.99 | 1.12 | 1.03 | † | † | † | † | 2,247 | 0.996 | 1.01 | 1.00 |
| Yes | 1,260 | 0.97 | 1.12 | 1.02 | † | † | † | † | 8,633 | 0.994 | 1.01 | 1.00 |
| Race/ethnicity | | | | | | | | | | | | |
| Hispanic | 313 | 0.97 | 1.12 | 1.02 | 143 | 1.00 | 1.00 | 1.00 | 2,596 | 0.994 | 1.01 | 1.00 |
| Non-Hispanic | 367 | 0.97 | 1.12 | 1.02 | † | † | † | † | 2,402 | 0.994 | 1.01 | 1.00 |
| Other | 1,050 | 0.97 | 1.12 | 1.02 | 1,144 | 1.00 | 1.00 | 1.00 | 5,882 | 0.995 | 1.01 | 1.00 |

† Not applicable.
NOTE: All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as other include non-Hispanics of all other races including multiracial. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

The final weights were computed as the product of the composited background questionnaire weight and the raking and trimming factors, as shown below:

$$W_{ijkl}^{final,BQ} = W_{ijkl}^{comp,BQ} F_{\gamma}^{rake1,BQ} F_{ijkl}^{trim2,BQ} F_{\gamma}^{rake2,BQ} ,$$

where

$F_{\gamma}^{rake1,BQ}$ = the first-round raking factor for sample persons in raking cell γ;

$F_{ijkl}^{trim2,BQ}$ = the trimming factor after compositing for sample person l; and

$F_{\gamma}^{rake2,BQ}$ = the second-round raking factor for sample persons in raking cell γ.

The distribution of the final background questionnaire weights is shown in table 12-11 for the combined NAAL and SAAL sample. The coefficient of variation, shown in the table, measures the variation of the sampling weights, which can affect the precision of survey outcome statistics. The total combined NAAL and SAAL sample has a relatively high coefficient of variation of the weights due to a mixture of oversampling Blacks and Hispanics in the national sample and higher sampling rates for SAAL states.

**Table 12-11. Distribution of final background questionnaire weights, by domain: 2003**

| Domain | Sample size | Minimum | Sum | Median | Maximum | Coefficient of variation |
|---|---|---|---|---|---|---|
| Total | 18,541 | 149 | 221,021,328 | 7,697 | 72,267 | 95 |
| | | | | | | |
| Kentucky | 1,545 | 149 | 3,222,654 | 2,052 | 6,482 | 44 |
| Maryland | 1,016 | 887 | 4,228,643 | 3,856 | 11,899 | 50 |
| Massachusetts | 1,074 | 568 | 5,155,801 | 4,316 | 14,268 | 53 |
| Missouri | 1,009 | 739 | 4,355,187 | 3,835 | 11,673 | 60 |
| New York | 1,730 | 809 | 15,119,508 | 7,008 | 27,174 | 67 |
| Oklahoma | 1,287 | 221 | 2,706,561 | 1,959 | 5,963 | 46 |
| Rest of U.S. | 10,880 | 1,567 | 186,232,851 | 13,176 | 72,267 | 69 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 12.1.6.3 Quality Control

When several stages of adjustments are used to produce sampling weights, it is essential that quality control (QC) checks are done throughout the weighting process. The checks done for NAAL included reviewing the computer code used to calculate the weights, validating the implementation of weighting specifications, and calculating and checking descriptive statistics on the weights.

A standardized weighting system was used to perform the tasks of nonresponse adjustment, raking, and the creation of replicate weights for variance estimation. The system has been used on numerous surveys and has been well tested through careful review procedures.

Despite careful review of the specifications and program, there still needs to be careful review of the resulting weights. After each weighting step, the weighted totals and percent distribution for several key domains was compared to the weighted totals and percent distribution prior to the weighting step. This allowed the statistician to identify any large changes in distribution and investigate into the reasons for the changes.

### 12.1.7 Replicate Weights

Variance estimation must take into account the sample design. In particular, the sampling variance estimate for any statistic should account for the effects of clustering; the use of nonresponse, trimming, and poststratification adjustments; and the component of sampling variability arising from the variation in the weights used to compute the statistic. Treating the data as a simple random sample will produce underestimates of the true sampling variability for the NAAL area sample design.

The stratified jackknife method can be used to estimate the variance for most statistics. Jackknifing estimates the sampling variability of any statistic Y, as the sum of components of variability that may be attributed to individual pairs of first-stage sampling units. The variance attributed to a particular pair is measured by estimating how much the value of the statistic would change if only one unit in the pair had been sampled. When using replication techniques such as jackknifing to calculate standard errors, it is necessary to establish a number of subsamples (or replicates) from the full sample, calculate the estimate from each subsample, and sum the squared difference of each replicated estimate from the full-sample estimate. The 61 replicates formed for the NAAL and SAAL combined household sample provided the degrees of freedom necessary for the production of stable estimates of variance.

Three steps were involved in facilitating variance estimation: (1) forming the replicates, (2) constructing the replicate weights, and (3) computing estimates of variance for survey statistics. Under the stratified jackknife approach employed for NAAL, the number of replicates is equal to the number of degrees of freedom, which is directly related to the stability of variance estimates (Rust 1985). The formation of replicates is discussed in detail in sections 12.1.7.1 through 12.1.7.3. After the replicates were formed, a replicate factor was constructed for each variance stratum. Let $f_{ijk}(r)$ denote the rth replicate factor for sampled dwelling unit k in variance unit i′ of variance stratum h′ (the prime symbol is used to distinguish between the variance unit and PSU i, and also between the variance stratum and PSU stratum h). Then, in general,

$$f_{ijk}(r) = \begin{cases} 2 & \text{if } h' = r \text{ and } i' = 1; \\ 0 & \text{if } h' = r \text{ and } i' = 2; \text{ and} \\ 1 & \text{if } h' \neq r, \end{cases}$$

and the replicated screener base weight, $W_{ijk}^{base,SCR}(r)$, was obtained as

$$W_{ijk}^{base,SCR}(r) = W_{ijk}^{base,SCR} f_{ijk}(r)$$

for r = 1, 2,˜, 61. For SAAL, 13 to 19 replicate weights were formed, depending on the state.

After obtaining a screener base weight for each replicate, all remaining full-sample weighting steps leading to the final person weight were performed on each replicate. The repetition of the various weight adjustment procedures on each set of replicate base weights ensures that the impact of these procedures on the sampling variance of the estimator Y is appropriately reflected in the variance estimator, v(Y).

### 12.1.7.1 Replicate Weights for the National Sample

The national sample contained 100 PSUs, 16 of which were selected with certainty. The remaining 84 PSUs were selected one per stratum with probabilities proportional to size. Replicates were formed by pairing first-stage sampled units; that is, segments were paired in PSUs selected with certainty, and whole PSUs were paired in noncertainty strata. For the 100-PSU sample, the natural pairing led to 61 replicates.

### 12.1.7.2 Replicate Weights for the State Sample

An independent sample of PSUs was selected in each of the six participating states. The largest PSUs were taken with certainty. Within each state, the remaining PSUs were grouped into strata, and from each stratum a single PSU was sampled with probability proportional to size. In PSUs selected with certainty, segments were paired to form replicates. The number of replicates formed for each state was as follows:

| | |
|---|---|
| Kentucky | 16 |
| Maryland | 15 |
| Massachusetts | 13 |
| Missouri | 13 |
| New York | 19 |
| Oklahoma | 13 |

For Maryland and Massachusetts, segments were the first-stage sampling units. Therefore, the selected segments were grouped and paired to form replicates. In addition, Kentucky and Oklahoma each had one triplet; that is, one variance stratum had three variance units. The triplets generate one additional replicate, while affecting two replicates in total, using a factor of 1.5 to construct the replicate weights.

### 12.1.7.3 Final Replication for the National and State Samples

The NAAL analysis combined data from a nationally representative sample of 100 PSUs with data from 6 independently selected state PSU samples. The threefold objective of the replication scheme was to (1) reflect the actual sample design of each sample; (2) ensure the production of stable estimates of standard errors by having sufficient degrees of freedom for national estimates, individual SAAL state estimates, and regional estimates; and (3) limit the total number of replicates so that variance estimation could be done more efficiently. The general approach in setting up the replication was to devise an appropriate scheme for each component of the sample, the national sample, and the six states, and then to collapse replicates to a reasonable number.

#### 12.1.7.3.1 Active Replicates

A total of 150 replicates had been formed at this point: 61 from the national sample and 89 from the state samples. These replicates reflected the actual design of each sample and provided sufficient degrees of freedom to produce stable estimates of variance for the nation, each state, and the four census regions. However, using 150 replicates to estimate variances would greatly increase the computer processing time for data analysis, while providing only a slight gain in the precision of the overall variance estimates. This is analogous to increases in sample size providing diminishing returns with respect to the precision of estimates. Refer to Rust and Rao (1996) for further discussion. Therefore, the number of replicates was collapsed to 61. To preserve the total number of replicates for each state, replicates from the same state were never collapsed. To the extent possible, replicates from the same region were not combined either.

Table 12-12 presents the results of the replication scheme, showing which replicates are active for the major subdomains of analysis.

#### 12.1.7.3.2 Replicated Control Totals

As mentioned above, the 2003 CPS March Supplement was used to create the population control totals for the household component. In general, control totals derived from the CPS have some variance associated with them because the CPS is a survey (not a census). Usually, the sampling errors are ignored when using the control totals for the U.S. population (or main subgroups of the U.S. population) because the very large size of the CPS sample results in very small variances. However, the state sample sizes in the CPS are smaller, and the variances associated with the state-level control totals are relatively high. This section describes the approach used to add variation to the replicate totals, reflecting the CPS variances. The CPS variances are measured using generalized variance functions (GVFs). The GVF model is a regression model fit to the survey relative variance, V2, as follows,

$$V^2=A+B/X$$

where A is the intercept, B is the slope, and X is a set of survey estimates. The resulting parameters, A and B, can be used to approximate the standard error associated with any survey estimate X. The GVF parameters are found in the source and accuracy statement of the 2003 CPS March Supplement.

**Table 12-12. Active replicates for subdomains of the NAAL analysis file, by replicate, region, and state: 2003**

| Replicate | Region and state | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | U.S. | Northeast | Midwest | South | West | Massachusetts | New York | Kentucky | Oklahoma | Maryland | Missouri |
| 1 | x | x | | x | | | | x | | | |
| 2 | x | x | | x | | x | | x | | | |
| 3 | x | x | | x | | | x | x | | | |
| 4 | x | x | | x | | | x | x | | | |
| 5 | x | x | | x | | | x | x | | | |
| 6 | x | x | x | x | | | x | x | | | |
| 7 | x | x | x | x | | | x | x | | | |
| 8 | x | x | x | x | | | x | x | | | |
| 9 | x | x | x | x | | | x | x | | | |
| 10 | x | x | x | x | | | x | | | x | |
| 11 | x | x | x | x | | | x | | | x | |
| 12 | x | x | x | x | | | x | | | x | |
| 13 | x | x | x | x | | | x | | | x | |
| 14 | x | x | | x | | | x | | | | |
| 15 | x | x | | x | | | x | | | | |
| 16 | x | x | | x | | | x | | | | |
| 17 | x | x | | x | | | x | | | | |
| 18 | x | x | | x | | | x | | | x | |
| 19 | x | x | | x | | | x | | | | |
| 20 | x | x | | x | | | x | | | | |
| 21 | x | x | | x | | | x | | | | |
| 22 | x | x | | x | | | x | x | | | |
| 23 | x | x | | x | | | x | | | | |
| 24 | x | x | | x | | | x | | x | | |
| 25 | x | x | | x | | x | | | | | |
| 26 | x | x | | x | x | x | | | | x | |
| 27 | x | x | | x | x | x | | | | x | |
| 28 | x | x | | x | x | x | | | | x | |
| 29 | x | x | | x | x | x | | | | x | |
| 30 | x | x | | x | x | x | | | | x | |
| 31 | x | x | | x | x | x | | | | x | |
| 32 | x | x | | x | | | | | | x | |

See notes at end of table.

**Table 12-12. Active replicates for subdomains of the NAAL analysis file, by replicate, region, and state: 2003—Continued**

| Replicate | Region and state | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | U.S. | Northeast | Midwest | South | West | Massachusetts | New York | Kentucky | Oklahoma | Maryland | Missouri |
| 33 | x | x | x | x | | x | | | | x | |
| 34 | x | x | x | x | | x | | | | x | x |
| 35 | x | x | x | x | | x | | | | x | |
| 36 | x | x | | x | | x | | | | | |
| 37 | x | x | | x | | x | | | | | |
| 38 | x | x | | x | | x | | | | | |
| 39 | x | | x | x | | | | x | | | x |
| 40 | x | | x | x | | | | | x | | x |
| 41 | x | | x | x | x | | | | | x | x |
| 42 | x | | x | x | x | | | | x | | x |
| 43 | x | x | x | x | | x | | | x | | x |
| 44 | x | x | x | x | | | | | x | | x |
| 45 | x | x | x | x | | | x | | x | | x |
| 46 | x | x | x | x | | | x | | x | | x |
| 47 | x | x | x | x | | | x | | x | | x |
| 48 | x | x | x | x | | | x | | x | | x |
| 49 | x | | x | x | | | | | x | | |
| 50 | x | | x | x | | | | | x | | |
| 51 | x | | x | x | | | | | x | | x |
| 52 | x | | x | x | | | | | x | | x |
| 53 | x | | x | x | | | | | x | | x |
| 54 | x | | | x | | | | | x | | |
| 55 | x | | | x | x | | | x | | | |
| 56 | x | | | x | x | | | x | | | |
| 57 | x | | | x | x | | | x | | | |
| 58 | x | | | x | x | | | x | | | |
| 59 | x | | | x | x | | | x | | | |
| 60 | x | | | x | x | | | x | | | |
| 61 | x | | | x | x | | | x | | | |
| Number active | 61 | 44 | 26 | 61 | 15 | 15 | 26 | 18 | 15 | 16 | 14 |

X Active replicate.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

Control totals were created for each replicate and for each subgroup. Variation was controlled for 10 subgroups, which identified the six SAAL states and the remainder in each census region. The replicated control totals were created by adding error to the current totals for each subgroup. The amount of error depended on the subgroup. The variation comes in the form of the GVF model and uses parameters derived from information in the technical report for the 2003 CPS March Supplement. Table 12-13 provides the GVF parameters for the GVF formula for the standard error of an estimate X related to the NAAL control totals: $SE = \sqrt{AX^2 + BX}$ , where B=1,586.

**Table 12-13. GVF parameter, by subgroup: 2003**

| Subgroup | GVF intercept parameter estimate (A) |
|---|---|
| Rest of Northeast | –0.000008 |
| Massachusetts | –0.000248 |
| New York | –0.000083 |
| Rest of Midwest | –0.000008 |
| Missouri | –0.000284 |
| Rest of South | –0.000008 |
| Kentucky | –0.000395 |
| Maryland | –0.000295 |
| Oklahoma | –0.000464 |
| West | –0.000008 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

The new control totals were created for each replicate *r* as follows:

$$X(r) = X + \frac{z * SE}{\sqrt{60}},$$

where z = a random draw from a standard normal distribution. The error term was divided by the square root of 60 because the stratified jackknife (JK2) was used.

### 12.1.7.4 Jackknife and Taylor Series Variance Estimation

After the replicate weights had been constructed, the estimate of variance could easily be computed for any statistic. The statistic must be computed 62 times, once using the full-sample weight and an additional 61 times using each of the 61 replicate weights. The variance estimate is the sum of the 61 squared differences between the estimate derived using the full-sample weight and the estimate derived using each of the 61 replicate weights. That is, the estimate of the variance of a statistic Y is as follows:

$$v(Y) = \sum_{r=1}^{61} (Y_r - Y)^2,$$

where

$Y_r$ = the weighted estimate obtained using the rth replicate weight and

Y = the weighted estimate obtained using the full-sample weight.

The data user can use the variance stratum and variance unit that were developed for the stratified jackknife replicates to compute Taylor series estimates of variance. Taylor series expansion linearizes the estimator and then uses variance estimation methods to estimate the variance of the linearized estimate. The advantage of the linearization method is that it is applicable to general sampling design. However, it requires the derivation of a separate standard error for each nonlinear statistic. In contrast, the jackknife estimator employs a single standard error formula for all statistics.

### 12.1.7.5 Evaluation of Variance Estimates

Table 12-14 provides standard errors for selected background questionnaire items. The table also shows that the process of replicating control totals, to incorporate the variance associated with CPS estimates discussed in section 12.1.7.3.2, had very little effect on variances computed for the background questionnaire items at the national level. Standard errors were checked at the state level as well, and very little change to the standard errors was observed, as was expected and desired.

**Table 12-14. Standard errors for selected background questionnaire items, by background questionnaire item: 2003**

| Background questionnaire item | Sample size | Estimated percentage | Before replicating control totals | | After replicating control totals (final) | |
|---|---|---|---|---|---|---|
| | | | Standard error | Design effect | Standard error | Design effect |
| BQ1440: How many hours do you usually watch television, videotapes, or DVDs each day? | | | | | | |
| None | 378 | 2.0 | 0.17 | 2.53 | 0.17 | 2.53 |
| 1 hour or less | 3,642 | 22.0 | 0.48 | 2.45 | 0.48 | 2.45 |
| 2 hours | 4,970 | 28.0 | 0.52 | 2.48 | 0.52 | 2.48 |
| 3 hours | 3,591 | 19.6 | 0.47 | 2.58 | 0.47 | 2.58 |
| 4 hours | 2,311 | 12.2 | 0.32 | 1.69 | 0.32 | 1.70 |
| 5 hours | 1,146 | 5.8 | 0.25 | 2.09 | 0.25 | 2.09 |
| 6 or more hours | 2,133 | 10.3 | 0.43 | 3.69 | 0.43 | 3.69 |
| BQ1620: How often do you read books in English? | | | | | | |
| Every day | 5,677 | 31.3 | 0.47 | 1.84 | 0.47 | 1.84 |
| A few times a week | 3,591 | 20.0 | 0.44 | 2.17 | 0.44 | 2.17 |
| Once a week | 1,868 | 10.4 | 0.30 | 1.79 | 0.30 | 1.79 |
| Less than once a week | 4,338 | 25.1 | 0.49 | 2.28 | 0.48 | 2.27 |
| Never | 2,699 | 13.2 | 0.36 | 2.11 | 0.37 | 2.11 |
| BQ1865: How much do you think your math skills limit your job opportunities? | | | | | | |
| A lot | 1,715 | 9.0 | 0.39 | 3.02 | 0.39 | 3.02 |
| Some | 1,789 | 10.5 | 0.44 | 3.26 | 0.44 | 3.27 |
| A little | 2,098 | 12.0 | 0.44 | 2.94 | 0.44 | 2.95 |
| Not at all | 10,572 | 68.6 | 0.69 | 3.61 | 0.70 | 3.63 |
| BQ2000: There are 25 or more books in your home right now. | | | | | | |
| True | 15,195 | 88.1 | 0.36 | 2.28 | 0.36 | 2.28 |
| False | 2,962 | 11.9 | 0.36 | 2.28 | 0.36 | 2.28 |

See note at end of table.

**Table 12-14. Standard errors for selected background questionnaire items, by background questionnaire item: 2003—Continued**

| Background questionnaire item | Sample size | Estimated percentage | Before replicating control totals | | After replicating control totals (final) | |
|---|---|---|---|---|---|---|
| | | | Standard error | Design effect | Standard error | Design effect |
| BQ2040: Including yourself, how many people in your house are employed or work for pay or wages? | | | | | | |
| None | 3,675 | 16.9 | 0.40 | 2.08 | 0.40 | 2.08 |
| One | 7,099 | 33.6 | 0.75 | 4.63 | 0.75 | 4.62 |
| Two | 5,372 | 35.2 | 0.57 | 2.63 | 0.57 | 2.63 |
| Three or more | 1,997 | 14.3 | 0.52 | 3.95 | 0.52 | 3.94 |
| BQ2255: In general, how would you rate your overall health? | | | | | | |
| Excellent | 4,392 | 25.8 | 0.56 | 2.93 | 0.56 | 2.93 |
| Very good | 6,332 | 35.7 | 0.51 | 2.09 | 0.51 | 2.08 |
| Good | 4,465 | 24.5 | 0.50 | 2.50 | 0.50 | 2.50 |
| Fair | 2,172 | 10.5 | 0.34 | 2.24 | 0.34 | 2.25 |
| Poor | 800 | 3.5 | 0.17 | 1.48 | 0.17 | 1.48 |
| BQ2270: Have you ever been diagnosed or identified as having a learning disability? | | | | | | |
| Yes | 1,006 | 5.7 | 0.28 | 2.58 | 0.28 | 2.58 |
| No | 17,144 | 94.3 | 0.28 | 2.58 | 0.28 | 2.58 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

## 12.2 WEIGHTING THE PRISON STUDY SAMPLE

The prison study weighting process consisted of four main steps. First, prison base weights were constructed using the prison selection probability (section 12.2.1). Then, a nonresponse adjustment was made to prison base weights to account for nonparticipating prisons (section 12.2.2). Next, inmate base weights were formed using the prison nonresponse-adjusted weight and the within-prison sampling rate (section 12.2.3). Finally, the inmate base weights were raked to control totals to account for inmate nonresponse and noncoverage (section 12.2.4). Section 12.2.5 provides the distribution of the final weights.

Estimates of variance can be made using replication methods (such as the stratified jackknife procedure) or Taylor series linearization. Both methods can take into account the complex sample design of the prison study. To facilitate variance estimation of the prison study outcome measures, stratified jackknife replicate weights were created. The formation of the replicates is described in section 12.2.6, and the resulting variance estimates are evaluated for some background questionnaire items in section 12.2.7. For further information on jackknife and Taylor series variance estimation, refer to section 12.1.7.4.

### 12.2.1 Computing Prison Base Weights

The prison base weights were computed as the inverse of the prison probability of selection:

$$W_i = \frac{1}{P_i},$$

where

| | | |
|---|---|---|
| $W_i$ | = | the base weight for the ith prison and |
| $P_i$ | = | the probability of selection for the ith prison. |

The distribution of the full-sample base weights for eligible prisons is shown in table 12-15. Because prisons were selected with probability proportional to size, the weights vary with the size of the prison. Table 12-15 also compares the weighted number of inmates with the count of inmates for all prisons on the frame. The weighted number of inmates was calculated using the prison base weight and the count of inmates in each eligible prison, where the inmate count was updated on the basis of information obtained from the prison contacts immediately prior to data collection.

**Table 12-15. Distribution of Prison Study prison base weights: 2003**

| Prison base weights | | | | | Weighted number of inmates | Inmate count from frame |
|---|---|---|---|---|---|---|
| *N* | Mean | Min | Max | Coefficient of variation | | |
| 110 | 12.75 | 1.71 | 84.49 | 107.95 | 1,355,833 | 1,348,458 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 12.2.2 Prison Nonresponse Adjustment

Three prisons did not participate in the prison study. To adjust for the nonparticipating prisons, prison-level variables that are known for both participants and nonparticipants from the frame were used to form nonresponse adjustment cells, and an adjustment factor was applied to participating prisons within each cell. Because all three nonparticipating prisons were state prisons in the Midwest with male inmates only, they were assigned to a cell with eligible prisons of the same region, type of prison, and gender composition. The remaining eligible prisons were assigned to a second nonresponse adjustment cell.

The nonresponse adjustment factor in each cell, $F_\alpha^{nr}$, was computed as the sum of the weighted inmate population for eligible prisons divided by the sum of the weighted inmate population for participating prisons:

$$F_\alpha^{nr} = \frac{\sum_{i \in SE(\alpha)} W_i^{base} N_i}{\sum_{i \in SP(\alpha)} W_i^{base} N_i},$$

where

$Ni$ = *the inmate population count of the ith prison;*

$SE(\alpha)$ = *the set of eligible sampled prisons in nonresponse adjustment class* $\alpha$ *; and*

$SP(\alpha)$ = *the set of participating prisons in nonresponse adjustment class* $\alpha$.

The nonresponse adjustment cells and factors are summarized in table 12-16. The second cell contained no nonparticipating prisons, so the adjustment factor in the cell is equal to 1.000.

**Table 12-16. Prison Study prison nonresponse adjustment factors, by adjustment class: 2003**

| Nonresponse adjustment cell | Eligible prisons | | Responding prisons | | Nonresponse adjustment factor |
|---|---|---|---|---|---|
| | Sample size | Weighted number of inmates | Sample size | Weighted number of inmates | |
| 1: State prisons in the Midwest with male inmates only | 21 | 265,892 | 18 | 232,238 | 1.14 |
| 2: Other | 89 | 1,089,941 | 89 | 1,089,941 | 1.00 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 12.2.3 Computing Inmate Base Weights

The full-sample inmate base weight for inmate j of prison i was computed as the product of the prison nonresponse-adjusted base weight and the reciprocal of the inmate sampling rate, as given by

$$W_{ij}^{base} = W_i^{base} F_{\alpha}^{nr} S_i \frac{1}{P_{ij}} ,$$

where

$S_i$ = the adjustment factor for the subsampling of units within the $i^{th}$ prison[6] and

$P_{ij}$ = the initial probability of selection for the $j^{th}$ inmate in the $i^{th}$ prison.

The distribution of the inmate full-sample base weights is shown in table 12-17. The variation in the weights is due to the constraint of sampling 9 to 16 inmates per prison and to differences between

[6] One prison was found to have four separate units: one minimum security unit and three reception centers. Because of the difficulty of conducting interviews in all three reception centers, one reception center was sampled from the three with probability proportional to size. The base weights of inmates in the sampled unit were adjusted by the inverse of the selection probability of the unit, $S_i$. The factor $S_i$ was set to 1 for all other prisons. Inmates were also sampled at a higher rate within the reception center to maintain the same overall selection probability.

inmate counts provided at the time of negotiations with prisons and those determined through the within-prison sampling procedure conducted during data collection.

**Table 12-17. Distribution of inmate base weights: 2003**

| | Inmate base weights | | | | |
|---|---|---|---|---|---|
| *N* | Sum | Mean | Min | Max | Coefficient of variation |
| 1,298 | 1,358,771 | 1,047 | 159 | 1,423 | 11 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 12.2.4 Accounting for Nonresponse and Noncoverage

The inmate base weights were raked to the Bureau of Justice Statistics control totals to bring estimates for selected variables to known totals and reduce bias owing to inmate nonresponse and noncoverage. Refer to section 12.1.4.2 for a description of the raking algorithm.

Sampled inmates who completed the background questionnaire were included in the raking process. Literacy-related nonparticipants (those with a language problem or mental disability) were also included because the reasons for nonparticipation are highly related to literacy results. Because raking variables must be nonmissing, the background questionnaire variables of country of birth, educational attainment, and marital status were imputed for inmates who did not complete the background questionnaire for literacy-related reasons. Because there were only 10 cases with missing data, the imputation was done by forming cells on the basis of assessment status code (language problem or mental disability) and then taking the mode of the raking variable in the cell.

Weights were raked to control totals for the following dimensions: region/type of prison, prison security level, inmate gender, race/ethnicity, age category, educational attainment, country of birth, and marital status. The variables were selected in the same manner as discussed in the household weighting process in section 12.1.4.1. The raking factors for each raking dimension category are shown in table 12-18. Raking factors ranged from 0.48 to 1.93. Domains that have a relatively large range of factors will have more variation added to the weights. A maximum of 10 iterations was allowed for the raking process.

After raking, the 3× median rule,[7] discussed in section 12.1.3.3, was used within each region to detect extreme weights. No trimming of the weights was needed.

**Table 12-18. Raking factors by raking dimension for the Prison Study, by domain: 2003**

| Domain | Sample | | Control total | Raking factor | | |
|---|---|---|---|---|---|---|
| | Size | Estimate | | Mean | Minimum | Maximum |
| Overall | 1,173 | 1,231,421 | 1,380,776 | 1.13 | 0.48 | 1.93 |
| Region/prison type | | | | | | |
| Northeast | 137 | 139,493 | 174,204 | 1.25 | 0.73 | 1.77 |
| Midwest | 209 | 247,476 | 242,955 | 0.98 | 0.48 | 1.48 |
| South | 485 | 491,669 | 530,452 | 1.08 | 0.50 | 1.53 |
| West | 206 | 208,262 | 273,890 | 1.32 | 0.66 | 1.93 |
| Federal | 136 | 144,521 | 159,275 | 1.11 | 0.57 | 1.88 |
| Prison security level | | | | | | |
| Supermax, max/close/high | 375 | 392,821 | 466,991 | 1.19 | 0.65 | 1.92 |
| Medium | 539 | 559,789 | 666,387 | 1.19 | 0.69 | 1.93 |
| Min/low, admin, other | 259 | 278,811 | 247,398 | 0.89 | 0.48 | 1.38 |
| Gender | | | | | | |
| Male | 1,097 | 1,152,051 | 1,292,354 | 1.13 | 0.48 | 1.93 |
| Female | 76 | 79,370 | 88,422 | 1.11 | 0.61 | 1.77 |
| Race/ethnicity | | | | | | |
| Hispanic | 229 | 237,362 | 251,137 | 1.06 | 0.57 | 1.66 |
| Non-Hispanic Black only | 493 | 513,458 | 628,204 | 1.23 | 0.55 | 1.93 |
| Other | 451 | 480,601 | 501,435 | 1.05 | 0.48 | 1.77 |
| Age | | | | | | |
| 16–29 | 391 | 411,132 | 513,206 | 1.26 | 0.71 | 1.93 |
| 30–49 | 666 | 698,107 | 766,270 | 1.10 | 0.61 | 1.74 |
| 50+ | 116 | 122,182 | 101,300 | 0.83 | 0.48 | 1.23 |
| Education | | | | | | |
| Less than high school | 470 | 488,881 | 526,984 | 1.08 | 0.48 | 1.81 |
| High school or higher | 703 | 742,540 | 853,792 | 1.16 | 0.54 | 1.93 |
| Country of birth | | | | | | |
| U.S. | 1,051 | 1,104,383 | 1,236,811 | 1.12 | 0.48 | 1.93 |
| Other | 122 | 127,038 | 143,965 | 1.14 | 0.74 | 1.66 |
| Marital status | | | | | | |
| Never married | 631 | 661,732 | 763,735 | 1.16 | 0.55 | 1.92 |
| Other | 542 | 569,690 | 617,041 | 1.09 | 0.48 | 1.93 |

NOTE: All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as other include non-Hispanics of all other races including multiracial. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

[7] The cutoff value of three times the median weight for each cell was used as a guideline to limit the bias introduced by trimming.

### 12.2.5 Final Inmate Weights

The final inmate weights for inmate j of prison i were computed as the product of the inmate base weight and the raking factor:

$$W_{ij}^{final} = W_{ij}^{base} F_{\beta},$$

where

$F_{\beta}$ = the raking adjustment factor for raking cell $\beta$.

The distribution of the final inmate weights is shown in table 12-19. Note that the raking process increased the coefficient of variation (CV) of the weights from 11.47 percent to 22.63 percent. The overall effect, however, was a reduction in sampling variance. Refer to section 12.2.7 for the evaluation of variance estimates.

### 12.2.6 Replicate Weights for Variance Estimation for the Prison Study

Because of the clustering of inmates within prisons, simple random sample variance formulas would underestimate sampling variability. Therefore, replicates were formed to facilitate variance estimation. The variance estimation was carried out in three steps: (1) the formation of replicates, (2) the computation of replicate weights, and (3) the estimation of variances and design effects for some survey variables.

To create the replicates, the 110 eligible prisons were sorted in their order of selection. Prisons were paired consecutively and assigned to 55 variance strata. Within each variance stratum, one prison was randomly assigned to variance unit 1 and the other to variance unit 2. Replicates were then formed using the stratified jackknife approach. The rth replicate base weight for the prison associated with variance unit k of variance stratum h was calculated as

$$W_{hk}^{base}(r) = \begin{cases} 0 & \text{if } h = r \text{ and } k = 1 \\ 2W_{hk}^{base} & \text{if } h = r \text{ and } k = 2, \text{ and} \\ W_{hk}^{base} & \text{if } h \neq r, \end{cases}$$

where

$$r = 1, 2, \ldots, 55 \text{ and}$$

$$W_{hk}^{base} = \textit{the full-sample prison base weight for the prison in unit } k \textit{ of variance stratum } h.$$

**Table 12-19. Distribution of Prison Study final inmate weights, by raking dimension: 2003**

| Domain | Inmate final weights | | | | | |
|---|---|---|---|---|---|---|
| | *N* | Sum | Mean | Min | Max | CV[1] |
| Overall | 1,173 | 1,380,776 | 1,177.13 | 127.11 | 2,103.34 | 22.63 |
| Region/prison type | | | | | | |
| Northeast | 137 | 174,204 | 1,271.56 | 740.22 | 1,797.64 | 18.70 |
| Midwest | 209 | 242,955 | 1,162.46 | 127.11 | 2,103.34 | 28.77 |
| South | 485 | 530,452 | 1,093.72 | 505.18 | 1,583.99 | 17.76 |
| West | 206 | 273,890 | 1,329.56 | 672.21 | 1,962.41 | 19.83 |
| Federal | 136 | 159,275 | 1,171.14 | 656.27 | 1,914.11 | 23.56 |
| Prison security level | | | | | | |
| Supermax, max/close/high | 375 | 466,990 | 1,245.31 | 663.13 | 2,103.34 | 20.22 |
| Medium | 539 | 666,387 | 1,236.34 | 127.11 | 1,962.41 | 20.35 |
| Min/low, admin, other | 259 | 247,399 | 955.21 | 505.18 | 1,732.04 | 19.59 |
| Gender | | | | | | |
| Male | 1,097 | 1,292,354 | 1,178.08 | 127.11 | 2,103.34 | 22.83 |
| Female | 76 | 88,422 | 1,163.45 | 616.97 | 1,684.70 | 19.61 |
| Race/ethnicity | | | | | | |
| Hispanic | 229 | 251,137 | 1,096.67 | 127.11 | 1,592.49 | 19.62 |
| Non-Hispanic Black only | 493 | 628,204 | 1,274.25 | 186.37 | 2,103.34 | 21.93 |
| Other | 451 | 501,435 | 1,111.83 | 151.70 | 1,695.07 | 21.69 |
| Age | | | | | | |
| 16–29 | 391 | 513,206 | 1,312.55 | 186.37 | 2,103.34 | 19.48 |
| 30–49 | 666 | 766,270 | 1,150.56 | 127.11 | 1,798.30 | 20.26 |
| 50+ | 116 | 101,300 | 873.28 | 505.18 | 1,246.99 | 19.17 |
| Education | | | | | | |
| Less than high school | 470 | 526,984 | 1,121.24 | 127.11 | 1,845.37 | 21.99 |
| High school or higher | 703 | 853,792 | 1,214.50 | 151.70 | 2,103.34 | 22.46 |
| Country of birth | | | | | | |
| U.S. | 1,051 | 1,236,811 | 1,176.79 | 127.11 | 2,103.34 | 22.93 |
| Other | 122 | 143,965 | 1,180.04 | 715.96 | 1,732.04 | 19.98 |
| Marital status | | | | | | |
| Never married | 631 | 763,735 | 1,210.36 | 151.70 | 1,957.92 | 20.85 |
| Other | 542 | 617,041 | 1,138.45 | 127.11 | 2,103.34 | 24.35 |

[1] Coefficient of variation.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

For each subsequent stage of weighting, adjustments made to the full-sample weights were also made to the replicate weights. As a result of these adjustments, the effect of the weighting procedures was properly reflected in variance estimates. When the stratified jackknife estimation is applied, an estimate of the survey variable is produced using the full-sample weight and is computed 55 additional times using each of the replicate weights.

### 12.2.7 Evaluation of Variance Estimates

Table 12-20 provides standard errors and design effects for selected variables. Calculations were performed in WesVar (Westat 2002) using stratified jackknife variance estimation. Estimates are for selected variables from the prison study background questionnaire as well as variables used in raking. The table compares design effects using the inmate base weights with those for the final raked weights. The maximum base weight design effect is 2.04. Raking should reduce the variance of variables correlated with the raking variables. After raking, the final design effects of the weights for the raking variables are equal to 0. The maximum design effect of the remaining selected background questionnaire variables is 1.86. In general, due to clustering of inmates within prisons, the design effects are expected to be greater than 1.0. However, due to sampling error associated with variance estimates, some design effects are less than 1.0. For the majority of the selected background questionnaire items, raking resulted in a decrease in variance.

The 55 replicates for the prison study provide sufficient degrees of freedom for stable variance estimates. Table 12-21 shows the active replicates overall and by region and race/ethnicity subdomains.

**Table 12-20. Design effects for selected Prison Study background questionnaire and raking variables, by item: 2003**

| Background questionnaire item | Inmate base weights | | | | Final weights | | | |
|---|---|---|---|---|---|---|---|---|
| | Sample size | Estimated percentage | Standard error | Design effect | Sample size | Estimated percentage | Standard error | Design effect |
| BQ1495: Before your most recent admission to prison, did you ever serve time in prison, jail, or some other correctional facility as a juvenile or adult? | | | | | | | | |
| Yes | 845 | 73.3 | 1.46 | 1.26 | 845 | 73.4 | 1.42 | 1.19 |
| No | 313 | 26.7 | 1.46 | 1.26 | 313 | 26.6 | 1.42 | 1.19 |
| BQ1635: How much do you think your math skills limit your job opportunities? | | | | | | | | |
| A lot | 187 | 16.4 | 1.10 | 1.01 | 187 | 16.2 | 0.98 | 0.81 |
| Some | 213 | 18.7 | 1.17 | 1.03 | 213 | 18.7 | 1.12 | 0.94 |
| A little | 238 | 20.8 | 1.31 | 1.18 | 238 | 21.1 | 1.31 | 1.17 |
| Not at all | 501 | 44.1 | 1.65 | 1.26 | 501 | 44.1 | 1.53 | 1.08 |
| BQ1680: How many hours do you usually watch television, videotapes, or DVDs each day? | | | | | | | | |
| None | 127 | 10.7 | 1.02 | 1.27 | 127 | 10.8 | 1.04 | 1.30 |
| 1 hour or less | 271 | 23.1 | 1.76 | 2.02 | 271 | 22.7 | 1.67 | 1.84 |
| 2 hours | 244 | 21.0 | 1.22 | 1.05 | 244 | 20.6 | 1.22 | 1.05 |
| 3 hours | 176 | 15.3 | 1.06 | 1.00 | 176 | 15.7 | 1.13 | 1.12 |
| 4 hours | 127 | 11.2 | 0.91 | 0.98 | 127 | 11.1 | 0.87 | 0.89 |
| 5 hours | 75 | 6.4 | 0.76 | 1.10 | 75 | 6.6 | 0.81 | 1.23 |
| 6 or more hours | 140 | 12.3 | 1.25 | 1.68 | 140 | 12.5 | 1.28 | 1.75 |
| BQ1755: How often do you read books in English? | | | | | | | | |
| Every day | 583 | 50.0 | 1.63 | 1.24 | 583 | 50.2 | 1.56 | 1.12 |
| A few times a week | 248 | 21.5 | 1.38 | 1.31 | 248 | 21.8 | 1.32 | 1.19 |
| Once a week | 94 | 8.1 | 0.86 | 1.14 | 94 | 8.3 | 0.90 | 1.24 |
| Less than once a week | 135 | 11.7 | 0.95 | 1.01 | 135 | 11.5 | 0.96 | 1.05 |
| Never | 100 | 8.7 | 1.11 | 1.80 | 100 | 8.2 | 0.95 | 1.39 |
| BQ1785: Have you ever used a computer? | | | | | | | | |
| Yes | 292 | 24.9 | 1.65 | 1.69 | 292 | 25.0 | 1.73 | 1.86 |
| No | 869 | 75.1 | 1.65 | 1.69 | 869 | 75.0 | 1.73 | 1.86 |

See notes at end of table.

**Table 12-20. Design effects for selected Prison Study background questionnaire and raking variables, by item: 2003—Continued**

| Background questionnaire item | Inmate base weights | | | | Final weights | | | |
|---|---|---|---|---|---|---|---|---|
| | Sample size | Estimated percentage | Standard error | Design effect | Sample size | Estimated percentage | Standard error | Design effect |
| BQ1945: In general, how would you rate your overall health? | | | | | | | | |
| Excellent | 319 | 27.4 | 1.62 | 1.52 | 319 | 28.0 | 1.66 | 1.59 |
| Very good | 399 | 34.5 | 1.84 | 1.74 | 399 | 35.0 | 1.79 | 1.63 |
| Good | 259 | 22.3 | 1.24 | 1.03 | 259 | 21.7 | 1.18 | 0.96 |
| Fair | 132 | 11.4 | 0.90 | 0.93 | 132 | 11.1 | 0.88 | 0.92 |
| Poor | 51 | 4.5 | 0.54 | 0.78 | 51 | 4.2 | 0.51 | 0.76 |
| BQ1960: Have you ever been diagnosed or identified as having a learning disability? | | | | | | | | |
| Yes | 193 | 16.8 | 1.17 | 1.13 | 193 | 16.6 | 1.05 | 0.91 |
| No | 964 | 83.2 | 1.17 | 1.13 | 964 | 83.4 | 1.05 | 0.91 |
| AGECAT_R: Inmate age category | | | | | | | | |
| 16–29 | 427 | 33.0 | 1.87 | 2.04 | 391 | 37.2 | 0.00 | 0.00 |
| 30–49 | 736 | 56.9 | 1.77 | 1.65 | 666 | 55.5 | 0.00 | 0.00 |
| 50+ | 130 | 10.1 | 0.97 | 1.34 | 116 | 7.3 | 0.00 | 0.00 |
| BQRACETH_RAKE: Inmate race/ethnicity | | | | | | | | |
| Hispanic | 245 | 18.8 | 1.44 | 1.75 | 229 | 18.2 | 0.00 | 0.00 |
| Non-Hispanic Black only | 539 | 41.3 | 1.67 | 1.49 | 493 | 45.5 | 0.00 | 0.00 |
| Other | 508 | 39.9 | 1.86 | 1.87 | 451 | 36.3 | 0.00 | 0.00 |
| EDUCCAT_R: Inmate highest level of education | | | | | | | | |
| Less than high school | 458 | 39.1 | 1.64 | 1.31 | 470 | 38.2 | 0.00 | 0.00 |
| High school or higher | 703 | 60.9 | 1.64 | 1.31 | 703 | 61.8 | 0.00 | 0.00 |

NOTE: All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as other include non-Hispanics of all other races including multiracial. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 12-21. Prison Study active replicates, by selected subdomains: 2003**

| Replicate | Total | Region | | | | Race/ethnicity | | |
|---|---|---|---|---|---|---|---|---|
| | | Northeast | Midwest | South | West | Hispanic | Non-Hispanic Black only | Other |
| 1 | x | x | | | | x | x | x |
| 2 | x | x | | | | x | x | x |
| 3 | x | x | | | | x | x | x |
| 4 | x | x | | | | x | x | x |
| 5 | x | x | | | | x | x | x |
| 6 | x | x | | | | x | x | x |
| 7 | x | x | | | | x | x | x |
| 8 | x | x | | | | x | x | x |
| 9 | x | | x | | | | x | x |
| 10 | x | | x | | | x | x | x |
| 11 | x | | x | | | | x | x |
| 12 | x | | x | | | x | x | x |
| 13 | x | | x | | | | x | x |
| 14 | x | | x | | | x | x | x |
| 15 | x | | x | | | x | x | x |
| 16 | x | | x | | | x | x | x |
| 17 | x | | x | | | x | x | x |
| 18 | x | | x | | | x | x | x |
| 19 | x | | x | | | x | x | x |
| 20 | x | | | x | | x | x | x |
| 21 | x | | | x | | x | x | x |
| 22 | x | | | x | | x | x | x |
| 23 | x | | | x | | x | x | x |
| 24 | x | | | x | | x | x | x |
| 25 | x | | | x | | x | x | x |
| 26 | x | | | x | | x | x | x |
| 27 | x | | | x | | x | x | x |
| 28 | x | | | x | | x | x | x |
| 29 | x | | | x | | | x | x |
| 30 | x | | | x | | x | x | x |
| 31 | x | | | x | | x | x | x |
| 32 | x | | | x | | x | x | x |
| 33 | x | | | x | | x | x | x |
| 34 | x | | | x | | x | x | x |

See notes at end of table.

**Table 12-21. Prison Study active replicates, by select subdomains: 2003—Continued**

| Replicate | Total | Region: Northeast | Midwest | South | West | Race/ethnicity: Hispanic | Non-Hispanic Black only | Other |
|---|---|---|---|---|---|---|---|---|
| 35 | x | | | x | | x | x | x |
| 36 | x | | | x | | x | x | x |
| 37 | x | | | x | | x | x | x |
| 38 | x | | | x | | | x | x |
| 39 | x | | | x | | x | x | x |
| 40 | x | | | x | | | x | x |
| 41 | x | | | x | | x | x | x |
| 42 | x | | | x | | x | x | x |
| 43 | x | | | x | | x | x | x |
| 44 | x | | | | x | x | x | x |
| 45 | x | | | | x | x | x | x |
| 46 | x | | | | x | x | x | x |
| 47 | x | | | | x | x | x | x |
| 48 | x | | | | x | x | x | x |
| 49 | x | | | | x | x | x | x |
| 50 | x | | | | x | x | x | x |
| 51 | x | | | | x | x | x | x |
| 52 | x | | | | x | x | x | x |
| 53 | x | | | | x | x | x | x |
| 54 | x | | | | x | x | x | x |
| 55 | x | | | | x | x | x | x |
| No. active | 55 | 8 | 11 | 24 | 12 | 49 | 55 | 55 |

NOTE: All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as other include non-Hispanics of all other races including multiracial. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

*This page is intentionally left blank.*

# CHAPTER 13

# SCORING

Justin Baer, American Institutes for Research, Jared Bernstein, Ordinate Corporation,
and Michelle Amsbary, Westat

This chapter describes the procedures followed for scoring the three main components of the 2003 assessment: the cognitive items, the Fluency Addition to NAAL (FAN), and the Adult Literacy Supplemental Assessment (ALSA). For the cognitive items and the ALSA, the scoring procedures used were similar to the procedures implemented for scoring the NAAL field test (chapter 4). Scoring the FAN was more complex because the scores were generated by an automatic speech recognition (ASR) system. To ensure the validity of the FAN data, a sample of tasks scored by the ASR were compared to a sample of tasks scored by human scorers.

## 13.1 COGNITIVE ITEMS

### 13.1.1 Refinement of Training Materials

The scorer training materials for the 2003 main study assessment were largely adapted from those used for the NAAL field test (see chapter 4). Prior to the scoring of the NAAL field test, AIR staff compiled sample responses to each of the cognitive items in the field-test assessment booklets. In developing the training materials for the 2003 assessment, AIR staff began by reviewing the inter-rater reliability statistics for the field-test items selected for the main study.[1] Sample responses to items with high inter-rater reliability during the field test were reused as scorer training materials. For items with low inter-rater reliability and those that were challenging to score, AIR conducted additional range finding to locate both straightforward and ambiguous responses. During range finding, AIR staff reviewed a sample of booklets that had been returned from the field to the data collection facility, searching for responses that would be valuable for training purposes.

Sample responses to the targeted items were then photocopied and compiled with the existing sample responses to items from the field test with high inter-rater reliability. The complete collection of sample papers consisted of a mixture of responses that closely matched the scoring rubrics and more ambiguous responses to items. Because the sample responses were to be used as scorer training papers, the number of example papers selected also varied across the items on the basis of their scoring difficulty.

[1] See chapter 4 (section 4.2.2.2) for a discussion of procedures for items with low inter-rater reliability. Inter-rater reliability statistics for field test items are also presented in chapter 4.

For example, fewer examples of responses to quantitative questions were photocopied because correct answers were typically numeric and simple to score.

As noted in chapter 2, 6 blocks of items from the 1992 National Adult Literacy Study (NALS) were included in the 2003 NAAL. When items are re-used across assessments, trend scoring is usually conducted to ensure that the common items are scored consistently from one assessment to the other. Unfortunately, with the exception of the scoring rubrics, none of the training material from the 1992 survey was available. Consequently, trend scoring between the 1992 NALS and the 2003 NAAL could not be conducted. To ensure consistency in scoring, AIR staff consulted with a project member from the 1992 survey who reviewed training papers for the 1992 items reused in 2003. Her comments and interpretation of the 1992 scoring rubrics were carefully documented so they could be included in the materials used to the train the scorers.

For the new blocks developed for the 2003 assessment, AIR convened an expert panel to review the scores assigned to the training papers.[2] The panel was directed to closely review the rubrics and training papers for items with low inter-rater reliability in the field test. This meeting also provided a final opportunity to review and modify the scoring rubrics for the new 2003 blocks. On the basis of decisions made by the expert panel, several of the scoring rubrics were revised further. Comments from the panel about how responses to particular assessment items should be scored were also documented and included in the scorer training materials.

### 13.1.2 Scorer Training

Once the scores assigned to the training papers had been reviewed and agreed on by the members of the expert panel, the papers were compiled in training binders. For each item, one or two training papers were designated anchor papers. Anchor papers were straightforward responses to assessment questions that clearly corresponded to the rubrics. The remaining training papers were a mix of straightforward and more challenging responses designed to expose scorers to the range of responses they might encounter and to make certain that they demonstrated sufficient aptitude to score the assessment.

Scoring was conducted at the scoring contractor's facility in Tucson, Arizona. A total of 142 scorers were hired to score the exercise booklets; all scorers were required to hold a bachelor's degree from a college or university. Each scorer was assigned to a table that was responsible for scoring a

[2] The panelists had expertise in reading instruction and assessment as well as the alignment of curriculum, instruction, and assessment.

specific assessment block, for a total of 13 tables with 9 to 13 scorers per table. A scoring supervisor with previous scoring experience was also assigned to each table to answer questions and monitor the scoring. The 13 scoring supervisors all had previous experience working at the scoring center and had demonstrated proficiency scoring assessments. AIR staff trained the scoring supervisors, who in turn trained the scorers.

To guide the scorer training, each item was accompanied by a script written to incorporate a discussion of the items as well as a review of the training papers. The script for each item described the purpose of the question and explained how the item should be scored. For items for which a correct answer could be expressed a variety of ways, the scripts advised the scorers that the rubrics and training papers did not include every possible correct response. For these items, the scorers were instructed to use the rubrics and training papers as a guide for determining whether a response should be scored as correct.

Scorers were given time to read the item rubrics before reviewing and discussing the anchor papers as a group. Following the review, the scorers independently scored the training papers, which were then discussed by all the scorers assigned to a particular block. NAAL staff monitored the training to make certain that the scorers correctly scored the training papers. NAAL staff remained at the scoring center to answer questions during scorer training and through the first two days of scoring and were available by phone for consultation until scoring concluded.

### 13.1.3 Scoring Procedures and Quality Control

Scoring of the household assessment began in late January 2004 and was completed by mid-February 2004. The assessment booklets were rotated among the tables of scorers, with each table scoring its assigned block and then passing the booklets along to another table for scoring. To ensure consistency in scoring, half the booklets were rescored by two different scorers. In addition, the scoring supervisor for each block "backread" approximately 10 percent of all item responses scored by each rater. During "backreading," the scoring supervisor compared the scores awarded by one rater to the scorers awarded by a second rater to the same booklet. The scoring supervisor then discussed any discrepancies with the individual scorers as well as with the other scorers at the table.

Following the end of scoring for each day, the scoring contractor generated inter-rater reliability statistics for the sample of rescored booklets. Reliability was measured as the frequency of agreements between two scorers for each item scored. NAAL staff reviewed the reliability statistics daily for each assessment item to make certain that the rubrics were applied consistently across scorers. The number of possible score points and final inter-rater reliability for each item are summarized in table 13-1. The number of score points per item ranged from 2 to 10. Of the 153 items, all but 3 had inter-rater reliability agreement greater than 95%.[3] The average reliability of the three remaining items was 94%. For the entire pool of items in the assessment, the final inter-rater reliability across all items in the household and prison samples, including the core, was 99%. Inter-rater reliability ranged from a low of 92.6% (item N011101) to a high of 100% (items CC001, C030601, and C060101).

In addition to scoring the 13 blocks of assessment items, the scorers also scored the 7 core items originally scored in the field by the interviewers. In the field, performance on the core was used to assign respondents to either the main assessment (NAAL) or the supplementary assessment (ALSA). Although the determination of whether a respondent should be assigned to NAAL or to ALSA had already been made in the field, the NAAL scorers rescored the core items in each assessment booklet. As noted in table 13-1, the inter-rater agreement between the scorers for each of the 7 items was greater than 99%. The core scores from the NAAL scorers, rather than the interviewer core scores, were used when the data were scaled and proficiency scores were generated. Core scores from the NAAL scorers were used because scoring consistency could be assessed through the inter-rater reliability statistics calculated for all items scored by the scoring contractor staff.

**Table 13-1. Inter-rater reliability statistics for NAAL household and prison items, by block: 2003**

| Block | Item | Score points | Reliability |
|---|---|---|---|
| Core | CC001 | 2 | 100.0 |
| | CC002 | 2 | 99.7 |
| | CC003 | 2 | 99.4 |
| | CC004 | 2 | 99.5 |
| | CC005 | 2 | 99.6 |
| | CC006 | 2 | 99.6 |
| | CC007 | 2 | 99.6 |
| | | | |
| Block 1 | N010101 | 2 | 98.6 |
| | N010201 | 2 | 98.6 |
| | N010301 | 2 | 96.9 |
| | N010401 | 2 | 99.6 |
| | N010501 | 3 | 98.1 |
| | N010601 | 2 | 97.1 |
| | N010701 | 3 | 97.1 |
| | N010801 | 2 | 98.1 |
| | N010901 | 2 | 97.0 |
| | N011001 | 4 | 98.0 |
| | N011101 | 10 | 92.6 |
| | | | |
| Block 2 | C020101 | 2 | 98.4 |
| | C020201 | 2 | 98.3 |
| | C020301 | 5 | 99.0 |

See notes at end of table.

[3] This count is based on rounding the inter-rater reliability statistics to the nearest integer.

**Table 13-1. Inter-rater reliability statistics for NAAL household and prison items, by block: 2003—Continued**

| Block | Item | Score points | Reliability |
|---|---|---|---|
| | C020401 | 2 | 98.1 |
| | C020501 | 2 | 98.0 |
| | C020601 | 3 | 99.0 |
| | C020701 | 6 | 98.3 |
| | C020801 | 8 | 98.2 |
| | C020901 | 3 | 97.3 |
| | C021001 | 2 | 99.0 |
| | C021101 | 2 | 98.6 |
| Block 3 | C030101 | 3 | 97.2 |
| | C030201 | 2 | 99.9 |
| | C030301 | 3 | 97.3 |
| | C030401 | 2 | 99.9 |
| | C030501 | 3 | 98.9 |
| | C030601 | 2 | 100.0 |
| | C030701 | 3 | 98.1 |
| | C030702 | 3 | 97.3 |
| | C030703 | 3 | 99.6 |
| | C030704 | 3 | 99.1 |
| | C030705 | 3 | 98.9 |
| | C030706 | 4 | 96.0 |
| | C030707 | 3 | 99.6 |
| | C030708 | 3 | 99.4 |
| | C030709 | 3 | 98.7 |
| Block 4 | C040101 | 2 | 99.8 |
| | C040201 | 2 | 99.9 |
| | C040301 | 2 | 97.2 |
| | C040401 | 4 | 98.4 |
| | C040501 | 4 | 94.4 |
| | C040502 | 2 | 99.2 |
| | C040503 | 3 | 98.3 |
| | C040504 | 2 | 98.9 |
| | C040601 | 2 | 98.8 |
| | C040701 | 3 | 96.0 |
| | C040801 | 6 | 98.0 |
| Block 5 | C050101 | 2 | 99.7 |
| | C050201 | 3 | 99.0 |
| | C050301 | 3 | 99.3 |
| | C050401 | 3 | 98.4 |
| | C050501 | 2 | 99.4 |
| | C050601 | 4 | 99.0 |

See notes at end of table.

**Table 13-1. Inter-rater reliability statistics for NAAL household and prison items, by block: 2003—Continued**

| Block | Item | Score points | Reliability |
|---|---|---|---|
| | C050701 | 3 | 98.9 |
| | C050801 | 3 | 99.2 |
| | C050901 | 2 | 94.6 |
| | C051001 | 2 | 94.7 |
| | C051101 | 2 | 96.6 |
| Block 6 | C060101 | 2 | 100.0 |
| | C060201 | 2 | 99.3 |
| | C060301 | 4 | 99.4 |
| | C060401 | 3 | 99.3 |
| | C060501 | 2 | 99.3 |
| | C060601 | 2 | 99.5 |
| | C060701 | 6 | 99.4 |
| | C060801 | 4 | 99.6 |
| | C060901 | 2 | 97.2 |
| | C061001 | 3 | 98.9 |
| | C061101 | 4 | 97.3 |
| Block 7 | C070101 | 2 | 99.7 |
| | C070201 | 2 | 99.2 |
| | C070301 | 4 | 99.4 |
| | C070401 | 2 | 99.1 |
| | C070501 | 2 | 99.8 |
| | C070601 | 4 | 99.6 |
| | C070701 | 3 | 98.9 |
| | C070801 | 6 | 99.5 |
| | C070901 | 2 | 98.9 |
| | C071001 | 2 | 99.8 |
| | C071101 | 3 | 97.4 |
| Block 8 | C080101 | 3 | 98.9 |
| | C080201 | 2 | 99.0 |
| | C080301 | 2 | 99.2 |
| | C080401 | 2 | 99.1 |
| | C080501 | 2 | 99.8 |
| | C080502 | 3 | 99.4 |
| | C080503 | 3 | 99.4 |
| | C080504 | 2 | 99.8 |
| | C080601 | 2 | 99.6 |
| | C080701 | 4 | 98.2 |
| | C080801 | 2 | 99.6 |

See notes at end of table.

**Table 13-1. Inter-rater reliability statistics for NAAL household and prison items, by block: 2003—Continued**

| Block | Item | Score points | Reliability |
|---|---|---|---|
| Block 9 | N090101 | 3 | 98.2 |
| | N090201 | 3 | 97.0 |
| | N090301 | 2 | 97.7 |
| | N090401 | 2 | 98.0 |
| | N090501 | 2 | 97.5 |
| | N090601 | 2 | 98.3 |
| | N090701 | 2 | 99.5 |
| | N090801 | 2 | 97.7 |
| | N090901 | 3 | 95.7 |
| | N091001 | 5 | 98.6 |
| Block 10 | N100101 | 2 | 99.9 |
| | N100201 | 2 | 99.9 |
| | N100301 | 2 | 99.9 |
| | N100401 | 3 | 97.1 |
| | N100501 | 2 | 99.6 |
| | N100601 | 3 | 99.6 |
| | N100701 | 3 | 97.5 |
| | N100801 | 2 | 99.8 |
| | N100901 | 2 | 99.8 |
| | N101001 | 2 | 99.8 |
| Block 11 | N110101 | 2 | 98.9 |
| | N110201 | 2 | 98.7 |
| | N110301 | 4 | 97.6 |
| | N110302 | 5 | 99.2 |
| | N110303 | 6 | 98.7 |
| | N110401 | 3 | 98.1 |
| | N110501 | 2 | 98.8 |
| | N110601 | 3 | 97.4 |
| | N110701 | 2 | 99.5 |
| | N110801 | 3 | 99.2 |
| | N110901 | 3 | 96.2 |
| Block 12 | N120101 | 2 | 98.3 |
| | N120201 | 2 | 98.8 |
| | N120301 | 2 | 97.8 |
| | N120401 | 3 | 93.5 |
| | N120501 | 2 | 97.9 |
| | N120601 | 3 | 99.6 |
| | N120701 | 4 | 99.0 |
| | N120801 | 4 | 98.4 |

See notes at end of table.

**Table 13-1. Inter-rater reliability statistics for NAAL household and prison items, by block: 2003—Continued**

| Block | Item | Score points | Reliability |
|---|---|---|---|
| | N120901 | 2 | 99.7 |
| | N121001 | 3 | 97.9 |
| | N121101 | 4 | 98.5 |
| Block 13 | N130101 | 5 | 98.0 |
| | N130102 | 4 | 99.6 |
| | N130103 | 5 | 99.8 |
| | N130104 | 5 | 99.6 |
| | N130201 | 2 | 99.6 |
| | N130301 | 4 | 97.2 |
| | N130401 | 2 | 99.0 |
| | N130501 | 6 | 99.5 |
| | N130601 | 4 | 99.5 |
| | N130701 | 6 | 95.7 |
| | N130801 | 2 | 99.8 |
| | CN130901 | 6 | 99.4 |
| Mean | † | 3 | 98.6 |

†Not applicable.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 13.1.4 Scoring the Prison Sample

Although the scoring of the household assessment was completed by February 2004, data collection for the prison sample continued through July 2004. Because the household and prison assessments could not be scored concurrently, special provisions were implemented to make certain that the prison sample was scored consistently with the household sample. High scoring reliability between the two samples was especially important because the household and prison assessments were to be combined into a single reporting sample.

To maintain consistency in scoring, the same scorers recruited for the household sample were hired to score the prison sample and were assigned to the same block they had scored previously. Prior to the scoring of the prison booklets, the scoring supervisors for each block spent several hours reviewing the training materials with their table. Additionally, a sample of 882 household booklets scored earlier in the year was drawn for rescoring. The 882 household booklets were divided equally among the 26 combinations of assessment items, for a total of approximately 34 booklets for each combination of items. The booklets selected were stratified on the basis of their total booklet score, ensuring that the rescoring sample comprised assessments from respondents with varying levels of literacy.

The booklets were distributed to the scoring tables where each table scored its assigned block. Once each booklet had been scored, the scores assigned during the rescoring were compared with the original scores assigned in January and February 2004. Consistency in scoring was evaluated by examining the frequency with which the second scores agreed with the first scores, similar to an inter-rater reliability statistic. NAAL staff reviewed the statistics for each item and discussed items with low agreement (less than 95%) with the scorers. Once the review of training materials was completed, scoring began for the prison sample. Similar to the household sample, half of all assessment booklets were rescored to monitor inter-rater scorer agreement. With the exception of booklets from California prisons (discussed below), all prison assessments were scored by August 2004.

Assessments from California prisons could not be scored concurrently with those from other prisons because data collection in California was extended into the summer of 2004. Because all other assessments for the prison sample had already been scored by August 2004, the sample of 116 assessments from California prisons was scored directly by the two AIR staff members who developed the scoring materials and supervised the household and prison scoring.

AIR staff followed procedures for scoring the California assessments that were similar to those implemented to score the other assessments in the prison sample. First, to ensure consistency in scoring, AIR scorers independently rescored a sample of booklets and compared their scores against each other. The California prison assessments were then scored, with half the booklets randomly selected to be rescored to check the inter-rater scoring reliability. On completion of scoring, the California assessments were then combined with the remaining prison sample booklets. Scoring of the California prison assessments was completed in September 2004.

## 13.2 FLUENCY ADDITION TO NAAL (FAN)

### 13.2.1 Background

To evaluate the level of oral reading fluency of NAAL participants, the Fluency Addition to NAAL (FAN) was administered. For the FAN, each respondent read aloud from lists and passages of text.[4] The oral reading responses were digitally recorded and subsequently analyzed for measures of accuracy and fluency (accurate reading rate). Each list and each passage was digitized and saved to a separate audio file referred to as a response recording. The response recordings were then sent to a scoring contractor for machine scoring. Due to technical difficulties with the FAN software and

[4] Additional information about the background and design of the FAN is presented in chapter 2.

associated hardware, as well as transmission issues between the data collection and scoring contractors, there were approximately 40 cases for which primary machine scores are not available.

### 13.2.2 Scoring the FAN Data

To automatically score the FAN responses, each respondent's oral reading of the FAN material was first digitally recorded during the FAN portion of the NAAL assessment. Respondents read into a microphone that was connected to a laptop computer. Special audio recording software from the contractor for the fluency assessment was installed on each laptop and allowed the audio from the microphone to be digitized and recorded. The response recordings were then downloaded from the computer and sent to the contractor for automatic scoring.

The first step in scoring the FAN recordings using an automatic speech recognition (ASR) system was the development of a language model rule set. As part of the language model training, professional transcribers were hired to transcribe about 1,000 responses for each target item in the FAN and also all data used to validate the ASR system. A transcription is a string of words and symbols that represent the recorded response of the respondent. From these transcriptions, words were identified that were not already in the dictionary of the FAN contractor's speech recognition system. Pronunciations were created for missing words and were inserted into the dictionary.

The transcriptions were then divided into two sets: a training set of 4,681 responses and a test set of 2,170 responses. The training set was used to build the language model rule set and the test set was used to test it. The two sets did not intersect. The transcriptions were tagged for part of speech. For example, the word *dog* was tagged as a singular noun and the word *the* was tagged as a determiner. The actual tags used in this process were from the well-established Penn Treebank Tag Set (Marcus et al. 1993) and were rich in grammatical information. The tagging was done such that specific word-level rules, such as *trained* goes to *train* and *visited* goes to *visit*, could be generalized as "Verb-ed goes to Verb."

After the tagging was completed, a preliminary language model was created for each passage. (The language model is specific to each passage or list). This preliminary language model consisted of the answer choice or the text of the passage, for example, the string of words "Curly is my big black dog …" and a small set of rules. These rules were written by linguists and were intended to account for as many reading errors and disfluencies as possible. Then all the transcriptions for the passage were run through this language model. When the transcription found a path through the language model, an output file recorded any rules that fired so that the information could be reviewed by a linguist and modified if necessary. Specifically, more descriptive rules, such as “Noun goes to Noun-s,” were added to account for more reading errors. If a transcription could not be matched to a path through the language model, a new rule was written to account for it. This process was iterated until the number of transcriptions accounted for was maximized.

After the language models were created, the responses were machine scored. The speech recognition engine was used to identify the string of words that best matched the speech in each response recording. An alignment algorithm was then applied to the string of words to determine the number of omissions, substitutions, and insertions. A subset of respondents completed the Adult Literacy Supplemental Assessment (ALSA)[5] instead of the main NAAL study. For ALSA respondents, all the response recordings were transcribed, and the alignment algorithm was applied to the human transcriptions instead of to the output of the speech recognizer. On the basis of a scoring algorithm, the number of reading errors was tallied and weighted to produce the number of words read correctly for each response recording.

To evaluate the validity of the scores generated by the ASR system, a sample of scores from the ASR system was compared to a sample of scores from human scorers. The sample was comprised of recordings from 480 respondents. To ensure the validity of the ASR system across key population groups, the sample was stratified by performance on the NAAL cognitive tasks (as well as respondents who completed the Adult Literacy Supplemental Assessment) and by the following linguistic/ethnic groups: Black adults, Spanish-speaking adults, and Other English-speaking adults. The validity analyses were performed on each of the passage included in the FAN as well as all the three English word lists. The final correlations between the ASR system scoring and the human scoring of the same tasks are presented in table 13-2.

[5]Respondents were administered either the main NAAL survey or the ALSA. The decision was based on the respondent’s performance on a set of screening items. The ALSA used concrete stimulus materials and visual input to support the assessment of the least-literate adults.

**Table 13-2. Correlations between human ratings and machine scores of number of words read correctly, by Fluency Addition to the NAAL (FAN) task: 2003**

| Task | Correlation |
|---|---|
| Passages | |
| 1st Grade 3 passage | .96 |
| 2nd Grade 3 passage | .98 |
| 3rd Grade 3 passage | .99 |
| 1st Grade 8 passage | .98 |
| 4th Grade 3 passage | 1.00 |
| 2nd Grade 8 passage | 1.00 |
| 3rd Grade 8 passage | .99 |
| 4th Grade 8 passage | 1.00 |
| | |
| Word lists | |
| 1 | .98 |
| 2 | .99 |
| 3 | .98 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 13.2.3 FAN Data Products

Following the completion of scoring, six primary data products were generated for each response recording:

1. Span summary
   - index of first word attempted (first)
   - index of last word attempted (last)
   - number of words read correctly (nwordcorr)
   - narrow time in centiseconds (narrowt)
2. Articulation rate (pros)
3. Number of short pauses (nsp)
4. Number of long pauses (nlp)
5. Number of words completely deleted (ndel)
6. Number of false starts (fstart)

#### 13.2.3.1 Span Summary

In the source text, each word has an index value that corresponds to its sequential position in the text. The first word is 0, the next word is 1, the next is 2, and so on. The index values of the first word attempted and the last word attempted by the respondent are the first two items in the span summary.

The number of words read correctly is the third item in the span summary. This number is an estimate of the number of words the respondent read correctly in the source text and is optimized to match human ratings of reading accuracy.

The last number in the span summary is narrow time. Narrow time is the time from the onset of the first word spoken in the response recording through the offset of the last word spoken. The value is in centiseconds (e.g., 6000 centiseconds equals one minute).

Note that only the span summary was analyzed as part of the validation of the ASR system. The other measures listed below are for research purposes only.

#### 13.2.3.2 Articulation Rate

The articulation rate, or phonemic rate of speech (pros), is defined as the number of phonemes per second of speech. The articulation rate is computed by counting the number of phonemes in the response and dividing by the total speech duration. Total speech duration is the sum of the elapsed time of the relevant spoken material and does not include inter-word pauses or leading or trailing silence in the response file. For the purposes of computing pros, all filler material (e.g., hesitations, mouth noise) is treated as a part of inter-word pauses and is not included in the total duration of speech.

#### 13.2.3.3 Number of Short Pauses

The number of short pauses (nsp) is the count of pauses with durations that are greater than 200ms but less than 1000ms. The pause duration is a measurement of the time between the speech sounds of two contiguous words. The pauses at sentence boundaries are treated the same as pauses elsewhere in the respondent's speech.

#### 13.2.3.4 Number of Long Pauses

The number of long pauses (nlp) is the count of pauses with durations equal to or greater than 1000ms. The duration is computed from the end of one word to the beginning of the next. As with nsp, pauses at sentence boundaries are treated the same as all other pauses.

#### 13.2.3.5 Number of Deletions

The number of deletions (ndel) is the number of words completely omitted in the oral reading response (between the first word and last word attempted) in addition to the words that were deleted because of a substitution.

The process of aligning the most likely hypothesis of what the respondent said with the source text produces an estimate of the number of deletions, insertions, and substitutions present in the word sequence. From this estimate, the number of deletions is added to the number of substitutions to produce the total ndel value.

#### 13.2.3.6 False Start Count

The false start count (fstart) is the count of how many times a respondent backed up in the text to repeat (or attempt to repeat) previous words. Even if the text was read incorrectly during either the first attempt or the repeat, the event is identified as a false start as long as there is enough correct content to not categorize the event as a substitution or insertion of irrelevant material.

The false start count is extracted from the path through the language models. The language model for each FAN task encompasses information about the errors and disfluencies that a respondent is likely to make when reading a text. These errors and disfluencies include such things as substitutions, insertions, and false starts, and they are referred to as "rule firings." When the FAN contractor's system determines the best path through the language model that matches a respondent's utterance, the system also tracks which rules are fired. The false start count is the number of times the false start rule fired in the language model to traverse the best path for a given response recording.

#### 13.2.3.7 Secondary Machine Scores

Most language models for passages contain many rules. As with the false start count, these rules can provide information about the type of disfluencies and errors that are made as a respondent reads a passage. The secondary machine scores simply list the number of times each of the most frequently Occurring Rules Was Fired For A Given Response Recording Of A Passage Reading.

## 13.3 ADULT LITERACY SUPPLEMENTAL ASSESSMENT

### 13.3.1 Background and Rationale

The Adult Literacy Supplemental Assessment (ALSA) was developed as an alternative assessment to the main exercise used in the National Assessment of Adult Literacy (NAAL) in an effort to gather as much information as possible about adults with limited English literacy skills. (See section 2.6 for a complete discussion of the development and content of ALSA.) On the basis of a respondent's performance on the seven core items in the exercise booklet, an algorithm in the computer-assisted personal interviewing (CAPI) system was implemented to determine whether the respondent should continue with the main exercise or be directed to ALSA. The ALSA tasks allowed adults with limited literacy skills to demonstrate their abilities to understand and use printed materials in ways that the traditional NAAL exercise items do not.

The algorithm used to determine whether respondents should complete the main exercise tasks or ALSA was incorporated into the CAPI interviewer guide. Interviewers were trained to score the core assessment items (as explained in section 13.3.2) immediately after the respondent completed the items. On the basis of the interviewer's evaluation of whether the core item responses were correct, incorrect, or not provided, the CAPI system indicated which assessment to administer. The algorithm took into account the response to the core item and the language in which the core items had been administered (English or Spanish). The algorithm directed the respondent to ALSA under three scenarios:

- Core items CC001 through CC006 were all wrong or were not answered, and the items were administered in English.
- Core item CC007 was wrong or was not answered and was administered in English, and no answer was provided for core items CC003 and CC004.
- The core items were administered in Spanish, fewer than five of core items CC001 through CC006 were correct, and core item CC007 was wrong or not answered.

Under all other circumstances, the interviewer was instructed to continue with the main assessment booklet tasks.

### 13.3.2 Field Scoring of Core Items to Identify Respondents for the Adult Literacy Supplemental Assessment

As mentioned above, responses to the seven core items were scored in the field by the interviewers. Immediately following the completion of these seven items, the CAPI interviewer guide instructed the interviewer to take the assessment booklet from the respondent. The interviewer guide then led the interviewer through the scoring process for each core item. For each task, the interviewer guide provided the interviewer with the acceptable response(s). After a quick review of the respondent's answer, the interviewer determined whether it was acceptable, was not acceptable, or had been left blank and entered this information into the CAPI system.

On the basis of the interviewer's determination of the core responses, and the language in which the core items were administered, the algorithm was invoked to determine whether the interviewer should administer ALSA or the main exercise.

#### 13.3.2.1 Interviewer Training

Three main household study training sessions were devoted to the scoring of the core items. The scoring determined whether NAAL or ALSA would be administered and was a focus of the assessment training sessions. It was crucial that interviewers apply the scoring guidelines uniformly to ensure that respondents were not routed to the wrong assessment.

Included in the core scoring training session were interactive exercises with examples, a thorough discussion of correct and incorrect answers to each core item, and practice exercises that used actual responses from the 2001 NAAL field-test booklets.

Interviewers were trained to give respondents the benefit of the doubt while still following the scoring guidelines in the interviewer guide. It was understood that the scoring rules in the interviewer guide could not anticipate every possible answer found in the field.

#### 13.3.2.2 Quality Control

As part of the standard NAAL quality control procedures, all seven core items in the completed assessment booklets returned to the home office of the data collection contractor were rescored by trained home office staff members. The validation results were entered into a specially designed core scoring program. The program compared the interviewer's scoring with that of the home office staff, enabling supervisors to provide interviewers with feedback on their performance.

Early in the data collection period, the core items were rescored for 100 percent of the receipted assessment booklets. When interviewers were determined to be proficient at scoring the core items (based on an 85 percent match between the interviewer and the home office staff), no further core validation was conducted for those interviewers. Home office staff continued to rescore 100 percent of the core items for the other interviewers.

The in-house validation of the core items continued through the middle of January 2004, as the end of data collection approached. As table 13-3 shows, home office staff rescored a total of 13,608 core assessments. Of the rescored core assessments, only 115 discrepancies (0.85 percent) between interviewer and home office scores resulted in misclassification; that is, the rescoring effort assigned the case to a different ALSA/main assessment route than the one determined through the interviewer's scoring. Therefore, although there was significant disagreement between the home office and interviewer scoring (19 percent), the respondent actually completed the incorrect assessment in fewer than 1 percent of these cases.

Core item CC004 produced the largest number of scoring discrepancies between the home office staff and the interviewers – slightly more than 1,000 discrepancies. The item required the respondents to underline a sentence in the 5-paragraph long stimulus material. There were two possible correct sentences, as well as text in both sentences that was optional. This ambiguity likely led to discrepancies in the application of the scoring rubrics.

### 13.3.3 Recording of Responses

The administration of ALSA required a higher level of interaction between the interviewer and the respondent than did the main assessment booklet. For the main assessment, interviewers were responsible only for guiding the respondents through the items. However, for ALSA, the interviewers read each question to the respondent and classified the response.

**Table 13-3. Summary of discrepancies in core item scoring between interviewers and home office staff, by core item: 2004**

| Core item | Number |
|---|---|
| Number of core assessments rescored | 13,608 |
| Number (%) of discrepancies | 2,534 (19%) |
| Number of discrepancies by item number | |
| CC001 | 169 |
| CC002 | 371 |
| CC003 | 690 |
| CC004 | 1,057 |
| CC005 | 324 |
| CC006 | 464 |
| CC007 | 444 |
| Number (%) of discrepancies resulting in misclassification | 115 (0.85%) |
| Rescored as ALSA | 49 |
| Rescored as NAAL | 66 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

ALSA was designed so that the response categories would be easy to find and the classification rules easy to follow. Response categories for each item were enclosed in a box. Each response box contained a question to help interviewers determine which response classification to select, such as "WHAT DID SP SAY?" The interviewer selected the response category that most accurately reflected the respondent's answer or action in response to each item.

#### 13.3.3.1 Interviewer Training

Interviewers were given extensive training in the administration of the ALSA instrument. A training DVD was developed to ensure that all interviewers received the same standardized training. Training concentrated on how to accurately classify respondent answers, follow skip patterns, use nondirective probing techniques, and gain cooperation, as well as the appropriate use of stimulus materials.

As part of a certification exercise at the end of the DVD, interviewers listened to a respondent provide an answer and then classified it on the questionnaire. This exercise was collected during training, and the results were reviewed with the interviewers.

Because the ALSA interview was administered in a small sample of cases, most interviewers did not administer it on a regular basis. Therefore, the interviewers were required to review the training DVD throughout the data collection period as refresher training.

#### 13.3.3.2 Quality Control

The ALSA booklets were receipted, reviewed, and edited in the home office. Trained staff reviewed notes written in the margins and ensured that every questionnaire item had a valid response and that the skip patterns had been followed correctly by the interviewer.

#### 13.3.3.3 Data Entry of Questionnaires

Completed ALSA booklets were entered by  the data collection contractor's data entry staff. Once all responses had been keyed, the codebooks and frequencies generated from the data were reviewed for accuracy and completeness and then reconciled. The final, clean data set was submitted to the American Institutes for Research (AIR) for analysis at the end of the field period.

*This page is intentionally left blank.*

# CHAPTER 14

# ITEM ANALYSIS, SCALING, AND ESTIMATES OF SUBPOPULATION PROFICIENCIES

Stéphane Baldi and Justin Baer, American Institutes for Research[1]

The National Assessment of Adult Literacy (NAAL) scales are reported on the same three proficiency scales—prose, document, and quantitative—used for the 1992 National Adult Literacy Survey (NALS). This chapter describes the procedures and models used to conduct item analysis, scale the results, estimate respondents' proficiency, and conduct statistical analyses.

## 14.1 PROCEDURES AND QUALITY CONTROL

After the assessment booklets were scored by a contractor, the scored data were sent to NAAL staff for item analysis and scaling. To ensure the accuracy of the item analysis and scaling, NAAL staff implemented two key quality control steps. First, as described below, all analyses were conducted by two independent teams of NAAL staff. Second, all scaling activities were performed using two software packages, PARSCALE (Muraki and Bock 2003) and AM (Cohen et al. 2000). PARSCALE is a software package capable of performing item response theory (IRT) scaling and scoring or rating scale data. AM is a statistical software package capable of IRT scaling and analyzing data from complex samples, especially large-scale assessment data such as the NAAL. Using both PARSCALE and AM allowed the analysis staff to check the reliability of the estimated item parameters and to make certain that the estimates were consistent regardless of the software package employed.

### 14.1.1 Analysis Teams

To provide independent verification of the results, all item analysis and scaling tasks were completed by two independent teams of analysts. Within the teams, analysts were allowed to consult with one another and compare results. For the item analysis, the two teams ran their analyses and then submitted them to a research assistant to compare the results. The research assistant flagged any inconsistencies between the two sets of analyses, which were then resolved in joint team meetings with senior project staff. The final set of item analysis statistics was verified by both analysis teams.

[1] Eugene Johnson contributed material to an early draft of this chapter.

For item analysis and scaling, one team scaled the assessment items using PARSCALE, while the other scaled the items using AM. The analysis teams and senior project staff jointly reviewed the item statistics and item parameters generated by PARSCALE and AM to ensure consistency. In the end, the item parameters generated by AM were used instead of those generated by PARSCALE, because of AM's greater precision in estimating standard errors.

## 14.2 ITEM ANALYSIS

NAAL staff calculated p-values and adjusted biserial/polyserial correlations between an item and the total booklet score in which the item appeared (with the item excluded from the score). P-values of the NAAL items can be found in appendix E. Following the quality control steps, all item analyses were conducted by two independent research teams.

On the basis of the examination of the item analysis statistics, one item, C060401, was dropped prior to scaling. Because of an error in the reproduction of the stimulus material accompanying the item, the item became much more difficult to interpret. A substantial number of respondents skipped the item (11%), far more than the number of respondents who omitted the preceding and succeeding items (5% and 4%, respectively). Further, the difficulty of the item increased from the field test, where the stimulus material was properly formatted. The analysis staff concluded that the formatting problem with the stimulus adversely affected respondents' ability to answer the question and decided to drop it from scaling. The item analysis revealed no problems with the remaining 152 items, which were retained for scaling.

### 14.2.1 Partial Credit Items

Although partial credit points were collected for some items in 1992, partial credit was not awarded when the items were scaled. Prior to scaling the 2003 survey data, NAAL staff reviewed the common 1992 and 2003 items to determine whether partial credit could be awarded to the items for which partial credit points were collected. In reviewing the items, analysts followed the same rules used to assign partial credit for the 2003 items:

- A score point assigned partial credit must substantively make sense as partial credit. In other words, to receive partial credit, respondents must show that they are able to do some meaningful part of the assigned task.
- On average, respondents who receive partial credit on an item should have average literacy scores (based on the other items they completed) between the scores of respondents who got the item completely right and respondents who got the item wrong.

The analysis staff examined the 1992 rubrics for the six blocks of items that were reused in 2003 to see which items had score points that made substantive sense as partial credit. Six out of some 65 items were identified with score points that were substantively appropriate for partial credit:

- N100401 (block 10, question 4), score point 2
- N100601 (block 10, question 6), score point 2
- N110601 (block 11, question 6), score point 2
- N110901 (block 11, question 9), score point 2
- N120401 (block 12, question 4), score point 2
- N130301 (block 13, question 3), score points 2 and 3

The analysis staff eliminated N100601 from consideration because respondents who received a partial credit score had literacy levels almost identical to respondents who got the item entirely wrong. This indicated that the item would not be likely to scale as partial credit. For the remaining five items, respondents who received the potential partial credit score point had total block scores that fell below the total block score for respondents who answered the question correctly and above the total block score for respondents who answered the question incorrectly. Once the potential partial credit items were identified, the five items were recoded as partial credit.

## 14.3 SCALING METHODOLOGY

Following the procedures used in 1992, the dataset used for scaling included respondents who completed five or more tasks on each of the prose, document, and quantitative literacy scales. Before scaling began, the analysis staff scored the data following the 1992 guidelines:[2]

- The correct key(s) for the item were considered **Right**.
- Nonresponses *that were followed by valid responses to other items in the same block* were considered **Omitted**. Items scored as **Omitted** were treated as though they were **Wrong**.
- Nonresponses *that occurred **after** the last item in a block with a valid response* were considered **Not Reached**. Items scored as **Not Reached** were treated as though they had never been presented to the respondent. This was done so as not to underestimate the literacy ability of respondents who did not complete an entire block.

[2]Exceptions to the scoring rules were made for partial credit items (discussed in section 14.1.2.3), which were considered partially correct rather than right or wrong.

- Multiple responses were considered **Wrong**.
- "I Don't Know" responses were considered **Wrong.**
- All other responses were considered **Wrong**.

This section reviews the scaling model employed in the analyses of the NAAL data and describes the marginal maximum likelihood (MML) methodology used for proficiency estimation.

### 14.3.1 The Scaling Model

Two distinct scaling models, depending on item type and scoring procedure, were used in the analysis. Each model is based on item response theory (IRT). Each is a "latent variable" model, defined separately for each of the scales, which expresses respondents' tendencies to achieve certain scores (such as correct/incorrect) on the items contributing to a scale as a function of a parameter that is not directly observed, called proficiency ($\theta$) on the scale.

The item response models used differ only in the form of the function $P_i(\theta)$. The two-parameter logistic (2PL) model (Hambleton, Swaminathan, and Rogers 1991), which was used for dichotomous items (that is, items that are scored either right or wrong), takes the form

$$P_i(\theta) = \frac{1}{1 + e^{-1.7a_i(\theta - b_i)}}$$

Where $P_i(\theta)$ is the probability that a randomly selected examinee with ability $\theta$ answers item $i$ correctly and $a_i$ and $b_i$ are parameters characterizing item $i$ in terms of its discrimination and difficulty. In the 1992 National Adult Literacy Survey, the three-parameter logistic (3PL) model was adopted to fit the dichotomous items. The 3PL model was not necessary for the NAAL 2003 calibration since no multiple choice items were included.

For the partial credit items, we used the Graded Response Logistic (GRL) model (Samejima, 1969, 1972). This model follows the 2PL model for the probability of a score of 1 (at least partially correct):

$$P_{i1}(\theta) = \frac{1}{1 + e^{-1.7a_i(\theta - b_{i1})}}$$

It also follows the 2PL model for the probability of a score of 2 (completely correct):

$$P_{i2}(\theta) = \frac{1}{1 + e^{-1.7a_i(\theta - b_{i2})}}$$

Linear transformation of the scales was used to link the NAAL scales to the 1992 NALS scales for gain purposes. The scale indeterminacy was resolved by setting an origin and unit size of theta to the reported scale means and standard deviations from 1992 NALS.

A basic assumption of item response theory is the conditional independence of the responses by an individual to a set of items, given the individual's $\theta$ score (a measure of proficiency). That is, item response probabilities depend only on the individual's $\theta$ and the specified item parameters, as opposed to depending on any demographic characteristics of examinees, the position of the item in the booklet, the content of items around an item of interest, or the assessment administration conditions. Conditional on $\theta$, the probability of a correct response on one item is unrelated to the probability of a correct response on another. This allows the following formula for the joint probability of examinee *i*'s response pattern $\mathbf{z}_i = (z_{i1}, z_{i2}, \ldots, z_{in})'$, where $z_{ij}$ is person *i*'s score on item *j*, across a set of *n* items for given ability $\theta$ and item parameters:

$$P(\mathbf{z}_i \mid \theta, \text{item parameters}) = \prod_{j=1}^{n} \prod_{k=0}^{m_j - 1} P_{jk}(\theta)^{v_{jk}}$$

where $m_j$ is the number of score categories of item *j*, and $v_{jk} = \begin{cases} 1 \text{ if } z_{ij} = k \\ 0 \text{ otherwise} \end{cases}$.

After the hypothetical response pattern *z* is observed, the above function can be viewed as a likelihood function that is to be maximized with a given set of item parameters. These item parameters were treated as known for the subsequent calculations.

Another assumption underlying the model is unidimensionality—that is, performance on a set of items is accounted for by a single construct. Although this assumption may be too strong, the use of the model is motivated by the need to summarize overall performance parsimoniously within a single domain. Hence, item parameters were estimated for each scale separately.

Testing the assumptions of the item response theory model is a critical part of the data analyses. A number of checks were made to detect multidimensionality and certain condition dependencies. Differential item functioning (DIF) analyses were used to examine issues of dimensionality (see section 4.2 in chapter 4), and item fit was examined to flag responses with serious departures from the IRT model. The fit of the IRT models to the observed data was examined within each scale by comparing the empirical item response functions (IRFs) with the theoretical curves. The theoretical curves were plots of the response functions based on the estimates of the item parameters. The empirical proportions were calculated from the posterior distributions of the theta for each student who received the item. For good fitting items, the empirical and theoretical curves were close together. Items for which this was not true indicated poor fit and were examined carefully. When warranted, remedial efforts, such as collapsing categories of polytomous item or removing items from the test, were made to mitigate the effects of such violations on inferences.

### 14.3.2 Linking to the 1992 NALS

As already noted, the prose, document, and quantitative literacy results from the NAAL are reported on scales that were established in the 1992 NALS; 65 (43 percent) of the tasks administered in the 2003 NAAL were originally administered in 1992. The linkage between the scales from the two studies is based on these tasks. In addition, 88 new tasks were developed for the NAAL and therefore a total of 153 tasks were administered in the 2003 assessment. However, out of the 88 new tasks, one task (C060401) was dropped prior to scaling on the basis of the examination of the item analysis statistics. A total of 152 tasks were retained for scaling.

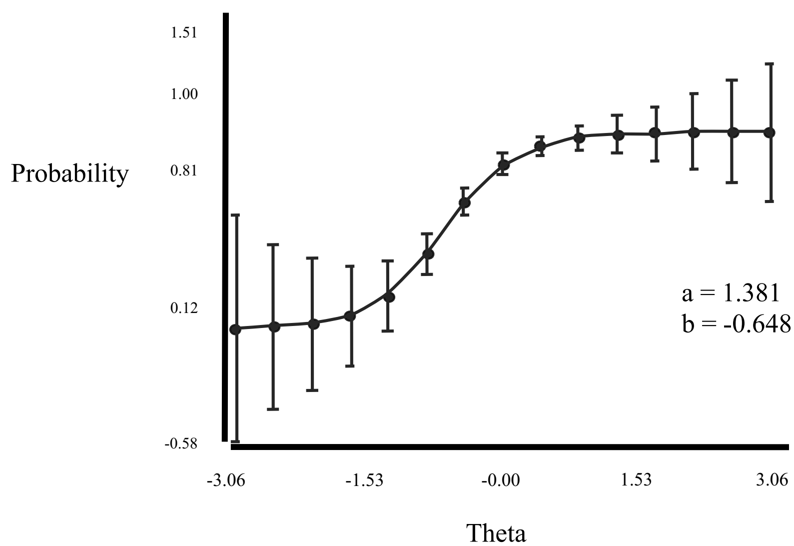
### 14.3.3 Item Parameter Estimation

Identical item calibration procedures were carried out separately for each of the three literacy scales. Analysts used the IRT package of the AM software developed by Cohen et al. (2000). The two-parameter logistic item response theory model was fit to dichotomous items and the Graded Response Logistic item response theory model was fit to partial credit items. Preliminary sample weights were used during the calibration procedures.

After operational data were received by NAAL staff, all items were put into the scaling, with the exception of the one item that was identified as flawed on the basis of the item analysis (C060401). During the scaling, any items identified as problematic on the basis of the item analysis were closely watched to ensure that the scaling model was correctly fitting the data. At this stage, certain items scored as partial credit were not well fit by the IRT Graded Response Logistic model. These items were collapsed back to right/wrong items and the scaling of the full item set was repeated.

In addition, linking the 1992 and 2003 assessments required that items presented in both years could be fit by using the same IRT item parameters, based on data from both years. The fit of the model to any item that was given in both assessments was evaluated by checking the fit of the IRT model estimated using both years of data to the data from each individual assessment. Any item showing lack of fit for either or both assessments was "split"; that is, it was treated as two distinct items: one for 1992 and another for 2003. Model fit was evaluated at the task level by inspecting residuals from fitted item response curves from AM. The item response curves were visually examined by comparing the empirical item response functions (IRFs) with the theoretical curves.  An example of item response curve is presented in figure 14-1.

**Figure 14-1.  Example of item response curve (item N100201) from the NAAL: 2003**




NOTE: Dots represent the 2003 NAAL data. They indicate estimated conditional probabilities obtained without assuming a specific model form. The curve indicates the estimated item response function (IRF) assuming a logistic model form. The bars around the dots indicate the standard errors around the dots.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

As a result of the waves of analysis, NAAL staff were able to successfully scale together the 1992 and the 2003 data. Only one item was dropped from scaling (C060401). Nine items in the 2003 dataset that were scored as partial credit were collapsed as a result of misfit when they were scaled with the graded response model. And only one item needed to be split (N130901 which was renamed CN130901 in the 2003 dataset where it was treated as a new item). Table 14-1 summarizes the treatment of the items (i.e., whether they were dropped, collapsed, or split). Estimated item parameters for each literacy scale are presented in tables D-1 through D-3 in appendix D. As shown in appendix D, the slope or discrimination parameters (parameter a) range from 0.41 to 2.55 for the prose literacy scale, from 0.41 to 2.63 for the document literacy scale, and from 0.34 to 4.60 for the quantitative literacy scale. The difficulty parameters (parameter b) for dichotomous items range from -2.50 to 1.71 for the prose literacy scale, from -6.34 to 1.52 for the document literacy scale, and from -2.82 to 1.77 for the quantitative literacy scale. The ranges of the step parameters for polytomous items are from -1.92 to 1.62 for the prose literacy scale, from -2.03 to 1.16 for the document literacy scale, and from -1.76 to 0.74 for the quantitative literacy scale.

**Table 14-1. NAAL items, by item treatment during scaling: 2003**

| Item treatment | Item |
|---|---|
| **Dropped** | C060401 |
| **Collapsed** | C030301, C030707, C040401, C040503, C040701, C040801, C061001, C061101, C080503 |
| **Split** | CN130901 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

## 14.4 DIRECT ESTIMATION OF SUBPOPULATION DISTRIBUTIONS OF PROFICIENCY USING DIRECT ESTIMATION WITH THE AM SOFTWARE

### 14.4.1 Background

As with most survey programs, NAAL faces competing demands. NAAL must contend with the paradoxical requirement that the test be both long enough to assess proficiency on a broad set of literacy skills and knowledge and short enough to ensure that the test can be completed in a reasonable length of time. The requirement for a short test comes from the very nature of the program: for examinees, NAAL is a low-stakes test, so they are unlikely to expend much effort. Individual examinees never receive any feedback about their individual performance; in fact, individual scores are never assigned at all. Therefore, the test must be short to avoid exceeding the effort that examinees are likely to expend.

NAAL fulfills both of these competing objectives by using an incomplete-block test design. Under this design, test items are arranged into large number of “blocks,” and only a small number of the blocks appear in each test booklet. Data from the different booklets are tied to a common scale through statistical methods based on item response theory (IRT; Rasch, 1960; Lord, 1952; Birnbaum, 1968; Hambleton and Swaminathan, 1995). Under this design, each examinee completes a single, short booklet containing a subset of blocks. Across all booklets, the assessment includes many items, enough to cover the full extent of the underlying trait being measured. Hence, NAAL appears short to the individual examinees—where it needs to be short—and long from the perspective of aggregate content coverage.

The use of IRT methods enables NAAL to calibrate items from different test booklets to reveal a common, partially observed latent trait, which we will call *proficiency*. Hence, although the measurement properties of the booklets vary both cross-sectionally and over time, the measurement models putatively yield access to a consistent underlying construct.

NAAL's incomplete-block design exacerbates a problem shared by all assessment programs that report aggregate proficiency statistics. Tests measure proficiency imperfectly, and the measurement error in individual scores can bias estimates of underlying proficiency, even when the individual-level instrument is very precise. For example, Mislevy, Beaton, Kaplan, and Sheehan (1992) estimate that the variability among individual scores from a 30-item test would have a reliability of 0.80 and would overestimate the population variance by 25 percent. Individual scores from an 85-item individual test with a reliability of 0.90 would still overestimate the variance by 11 percent. Similarly, the measurement error in these assessments would cause underestimates of correlation coefficients and bias other aggregate statistics such as the proportion of the population within specified score ranges. With the incomplete-block design, scores from the individual-level instruments tend to be quite imprecise, increasing the biases owing to measurement error. Further, the measurement properties typically differ across booklets, making the exact impact of the bias somewhat unpredictable.

In 1992, NALS used methods derived from Marginal Maximum Likelihood (MML) estimation to obtain unbiased estimates of the target statistics. Until recently, however, appropriate MML software was not widely available. Therefore, the contractor in charge of the 1992 NALS used the plausible values method developed for National Assessment of Educational Progress (NAEP; Mislevy 1984, 1985, 1991; Thomas, 1993) to allow secondary users to estimate statistics derived from individual data.

Plausible values are multiple imputations randomly drawn from a distribution derived from the MML parameter estimates for an extensive conditioning model (Allen, Carlson, and Zelenak, 1999). When analyzed as though they are test scores free of measurement error, plausible values yield good approximations of many aggregate statistics. However, it is important to understand that they remain approximations not estimated directly from the data.

Plausible values were introduced in the 1980s to provide secondary analysts as a way to analyze incomplete block-design assessment data using existing software and computers available at that time. Times have changed, however, in that most desktop computers provide more than enough power to directly estimate statistical models that plausible values approximate. The following is a brief overview of the direct estimation of models based on MML using the AM software, along with a Taylor-series approximation to the standard error for these models.

### 14.4.2 Direct Estimation of MML Models

This section describes MML regression, as well as the numerical optimization method that AM uses to estimate the model. The model itself is not new, and the optimization method is well known; however, this section describes some of the computational advantages of this approach over the now-standard EM optimization. In addition, this section describes the application of Binder's (1983) implicit differentiation method for variance estimation in complex samples to the MML regression. The development of the variance estimator addresses the fact that virtually all large-scale assessments involve complex sample designs, generally invalidating standard variance estimates.

Based on section 14.3.1, the conditional probability of examinee *i*'s response pattern $\mathbf{z}_i = (z_{i1}, z_{i2}, \ldots, z_{in})'$, where $z_{ij}$ is person *i*'s score on item *j*, across a set of *n* items for given ability $\theta$ and item parameters is:

$$P(\mathbf{z}_i \mid \theta, \text{item parameters}) = \prod_{j=1}^{n} \prod_{k=0}^{m_j - 1} P_{jk}(\theta)^{v_{jk}}$$

where $m_j$ is the number of score categories of item *j*, and $v_{jk} = \begin{cases} 1 \text{ if } z_{ij} = k \\ 0 \text{ otherwise} \end{cases}$. Suppose that the $\theta$ is a random draw from a population distribution with probability density function $f(\theta \mid \text{population parameters})$, then following Bayes rule, the marginal likelihood function of the population parameters for person *i* is:

$$L_i(\text{population parameters} \mid \;_i, \text{item parameters}) \propto$$
$$\int p(\;_i \mid \theta, \text{item parameters}) f(\theta \mid \text{population parameters}) d\theta$$

In practice, a normal distribution with mean $\mu$ and standard deviation $\sigma$ is often assumed for $f(\theta \mid \text{population parameters})$, hence, the marginal likelihood function becomes:

$$L_i(\mu, \sigma \mid \mathbf{z}_i, \text{item parameters}) \propto \int p(\mathbf{z}_i \mid \theta, \text{item parameters}) f(\theta \mid \mu, \sigma) d\theta \qquad (1)$$

Recall that in the regression case, $\mu_i = \mathbf{x}_i'\mathbf{b}$, where $\mathbf{x}_i$ is the design matrix formed by the independent variables of person *i*, and $\mathbf{b}$ is the regression parameters . In this case, the marginal likelihood function becomes:

$$L_i(\mathbf{b}, \sigma \mid \mathbf{z}_i, \mathbf{x}_i, \text{item parameters}) \propto \int p(\mathbf{z}_i \mid \theta, \text{item parameters}) f(\theta \mid \mathbf{x}_\text{i}; \mathbf{b}, \sigma) d\theta \qquad (1')$$

Note that likelihood function never actually requires a point estimate of proficiency (theta) for each individual. Rather, the method essentially "tries" all possible values, weighting each one by the probability that a random draw from a population with mean $\mu_i$ and variance $\sigma^2$ would yield it. In this way, the method estimates the distribution of proficiency in the population without ever estimating proficiency for each individual. This is why the method is called "marginal": it yields point estimates of the group or subgroup parameters without requiring point estimates for individual students by integrating the $\theta$ parameter out.

The estimation task for computing statistics based on $\theta$ involves specifying and maximizing the likelihood function across observations. For numerical reason, the log-likelihood function is often used in place of the likelihood function during the maximization procedure. We discuss the algorithm for doing so below, followed by a discussion of how we obtain estimates of the standard errors that are appropriate under a complex sample design.

#### 14.4.2.1 Estimation

We estimate this model by using an algorithm that combines quasi-Newton optimization with numeric quadrature over about 30 equidistant points along the feasible range of $\theta$. Although an EM algorithm is often used for models with the general form of equation 1, in this particular model it proves inefficient. The first term on the right-hand side of equation 1 is a product over many IRT link functions with fixed parameters. These are time-consuming computations. With fixed-distance quadrature points, this set of computations is required just once. Most EM algorithms would require recomputation at each iteration.

We use numeric quadrature to approximate the integral in equation 1, which is difficult to evaluate analytically. We identify a range within which all observations are virtually certain to fall and then select equally spaced points along this interval. For example, if theta is a standard normal (0, 1) variate, it makes sense to have the quadrature points range between about –4 and 4 (99.994% of cases will fall in this range). Thus, letting $q = \{1,2,...Q\}$ for $Q$ quadrature points ($\theta_q$), we can rewrite the individual likelihood function (1') as

$$L_i(\mathbf{b}, \sigma \mid \mathbf{z}_i, \mathbf{x}_i, \text{item parameters}) \approx \sum_{q=1}^{Q} p(\mathbf{z_i} \mid \theta_q, \text{item parameters}) f(\theta_q \mid \mathbf{x}_i; \mathbf{b}, \sigma). \qquad \textbf{(2)}$$

The goal then is to find the values of the parameters ($\mathbf{b}, \sigma$) that maximize the likelihood function. This is typically done by using iterative methods that try various values of the parameters, evaluate the likelihood function, then adjust the provisional parameter estimates to values that increase the function.

The method of Berndt, Hall, Hall, and Hausman (BHHH; 1974) was used to find values of $\mathbf{b}$ and $\sigma$ that maximize the likelihood functions. (The algorithm employed here offers the option of taking a single steepest-descent step when a BHHH step fails to yield improvement in the likelihood function. Alternatively, users can select the slower steepest-ascent algorithm.) This method modifies Newton's method and has proven quite successful in a range of MML problems. Letting $\mathbf{\Gamma} = (\mathbf{b}', \sigma\ '$ represent a vector of the parameters, Newton's method uses the following iterations to update $\mathbf{\Gamma}$: $\mathbf{\Gamma_{t+1}} = \mathbf{\Gamma_t} - \mathbf{G_t^{-1} g_t}$ for iteration $t+1$, based on values from iteration $t$. In this equation, $\mathbf{g}_t = \sum_i \mathbf{g}_{it}$ is the vector of first derivatives of the log-likelihood function with respect to the parameters from iteration $t$, and $\mathbf{G_t}$ is the matrix of second derivatives of the log-likelihood function with respect to the parameters from iteration $t$. Although Newton's method has many desirable properties in a range of maximum likelihood problems, it only works where $-\mathbf{G_t^{-1}}$ is positive definite. Unfortunately, this is not always the case, especially in early iterations. The BHHH method substitutes $\mathbf{H_t}$ for $-\mathbf{G_t^{-1}}$ where $\mathbf{H_t} = \left[\sum_i \mathbf{g}_{it} \mathbf{g}'_{it}\right]^{-1}$, thereby guaranteeing a positive definite matrix everywhere. At convergence, either matrix consistently estimates the inverse of the information matrix.

Using this method, it is necessary to specify the likelihood function and the first-order derivatives with respect to the vector $\mathbf{b}$ and the scalar $\sigma$ (which are appended into a vector, **g**). These derivatives take a relatively simple form.

The first $k$ elements of **g** represent $\dfrac{\partial \ln(L)}{\partial b_k}$. This portion of **g** is given by

$$\frac{\partial \ln(L)}{\partial \mathbf{b}} = \sum_i \mathbf{x}_i \left( \frac{1}{\sigma^3 L_i} \sum_q - \exp^{-\frac{(\mathbf{x}_i'\mathbf{b}-\theta_q)^2}{2\sigma^2}} \left(\mathbf{x}_i'\mathbf{b} - \theta_q\right) p_{iq} \right)$$

where $L_i$ is the value of the likelihood function for examinee $i$ and $p_{iq} = p(\mathbf{z}_i \mid \theta_q, \text{item parameters})$. The final element of $\boldsymbol{g}$, the first derivative with respect to sigma, is given by

$$\frac{\partial \ln(L)}{\partial \sigma} = \sum_i \left( \frac{1}{\sigma^4 L_i} \sum_q p_{iq} \left( \exp^{-\frac{(\mathbf{x}_i'\mathbf{b}-\theta_q)^2}{2\sigma^2}} \left(\mathbf{x}_i'\mathbf{b} - \theta_q\right)^2 - \exp^{-\frac{(\mathbf{x}_i'\mathbf{b}-\theta_q)^2}{2\sigma^2}} \sigma^2 \right) \right).$$

This approach works well even when starting with far less than optimal starting values. To improve performance, we begin with a rough approximation: we calculate an approximate score for each respondent as the weighted average of the values of the quadrature points, where the weights are given by $p_{iq}$. We then estimate an OLS regression against these “pseudo-scores” and take these as starting values for the **b** parameters. The regression root MSE, appropriately adjusted for the unreliability of the pseudo-scores, provides a starting value for $\sigma$.

### 14.4.3 Weighting and Variance Estimation in Complex Samples

The roots of the score functions constitute sets of *estimating functions* (Godambe, 1960, 1991). Godambe (1960) proves an optimal property of maximum likelihood by estimating functions in simple random samples—and Godambe and Thompson (1986) show that the standard π-weighted estimating function retains this optimal property in unequal probability samples. Thus, the log-likelihood function and its derivatives at each observation can be multiplied by the inverse probability of inclusion in the sample, yielding optimal estimating functions.

The score function provides an estimating equation by which consistent estimates of the finite population parameters may be obtained, even though their superpopulation counterparts may be fictional:

$$\hat{W}(\hat{\boldsymbol{\Gamma}}) = \mathbf{0} = \sum_{i=1}^{n} w_i \frac{\partial \log(L_i(\hat{\boldsymbol{\Gamma}} \mid \mathbf{z}_i, \mathbf{x}_i, \text{item parameters}))}{\partial \hat{\boldsymbol{\Gamma}}}$$

where $w_i$ represents the sample weight, usually the inverse of an estimate of the probability of selection.

The lack of independence among the observations and the misspecification of the model render the traditional maximum likelihood variance estimator based on the inverse of the observed information matrix useless. More appropriate is an approximate variance estimator based on Binder's (1983) method of implicit differentiation, which Godambe and Thompson (1986) suggest be applied to estimating functions in the presence of unequal weights. Binder begins with a Taylor series expansion of the estimation function around the true-value . A first-order expansion of a linear estimating equation $\hat{W}(\hat{\mathbf{\Gamma}}) = 0$ yields

$$0 \approx \hat{W}(\mathbf{\Gamma}) + \frac{\partial \hat{W}(\mathbf{\Gamma})}{\partial \mathbf{\Gamma}}(\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}).$$

Solving for $\hat{\mathbf{\Gamma}} - \mathbf{\Gamma}$ yields

$$\hat{\mathbf{\Gamma}} - \mathbf{\Gamma} = -\left[\frac{\partial \hat{W}(\mathbf{\Gamma})}{\partial \mathbf{\Gamma}}\right]^{-1} \hat{W}(\mathbf{\Gamma}),$$

and taking variance of both sides,

$$Var(\hat{\mathbf{\Gamma}}) = \left[\frac{\partial W(\mathbf{\Gamma})}{\partial \mathbf{\Gamma}}\right]^{-1} \Omega(\mathbf{\Gamma}) \left[\frac{\partial W(\mathbf{\Gamma})'}{\partial \mathbf{\Gamma}}\right]^{-1}$$

where $\Omega(\mathbf{\Gamma})$ is the variance of $W(\mathbf{\Gamma})$ across observations. Substituting expectations (estimates) in place of true values yields the proposed standard error estimator:

$$Var(\hat{\mathbf{\Gamma}}) = \hat{G}^{-1}(\hat{\mathbf{\Gamma}}) Var(\hat{g}(\hat{\mathbf{\Gamma}})) \hat{G}^{-1}(\hat{\mathbf{\Gamma}})$$

This is popularly known as a *sandwich estimator*. In this case, the outer terms are approximated by $\mathbf{H} = \left[\sum_i \mathbf{g}_i \mathbf{g}_i'\right]^{-1}$, where $\mathbf{g}_i$ is the vector of first derivatives of the log-likelihood function with respect to the converged parameter values, i.e. $\mathbf{g}_i = w_i \dfrac{\partial \log(L_i(\hat{\mathbf{\Gamma}} \mid \mathbf{z}_i, \mathbf{x}_i, \text{item parameters}))}{\partial \hat{\mathbf{\Gamma}}}$ . The variance term in the center is the estimated variance/covariance matrix of the first derivatives. Note that this is simply the variance/ covariance matrix of a set of population totals (the summed first derivatives). In a stratified, clustered, unequally weighted sample, one can usually approximate this as the π estimator of the stratified, between primary sampling unit (PSU) variance.

Using the stratified, between- primary sampling unit (PSU) weighted estimator to estimate $\hat{\Omega}(\hat{\mathbf{\Gamma}})$, gives

$$\hat{\Omega}(\hat{\mathbf{\Gamma}}) = \sum_{h=1}^{H} \left(\frac{n_h}{n_h - 1}\right) \sum_{i=1}^{n_h} (\mathbf{g}_{hi} - \overline{\mathbf{g}}_h)(\mathbf{g}_{hi} - \overline{\mathbf{g}}_h)'$$

where $\mathbf{g}_{hi} = \sum_{k=1}^{m_{hi}} w_{hik} \frac{\partial \log(L_i(\mathbf{\Gamma} \mid \mathbf{z}_i, \mathbf{x}_i, \text{item parameters})}{\partial \mathbf{\Gamma}}$, and $\overline{\mathbf{g}}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} \mathbf{g}_{hi}$, in which $h$ indexes strata, $i$ indexes primary sampling units, and $k$ indexes individuals. In a simple random sample the PSUs are the sampled examinees.

## 14.5 LINKING THE 2003 SCALE TO THE 1992 SCALE

Initially, proficiency scores are estimated on the basis of a provisional scale with a mean of 0 and a standard deviation of 1. To be comparable to the scores from 1992, they need to be put on the same scale. This is accomplished through the use of linear transformation constants that match the mean and standard deviations of the 1992 sample based on the new item parameters.[3] The transformation that was applied is as follows:

$$\theta = \mathrm{A}\theta^* + \mathrm{B}$$

Where $\theta^*$ is the provisional scale from item calibration and $\theta$ is the reported 0 to 500 scale, and A and B are transformation constants. Table 14-2 presents the transformation constants (that is, the standard deviations and means) for the distributions of the three scales. These constants apply both to the 2003 data and to the 1992 data when the new item parameters are used.

**Table 14-2. Transformation constants (standard deviations and means) using new item parameters, by literacy scale: 1992 and 2003**

| Literacy scale | A (SD) | B (Mean) |
|---|---|---|
| Prose | 58.48056 | 280.705 |
| Document | 58.75546 | 274.8816 |
| Quantitative | 63.31159 | 280.4884 |
| Composite | 57.3496 | 280.6508 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 1992 National Adult Literacy Survey and 2003 National Assessment of Adult Literacy

[3]This is also based on the redefined sample of complete cases and rescored items from 1992 data.

## 14.6 MINIMUM SAMPLE SIZES FOR REPORTING SUBGROUP RESULTS

In the NAAL reports, the sample sizes were not always large enough to permit accurate estimates of proficiency and/or background results for one or more categories of variables. For results to be reported for any subgroup, a minimum sample size of 45 was required. This number was arrived at by determining the sample size needed to detect an effect size of 0.5 with a probability of 0.8 or greater, using a design effect of 1.5. This design effect implies a sample design-based variance 1.5 times that of a simple random sample. The effect size of 0.5 pertains to the true difference in mean proficiency between the subgroup in question and the total population, divided by the standard deviation of proficiency in the total population. An effect size of 0.5 was chosen following Cohen (1988), who classifies effect size of this magnitude as “medium” as well as to be consistent with what was done in the 1992 survey.

# CHAPTER 15

# THE LITERACY OF ADULTS WITHOUT COGNITIVE DATA

Ying Jin and Stéphane Baldi, American Institutes for Research

## 15.1 INTRODUCTION

Missing data are always expected in any large-scale assessment. Sampled individuals may not respond to an assessment for many reasons. For example, they may not respond because they have a language barrier or a physical disability, or they may simply refuse. A number of alternative ways are available to deal with missing data. The least desirable way is simply to ignore the missing data. This practice assumes that the data are missing at random and that the remaining observed samples are representative of the target population. However, if the pattern of missing data is correlated to the outcome of the study, this practice would yield both biased and inaccurate estimates of proficiency distributions for some subpopulations and consequently for the total population as well. For further information on nonresponse bias analysis and different approaches to treating missing data, the reader can refer to section 4-4 of NCES Statistical Standards (NCES 2003-601) and Groves et al. (2001).

Experience with the 1992 NALS and other assessments, as well as evidence from the 2003 NAAL assessment, indicates that adults with lower levels of literacy are more likely than adults with higher proficiencies to refuse to respond to the assessment. Ignoring the pattern of missing data would have resulted in overestimating the literacy skills of adults in the United States. This chapter describes the methods adopted to deal with the problem of missing cognitive data.

### 15.1.1 Missing Background Data

The target sample for the 2003 NAAL assessment included 35,365 nationally representative housing units, of which 4,671 were vacant. Approximately 18 percent of the households that were occupied at the time of data collection refused to participate, and no detailed background information is available on this group. For the households that agreed to participate in the study, the interviewers began by using a series of screening questions to obtain an accurate count of the number of age-eligible persons in the household. Depending on the number of eligible persons in the household, one or two persons were selected to participate in the interview to complete the background questionnaire (BQ). Respondents who did not answer a sufficient number of background questions were considered incomplete cases. Cases that were mostly incomplete could not be analyzed and were never incorporated into the database. Such cases were dealt with through weighting class and poststratification adjustments for instrument nonresponse (see chapter 12). For some background variables, such as education and country of birth, missing data for literacy-related reasons were imputed using logistic regression models (see chapter 12).

### 15.1.2 Missing Cognitive Data

The 19,714 adults ages 16 and older residing in households or prisons who agreed to respond to the assessment answered extensive background questions during the interview, as appropriate to the household or prison samples, about their age, country of birth, language(s) spoken or read, highest level of education completed, current educational aspirations, labor market status, current occupation and wages, voting behaviors, and reading habits. After answering the background questions, respondents were asked to complete the literacy tasks in the assessment booklet. Very easy tasks were placed first to encourage respondents to continue. (See chapter 2 for details on booklet and block design of the cognitive assessment.) Nevertheless, 1,155 (6 weighted percent) of these respondents did not complete any cognitive tasks on any of the three scales, and 256 (1 weighted percent) did not complete any cognitive tasks on at least one but not all scales. For individuals who refused to continue after answering the background questions, no information is available about their performance on the cognitive tasks. Completely omitting these individuals from the analyses would have resulted in overestimates of the literacy skills of the national population as a whole and particularly of certain subpopulations. Imputation procedures were applied to enable the estimation of their literacy proficiencies.

## 15.2 THE NORMAL TREATMENT OF MISSING COGNITIVE DATA

For the National Assessment of Adult Literacy, a distinction should be made between missing responses by design and missing responses by individual respondents who were presented with the cognitive questions. In population assessments, unlike in individual assessments, a matrix design for item sampling in which examinees respond to different subsets of cognitive questions is usually used to limit the burden on respondents. Because of the matrix sampling design of the NAAL assessment, each respondent received only a fraction (three-thirteenths) of the literacy tasks; therefore, most of the tasks were not presented to every respondent and could be considered missing. This type of missing data was intentional by design. Missing responses could also occur for the cognitive tasks that were presented to respondents.

### 15.2.1 Omitted, Not Reached

For the NAAL literacy tasks that were presented, the missing responses occurred in two distinct patterns: the respondent skipped over a question and responded to a subsequent question, or the respondent broke off the assessment and did not attempt to respond to any subsequent questions. The two types of nonresponse were called "omitted" and "not reached" tasks, categories that were based on the distinction between missing responses prior to the last question the respondent answered in each block and missing responses subsequent to the last observed response in each block.

- **Omitted.** In some cases, respondents skipped over a particular task but attempted or completed one or more tasks that followed. This kind of missing responses, called "omitted," occurred prior to the last observed response in each block and by definition could not be at the end of a block.

- **Not reached**. In other cases, respondents spent all their time responding to preceding tasks and did not reach tasks that appeared later in a block. This kind of missing responses was observed subsequent to the last question the respondent attempted in each block. Tasks that were not attempted were found consecutively at the end of the blocks and called "not reached**."**

In the omitted-response situation, there was a logical basis for assigning a wrong answer to a missing response. As in the 1992 NALS survey, in the 2003 NAAL assessment, omitted cognitive responses were treated as wrong answers, on the assumption that respondents decided to skip them intentionally because they found these tasks too difficult. The treatment of an omitted response as a wrong answer was a logical imputation that was based on the circumstances that surrounded the missing data.

In contrast, not-reached cognitive responses were not treated as wrong but were treated as if the questions had not been presented to the respondent. The assumption here was that respondents did not make a task-specific decision about whether to respond to tasks that were not reached. Because there was not a sufficient logical basis for assigning a wrong answer, these responses remained missing data. Moreover, it was unlikely that not-reached tasks occurred because of speededness—that is, because the respondent was not given enough time to answer them. The assessment booklet was not rushed and the NAAL was not a speeded test, but there were some practical time limits so that the interviewer would not have to spend an unreasonable number of hours collecting information.

### 15.2.2 Cognitive Data Missing by Design

The 2003 NAAL assessment had 153 individual literacy tasks. It was impractical to ask respondents to take the entire pool of the assessment questions, particularly for a low-stakes assessment such as the NAAL. Therefore, to allow maximum coverage of literacy materials and content while minimizing the time burden for any one respondent, the NAAL assessment used a matrix item sampling design (see chapter 7) in which individuals responded to different subsets of cognitive tasks. Because of the matrix design of the NAAL assessment, each respondent received only three-thirteenths of the literacy tasks. Therefore, for every respondent, most of the tasks were not presented. The tasks that were not presented could be considered missing, but unlike omitted or not-reached tasks, this type of missing data was intentional by design.

## 15.3 REASONS FOR MISSING DATA

As will be described in the next section, missing data were imputed on the basis of the reasons for nonresponse, in particular, whether or not the reason was literacy related. This section summarizes the reasons for missing data in the NAAL sample. Section 15.3.1 gives the distribution of the sample by presence of cognitive data and reasons for missing data. Section 15.3.2 provides support for the validity of reasons through a comparison of reasons for various demographic subgroups.

### 15.3.1 Disposition Codes and Literacy Skills

Table 15-1 shows the distribution of the final weighted NAAL sample into categories on the basis of the following criteria: presence of cognitive data, type of nonresponse, and response to the Adult Literacy Supplemental Assessment (ALSA). Literacy-related reasons for nonresponse consisted of language problems, mental disabilities (including mental retardation, learning disabilities, and mental/emotional conditions), and reading/writing barriers. Refusals, physical disabilities (including visual and hearing impairments), and other/unknown reasons were classified as not literacy related because the disposition code provided no direct evidence of low literacy skills.

The distribution in table 15-1 is provided for the total sample and separately for the household sample and the prison sample. Of the 19,714 cases in the total sample, 93 percent had complete cognitive data. Complete cognitive data were obtained for 93 percent of the household sample (total sample size 18,541) and for 96 percent of the prison sample (total sample size 1,173). For the remaining cases, literacy scores were imputed if the sample person had partial cognitive data, the reason for incomplete cognitive data was not literacy related, or the sample person had completed the ALSA. There were 456 cases (3 percent) that did not meet these requirements and for whom scores were not imputed.

**Table 15-1. Percent distribution of final weighted NAAL sample, by presence of cognitive data, reason for missing data, and sample type: 2003**

| | | | Incomplete cognitive data | | No cognitive data | | |
|---|---|---|---|---|---|---|---|
| Sample | Total | Complete cognitive data | Literacy related | Not literacy related[1] | Literacy related (no imputation) | Literacy related, ALSA[2] respondent | Not literacy related[1] |
| Total | 100.0 | 92.6 | 0.2 | 0.7 | 2.9 | 0.1 | 3.6 |
| Household | 100.0 | 92.5 | 0.2 | 0.7 | 2.9 | 0.1 | 3.6 |
| Prison | 100.0 | 96.0 | # | 0.9 | 1.4 | 0.2 | 1.5 |

# Rounds to zero.
[1] Disposition code provides no direct evidence of low literacy skills.
[2] Adult Literacy Supplemental Assessment.
NOTE: Incomplete cognitive data have at least one but not all scales. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

The final disposition code assigned to the case was used to determine whether cognitive data were missing for literacy-related reasons. The same procedure was used for the household and prison samples. Table 15-2 shows the weighted distribution of the 19,714 NAAL sample cases by presence of cognitive data and the reason for incomplete data. Among cases with no cognitive data, the predominant reason for nonresponse was refusal (2.2 percent), followed by language problems (1.9 percent).

**Table 15-2. Distribution of final weighted NAAL sample of 19,714, by presence of cognitive data and detailed reason for missing data: 2003**

| Presence of cognitive data and report reason (if no cognitive data) | Percent of adults | Presumed relation of reason to literacy skills |
|---|---|---|
| Total | 100.0 | † |
| Cognitive data | 93.4 | † |
| No cognitive data | 6.6 | † |
| Refused | 2.2 | Not literacy related[1] |
| Non-English language | 1.9 | Literacy related |
| Mental disability, including retardation, learning disability, and other mental/emotional condition | 0.9 | Literacy related |
| Other or unknown | 0.7 | Not literacy related[1] |
| Physical disability, including visual | 0.7 | Not literacy related[1] |
| Reading and/or writing difficulty | 0.2 | Literacy related |

† Not applicable.
[1] Disposition code provides no direct evidence of low literacy skills.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

### 15.3.2 Household Sample Internal Evidence for the Validity of Reasons

The reasons provided for incomplete assessments were used to determine the treatment of cases with missing cognitive data. This section presents an evaluation of the validity of those reasons for the household sample.[1] An analysis was performed for cases with no cognitive data, using demographic information from the screener and the background questionnaire. The evaluation follows the analysis presented in chapter 8 of the Technical Report and Data File User's Manual for the 1992 National Adult Literacy Survey (Yamamoto 2001). Table 15-3 displays the distribution of the interviewer-coded reasons for not providing cognitive data, by age, race/ethnicity, language spoken when growing up, and educational attainment. Standard errors for table 15-3 were computed in WesVar, using replicate weights, and are provided in table 15-4.

[1] The analysis was restricted to the household sample because of small sample sizes and unstable estimates for the prison sample.

**Table 15-3. Percentage of household study population by presence of cognitive data, reason for missing data, and population group: 2003**

| Population group | Total sample size | Percent with cognitive data present | Percent with no cognitive data present: Reasons related to literacy: Total | Reasons related to literacy: Non-English language | Reasons related to literacy: Mental disability | Reasons related to literacy: Reading or writing difficulty | Reasons unrelated to literacy: Total | Reasons unrelated to literacy: Refused | Reasons unrelated to literacy: Physical disability | Reasons unrelated to literacy: Other, unknown, no answer |
|---|---|---|---|---|---|---|---|---|---|---|
| Total | 18,500 | 93.4 | 3.0 | 1.9 | 0.9 | 0.2 | 3.6 | 2.2 | 0.7 | 0.7 |
| Age (years) | | | | | | | | | | |
| 16–29 | 4,700 | 96.0 | 1.8 | 1.0 | 0.7 | 0.1 | 2.2 | 1.8 | # | 0.4 |
| 30–49 | 7,300 | 94.3 | 2.7 | 2.1 | 0.5 | 0.1 | 3.1 | 2.2 | 0.1 | 0.7 |
| 50–69 | 4,600 | 93.3 | 2.7 | 1.8 | 0.7 | 0.2 | 4.0 | 2.6 | 0.8 | 0.6 |
| 70+ | 2,000 | 85.2 | 7.5 | 3.3 | 3.4 | 0.9 | 7.3 | 2.4 | 3.5 | 1.4 |
| Race/ethnicity | | | | | | | | | | |
| Hispanic | 3,200 | 91.0 | 3.8 | 2.3 | 0.8 | 0.7 | 5.2 | 2.7 | 0.8 | 1.8 |
| Black | 3,500 | 94.3 | 1.7 | 0.7 | 0.6 | 0.5 | 4.0 | 1.4 | 0.8 | 1.8 |
| Asian | 400 | 87.2 | 7.1 | 7.1 | # | # | 5.7 | 4.5 | # | 1.2 |
| White/other[1] | 11,400 | 94.0 | 2.9 | 1.8 | 1.0 | 0.1 | 3.1 | 2.2 | 0.6 | 0.3 |
| Language spoken while growing up | | | | | | | | | | |
| English | 15,700 | 96.5 | 0.3 | 0.1 | 0.1 | 0.1 | 3.2 | 2.1 | 0.7 | 0.5 |
| Other language | 2,500 | 89.7 | 3.1 | 2.4 | 0.1 | 0.7 | 7.2 | 4.1 | 0.6 | 2.6 |
| Missing | 400 | 1.1 | 98.8 | 65.1 | 33.7 | # | 0.1 | # | # | 0.1 |
| English spoken while growing up | | | | | | | | | | |
| Hispanic | 1,200 | 94.7 | 0.6 | 0.1 | 0.2 | 0.2 | 4.8 | 3.5 | 1.0 | 0.3 |
| Black | 3,400 | 95.5 | 0.5 | # | # | 0.5 | 4.0 | 1.4 | 0.8 | 1.8 |
| Asian | 200 | 94.3 | 2.7 | 2.7 | # | # | 3.0 | 3.0 | # | # |
| White/other[1] | 10,900 | 96.8 | 0.2 | # | 0.1 | 0.1 | 3.0 | 2.1 | 0.7 | 0.3 |

See notes at end of table.

**Table 15-3. Percentage of household study population by presence of cognitive data, reason for missing data, and population group: 2003—Continued**

| Population group | Total sample size | Percent with cognitive data present | Percent with no cognitive data present: Reasons related to literacy: Total | Non-English language | Mental disability | Reading or writing difficulty | Reasons unrelated to literacy: Total | Refused | Physical disability | Other, unknown, no answer |
|---|---|---|---|---|---|---|---|---|---|---|
| English not spoken while growing up | | | | | | | | | | |
| Hispanic | 1,900 | 92.7 | 1.5 | 0.5 | 0.1 | 1.0 | 5.8 | 2.2 | 0.7 | 2.9 |
| Black | 100 | 94.0 | # | # | # | # | 6.0 | 2.0 | 2.0 | 2.0 |
| Asian | 200 | 79.6 | 11.9 | 11.9 | # | # | 8.6 | 6.1 | # | 2.4 |
| White/other[1] | 300 | 86.5 | 1.9 | 1.6 | # | 0.3 | 11.7 | 9.3 | 0.5 | 1.9 |
| Education[2] | | | | | | | | | | |
| Less than high school | 4,600 | 87.7 | 7.3 | 4.7 | 1.7 | 0.9 | 5.0 | 1.7 | 1.3 | 2.0 |
| High school or GED | 4,700 | 93.3 | 2.5 | 1.7 | 0.7 | 0.1 | 4.2 | 2.8 | 0.8 | 0.7 |
| More than high school | 9,200 | 95.8 | 1.6 | 0.9 | 0.7 | # | 2.6 | 2.1 | 0.3 | 0.2 |

#Rounds to zero.
[1]Other includes non-Hispanic American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races.
[2]Includes values of educational attainment imputed during the weighting process for weighting purposes.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 15-4. Standard errors for table 15-3**

| Population group | Standard error of percent with cognitive data present | Standard error of percent with no cognitive data present | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Reasons related to literacy | | | | Reasons unrelated to literacy | | | |
| | | Total | Non-English language | Mental disability | Reading or writing difficulty | Total | Refused | Physical disability | Other unknown, no answer |
| Total | 0.44 | 0.26 | 0.26 | 0.09 | 0.04 | 0.37 | 0.35 | 0.08 | 0.08 |
| Age (years) | | | | | | | | | |
| 16–29 | 0.57 | 0.27 | 0.21 | 0.16 | 0.05 | 0.49 | 0.48 | 0.03 | 0.09 |
| 30–49 | 0.45 | 0.33 | 0.31 | 0.09 | 0.04 | 0.41 | 0.39 | 0.06 | 0.11 |
| 50–69 | 0.71 | 0.41 | 0.39 | 0.19 | 0.07 | 0.57 | 0.49 | 0.17 | 0.13 |
| 70+ | 1.33 | 0.87 | 0.70 | 0.49 | 0.25 | 0.78 | 0.49 | 0.45 | 0.31 |
| Race/ethnicity | | | | | | | | | |
| Hispanic | 1.05 | 0.83 | 0.79 | 0.25 | 0.20 | 0.77 | 0.70 | 0.26 | 0.25 |
| Black | 0.57 | 0.36 | 0.29 | 0.12 | 0.20 | 0.45 | 0.22 | 0.22 | 0.36 |
| Asian | 2.75 | 2.03 | 2.03 | # | # | 1.43 | 1.29 | # | 0.63 |
| White/other[1] | 0.49 | 0.28 | 0.27 | 0.10 | 0.03 | 0.40 | 0.36 | 0.10 | 0.06 |
| Language spoken while growing up | | | | | | | | | |
| English | 0.35 | 0.06 | 0.05 | 0.04 | 0.04 | 0.33 | 0.31 | 0.09 | 0.07 |
| Languages other than English | 1.44 | 0.71 | 0.64 | 0.04 | 0.20 | 1.08 | 1.03 | 0.21 | 0.38 |
| Missing | 0.65 | 0.65 | 3.66 | 3.54 | # | 0.07 | # | # | 0.07 |
| English spoken while growing up | | | | | | | | | |
| Hispanic | 1.34 | 0.26 | 0.06 | 0.20 | 0.17 | 1.26 | 1.17 | 0.43 | 0.13 |
| Black | 0.49 | 0.20 | 0.04 | 0.02 | 0.21 | 0.44 | 0.21 | 0.22 | 0.36 |
| Asian | 2.21 | 1.29 | 1.29 | # | # | 1.38 | 1.38 | # | # |
| White/other[1] | 0.38 | 0.06 | 0.03 | 0.05 | 0.03 | 0.36 | 0.32 | 0.10 | 0.06 |

See notes at end of table.

**Table 15-4. Standard errors for table 15-3—Continued**

| Population group | Standard error of percent with cognitive data present | Standard error of percent with no cognitive data present | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Reasons related to literacy | | | | Reasons unrelated to literacy | | | |
| | | Total | Non-English language | Mental disability | Reading or writing difficulty | Total | Refused | Physical disability | Other unknown, no answer |
| English not spoken while growing up | | | | | | | | | |
| Hispanic | 0.88 | 0.35 | 0.16 | 0.06 | 0.31 | 0.73 | 0.62 | 0.30 | 0.39 |
| Black | 3.03 | # | # | # | # | 3.03 | 2.06 | 1.67 | 1.55 |
| Asian | 4.64 | 3.70 | 3.70 | # | # | 2.27 | 2.00 | # | 1.18 |
| White/other[1] | 2.89 | 0.89 | 0.86 | # | 0.24 | 2.90 | 2.86 | 0.48 | 0.94 |
| Education[2] | | | | | | | | | |
| Less than high school | 0.78 | 0.72 | 0.66 | 0.29 | 0.19 | 0.51 | 0.37 | 0.24 | 0.26 |
| High school or GED | 0.64 | 0.38 | 0.34 | 0.13 | 0.04 | 0.55 | 0.47 | 0.15 | 0.13 |
| More than high school | 0.45 | 0.18 | 0.16 | 0.11 | 0.01 | 0.39 | 0.39 | 0.07 | 0.04 |

#Rounds to zero.
[1]Other includes non-Hispanic American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races.
[2]Includes values of educational attainment imputed during the weighting process for weighting purposes.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Age and Reasons.** For adults ages 70 and over, physical disabilities were the predominant reason for not providing any cognitive data. The prevalence of missing cognitive data related to physical disabilities increased with age, from less than 0.1 percent for sample persons ages 16 to 29 to 3.5 percent for those ages 70 and older. Mental disabilities, such as dementia, may also be associated with aging. As is evident in table 15-3, adults ages 70 and older with no cognitive data were more likely than younger adults to give mental disabilities and reading or writing difficulties as a reason for nonresponse. Overall, adults ages 70 and over were about as likely to give a literacy-related reason as a non-literacy-related reason, whereas the other age groups were more likely to provide reasons unrelated to literacy. The same pattern was evident in the 1992 NALS. Although the trend seen in the 1992 NALS of increasing refusals with increasing age was evident in the current study for the lower three age groups, it did not hold true for the group ages 70 and over.

**Race/Ethnicity and Reasons.** Four racial/ethnic groups were evaluated: Hispanic, Black, Asian, and White/other. The "other" group included non-Hispanic American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and multiple races. Non-Hispanic White and "other" were combined because of the small sample size for the "other" category. Speaking a non-English language was provided as a reason for having no cognitive data for 7.1 percent of the Asian population, compared with 2.3 percent of the Hispanic population, 1.8 percent of White/other populations, and 0.7 percent of the Black population. Asian adults were also the only racial/ethnic group more likely to be assigned a literacy-related than a non-literacy-related disposition code when cognitive data were missing. Differences among racial/ethnic groups may be attributable to differences in languages spoken while growing up.

**Native Language and Reasons.** For sample persons with no cognitive data, language problems were given as a reason for nonresponse for approximately one-third of adults who did not speak English while growing up, consistent with the 1992 NALS analysis. In contrast, less than 0.1 percent of adults who spoke English while growing up had language problems as a reason for not providing cognitive data.

**Native Language, Race/Ethnicity, and Reasons.** When the population was restricted to adults who spoke English while growing up, only minor differences were seen in the reasons for missing cognitive data among the Hispanic, Black, and White/other racial/ethnic groups. The percentage of language problems remained higher among the Asian population than among the other racial/ethnic groups, but the sample size for this group was small. For adults who did not speak English while growing up, the differences were greater, with 0.5 percent of the Hispanic population, 11.9 percent of the Asian population, and 1.6 percent of the White/other population having no cognitive data because of problems with the English language. Among the small Black population who did not speak English while growing up, no sample persons lacked cognitive data as a result of language problems.

**Education and Reasons.** The proportion of sample persons who were missing cognitive data because of literacy-related reasons was greater among those with less than a high school education than among those with higher educational attainment. Reading/writing difficulties contributed to the lack of cognitive data for less than 0.1 percent of the population with a high school degree or higher.

Disposition codes were related to the age, race, language, and education of the sampled persons. In addition, the relationships were consistent with those of the 1992 NALS evaluation as noted above. The notable exception was the greater percentage of Hispanic adults with cognitive data in the 2003 NAAL, which was a result of the implementation of the ALSA. Overall, the analysis lends support to the validity of reasons provided for missing cognitive data and to the potential bias that could be incurred if missing data were ignored.

## 15.4 POPULATION ESTIMATES FOR ADULTS WITH UNMEASURABLE LITERACY SKILLS

In the 2003 NAAL assessment, among the 19,714 persons who agreed to respond to the assessment, 1,155 (6 percent) of these respondents did not complete any cognitive tasks on any scale, and 256 (1 percent) did not complete any cognitive tasks on at least one scale. For such nonrespondents, no information was available about their performance on the literacy scale that they did not complete.

When missing data patterns are related to the outcome of a study, the missing responses, if ignored, will bias the results unless some adjustment can be made to counter the bias. Evidence from the 2003 NAAL assessment indicated that response rate were different among subpopulations, and adults with lower levels of literacy were more likely than adults with higher proficiencies either to decline to respond to the assessment at all or to begin the assessment but not complete it. Excluding these individuals from the analyses would have resulted in overestimates of the literacy proficiencies of the national population as a whole and particularly of certain subpopulations.

### 15.4.1 Using “Reasons” to Improve Treatment of Missing Cognitive Data

As in the 1992 NALS, nonrespondents in the 2003 NAAL were classified into two separate groups: people who did not answer any cognitive questions for literacy-related reasons and people who did not respond for non-literacy-related reasons, based on the adults’ self-reported reasons for nonresponse that Westat collected in the field. Literacy-related reasons included reading/writing barrier, language problem, and mental disability. Non-literacy-related reasons included physical disability (respondents were offered assistance with writing if necessary), refusals, unavailability for the field period, and other. Responses to the background variables indicated that those who did not respond to the cognitive items for literacy-related reasons were disproportionately likely to be foreign born, to have less than a high school education, to be Hispanic or Asian/Pacific Islander, and to be age 65 or older. These variables are known to relate to English language proficiency and cognitive skills. Combined with other background information, there was strong evidence to support the notion that nonresponse to the cognitive items was not a random occurrence.

To enable the estimation of the literacy proficiencies of the nonrespondents, AIR applied both logical and regression-based imputation procedures to the missing responses on the basis of the type of reasons for nonresponse (see Little and Rubin 1987 for various imputation methods). The data concerning reasons for missing cognitive data provided the basis for making logical imputations of what the missing answers would have been had the respondent completed the assessment booklet. In the regression-based imputation strategy, the NAAL background data for the nonrespondents were used to fit the logistic regression models. In each of the logistic regression models, the dependent variables were the items selected to be imputed. The variables used to perform nonresponse and poststratification adjustments were used as predictors. Two different sets of predictor variables were used for respondents from the household sample and from the prison sample because the background questionnaires were somewhat different. Tables 15-5 and 15-6 describe the predictor variables used in the logistic regression models in the Household Study and the Prison Study, respectively.

SAS procedures for logistic regression were used to conduct the logistic regression analyses for each item to be imputed. The estimated regression coefficients were used to predict missing values of the dependent variables. For each nonrespondent, the probability of answering the item correctly was computed and then compared with a randomly generated number between 0 and 1. If the probability of getting a correct answer was greater than the random number, the imputed value for the item was 1 (correct). Otherwise, it was 0 (wrong).

All respondents who answered at least one item on each scale (prose, document, and quantitative) were included in the main NAAL reporting sample only on the basis of their performance on the items they answered, including respondents who answered only core items and were put in the ALSA. For the remaining respondents, the specific imputation procedures and decisions were as follows[2]:

(A) For ALSA respondents who had missing responses to any core items:

[2] The decision to impute only the easiest item on each scale (and no more) was made so as to have the same number of respondents on every scale while limiting estimation error due to imputed data.

- Impute all missing core items as wrong on all three scales. It was logical to assign wrong answers to ALSA nonrespondents because they were the least literate adults and most of the ALSA respondents answered the NAAL questions incorrectly (which was how they were classified into ALSA).

**Table 15-5. Variables used as predictors in the logistic regression models for imputation in the NAAL household study, by value label: 2003**

| Variable (NAME) | Value label |
|---|---|
| Census Region (REGION) | 1: Northeast |
| | 2: Midwest |
| | 3: South |
| | 4: West |
| Metropolitan Statistical Area Status (MSAFLG) | 0: NonMSA |
| | 1: MSA |
| Age From Background Data (AGECAT) | 1: 16–29 |
| | 2: 30–49 |
| | 3: 50–69 |
| | 4: 70+ |
| Gender From Background Data (GENDER_R) | 1: Male |
| | 2: Female |
| Highest Education Level From Background Data (EDUCCAT) | 1: Less than high school |
| | 2: High school diploma or equivalent |
| | 3: More than high school |
| Country of Birth (BORNUSA) | 0: Born elsewhere |
| | 1: Born in the USA |
| Race-Ethnicity From Background Data (BQ_RACETH_RAKE) | 1: Hispanic |
| | 2: Non-Hispanic Black only |
| | 3: Other |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table 15-6. Variables used as predictors in the logistic regression models for imputation in the NAAL prison study, by value label: 2003**

| Variable (NAME) | Value label |
|---|---|
| Prison Security Level (SECURITY_R) | 1: Supermax, Max/close/high |
| | 2: Medium |
| | 3: Minimum/low, Admin, Other |
| Region/Prison Type (REGTYPE) | 1: Northeast |
| | 2: Midwest |
| | 3: South |
| | 4: West |
| | 5: Federal |
| Inmate Age Category (AGECAT_R) | 1: 16–29 |
| | 2: 30–49 |
| | 3: 50+ |
| Inmate Gender (GENDER_R) | 1: Male |
| | 2: Female |
| Inmate Highest Education Level (EDUCCAT) | 1: Less than high school |
| | 2: High school or higher |
| Inmate Country of Birth (BORNUSA) | 0: Born elsewhere |
| | 1: Born in the USA |
| Inmate Race-Ethnicity (BQ_RACETH_RAKE) | 1: Hispanic |
| | 2: Non-Hispanic Black only |
| | 3: Other |
| Inmate Marital Status (MARITAL) | 1: Never Married |
| | 2: Other |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

(B) For NAAL respondents (not put into ALSA) who did not attempt any item on at least one scale for non-literacy-related reasons, impute one item on the scale(s) for which they had no responses (with one exception for the document scale as described below):

- Impute easiest core item on prose (CC004) and quantitative (CC005) scales via the logistic regression models described above.

- Impute easiest health item on document scale (CC007) via the logistic regression model described above (even if another document item had been answered).

(C) For NAAL respondents (not put into ALSA) who did not attempt any item on at least one scale but not all scales for literacy-related reasons:

- Impute one item as wrong on the scale(s) for which they had no responses as discussed in (B) above.

(D) For NAAL respondents who did not attempt any item on any scale for literacy-related reasons:

- Do not impute. Leave as missing.

### 15.4.2 Maintaining Comparability of Population Coverage Between 1992 and 2003 Assessments

The NAAL 2003 reporting sample included respondents from the household survey and the prison survey, including those who were identified as least literate and were administered the ALSA assessment. It was possible and reasonable to include ALSA respondents in the NAAL reporting sample because ALSA respondents met the minimum requirement of answering at least one question on each scale because they all took the seven core NAAL questions before they were classified into ALSA. Another reason to include ALSA in the NAAL reporting sample was to facilitate the identification of a comparable population from the 1992 NALS in order to report trend between the two years.

In the 1992 NALS, approximately 5 percent (weighted) of respondents were unable to complete the simplest literacy tasks in the assessment and were classified in the lowest level of literacy. No separate assessment comparable to ALSA was developed in 1992 to measure the literacy skills of those lowest literacy adults. If the ALSA respondents were not to be included in the NAAL reporting sample, a population comparable to the 2003 ALSA would have to be identified from the 1992 NALS data to ensure that trend could be maintained between the two assessments. AIR explored various ways to identify such a comparable population in the 1992 NALS data and concluded that including ALSA respondents in the 2003 NAAL reporting sample was the most appropriate solution to the problem of identifying the matching 1992 reporting sample for trend study.

As indicated in the previous section, a different imputation strategy was applied in the 1992 NALS from that used in the 2003 NAAL. To report trend to 1992, AIR applied the same imputation procedures used in the 2003 NAAL to the 1992 NALS data and identified a population in the 1992 NALS comparable to the 2003 assessment (i.e., a population that included individuals who answered at least one item on any scale).

The section below describes the specific imputation decisions for the 1992 NALS:

(A) For respondents who did not answer any questions on any scale for non-literacy-related reasons:

- Impute the easiest core item on prose (NC00301), quantitative (NC00501), and document (SCOR100) scales via logistic regression models. The 1992 NALS background data were used to fit the logistic regression models in which the dependent variables were the three imputed items, respectively. The variables used for nonresponse and poststratification adjustments were used as predictors. Tables

15-6 and 15-7 list the predictor variables used in the logistic regression models for the 1992 NALS household study and prison study, respectively. The variables were listed separately for respondents from the household sample and from the prison sample because the background questionnaires were somewhat different.

(B) For respondents who did not answer any questions on any scale for literacy-related reasons:

- Do not impute. Leave as missing.

(C) For respondents who did not answer any questions on at least one scale but not all scales:

- If reasons for nonresponse were not literacy related: Impute the easiest core item on the scale(s) for which they had no responses via logistic regression models as discussed in (A) above.

- If reasons for nonresponse were literacy related: Impute the easiest core item on the scale(s) for which they had no response as wrong.

**Table 15-7. Variables used as predictors in the logistic regression models for imputation in the NALS household study, by value label: 1992**

| Variable (NAME) | Value label |
|---|---|
| Census Region (REGION) | 1: Northeast |
| | 2: Midwest |
| | 3: South |
| | 4: West |
| Age From Background Data (AGE) | 1: 16–29 |
| | 2: 30–49 |
| | 3: 50–69 |
| | 4: 70+ |
| Gender From Background Data (GENDER_R) | 1: Male |
| | 2: Female |
| Highest Education Level From Background Data (EDUCATION) | 1: Less than high school |
| | 2: High school diploma or equivalent |
| | 3: More than high school |
| Race-Ethnicity From Background Data (RACE-ETHNICITY) | 1: Hispanic |
| | 2: Non-Hispanic Black only |
| | 3: Other |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 1992 National Adult Literacy Survey.

**Table 15-8. Variables used as predictors in the logistic regression models for imputation in the NALS prison study, by value label: 1992**

| Variable (NAME) | Value label |
|---|---|
| Prison Security Level (FACTYPE) | 1: Maximum |
| | 2: Medium |
| | 3: Minimum |
| | 4: Other |
| | 5: Medical Center |
| Census Region (REGION) | 1: Northeast |
| | 2: Midwest |
| | 3: South |
| | 4: West |
| Inmate Age Category (AGE) | 1: <30 |
| | 2: Other |
| Inmate Gender (GENDER) | 1: Male |
| | 2: Female |
| Inmate Highest Education Level (EDUCATION) | 1: Less than high school |
| | 2: High school or equivalent |
| | 3: More than high school |
| Inmate Race-Ethnicity (RACE-ETHNICITY) | 1: Hispanic |
| | 2: Non-Hispanic Black only |
| | 3: Other |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 1992 National Adult Literacy Survey.

# CHAPTER 16

# VARIABLE CONSTRUCTION AND FILE DEVELOPMENT

Ying Jin, American Institutes for Research

This chapter discusses the variable construction and file development procedures for the National Assessment of Adult Literacy (NAAL). The discussion begins with the construction and the contents of the NAAL public use data for the Household Study and the Prison Study as well as the NAAL item parameter files, followed by a discussion on the construction procedures for the derived variables (see appendix A for details). Also in this chapter are instructions on how to analyze NAAL data by using AM software (see appendix B for details) and how to use electronic codebooks (see appendix C for details).

## 16.1 FILE CONSTRUCTION PROCEDURES AND FILE CONTENTS

The NAAL study includes three components: (1) a national household sample; (2) a state household sample from six states, called the State Assessment of Adult Literacy (SAAL); and (3) a national sample of prison inmates. The NAAL and SAAL household samples were combined to allow improved precision of statistical estimates. Data from respondents were collected through a Background Questionnaire (BQ) and a literacy assessment. Other types of data were also collected and derived during the NAAL sampling, data collection, and weighting processes. On the basis of the National Center for Education Statistics (NCES) guidelines for data release, two separate public use data files containing the data above were created: one for the Household Study and one for the Prison Study. The household and prison samples were combined for the national estimates.

Using the guidelines to limit the risk of data disclosure provided by NCES standard 4-2, Maintaining Confidentiality, staff conducted the NAAL disclosure control analysis to reduce the risk of data disclosure and to protect the confidentiality of all individuals responding to the NAAL (see chapter 10 for details). Public data files for both the Household Study and the Prison Study were constructed on the basis of the recommendations from this statistical disclosure control analysis and therefore could be released to the public. Restricted-use data files that contained more complete and detailed data, including data that were suppressed in the public data files because of disclosure risk reasons, were available at the NCES website for licensed users only. For procedures to obtain an NCES license for restricted use data, the reader can go to http://nces.ed.gov/StatProg/confid.asp.

### 16.1.1 Household Study

The public use data file for the Household Study contains the following four types of variables:

- ID: Randomly assigned unique case identifier.
- Assessment items: Cognitive tasks from the NAAL literacy assessment on each of the three literacy scales (prose, document, and quantitative). Each task measuring health literacy was also classified as a prose, document, or quantitative task and was included on one of those three scales. Each assessment item variable contains the scored data from respondents (right, wrong, omitted, not-reached, and missing).
- Background variables: Variables containing data collected from the NAAL Household Background Questionnaire (BQ). The NAAL household BQ covered a variety of areas including demographic information, language background, educational background and experiences, health, labor force participation, and so on, as discussed in section 2.2.
- Assessment design variables: Variables related to complex sampling design, including sampling weights, variance stratum variable, and cluster variable. These variables should be used appropriately for the analysis of the NAAL data to obtain the most accurate estimates.

### 16.1.2 Prison Study

Like the public use data file for the Household Study, the data file for the Prison Study contains the four types of variables discussed above. However, the background variables contain data from the NAAL Prison Background Questionnaire, which collected demographic data on inmates and contextual data on their experiences in and/or prior to admission to prison that were related to literacy (see section 2.3 for details).

### 16.1.3 NAAL Item Parameter Files

Two item parameter files, one for the prose, document, and quantitative literacy scales and one for the health literacy scale, were created for the analysis of NAAL data using the AM software (see section 16.3). Such item parameter files are also called AM dictionary files and have an extension of .dct in their file names. The item parameter file for the health literacy scale was created separately from the other three scales because the health items were also classified as prose, document, or quantitative items and confounding those items into one file would cause problems in AM. The item parameter files contain the mean, the standard deviation, and the item parameters from scaling procedures (see chapter 14) for each of the literacy scales. Appendix B describes how to use these files when using AM to analyze the NAAL data.

## 16.2 CONSTRUCTION OF DERIVED VARIABLES

The NAAL collected a large amount of information on many background variables. Some variables, such as gender, can be used directly for reporting purposes. Many variables, however, need to be derived from the raw data directly collected in the assessment for reporting purposes. For example, the NAAL collected information on respondents' date of birth. However, to report the literacy of adults in each age group of interest, a derived age variable needed to be constructed on the basis of the date of an interview and the respondent's date of birth. Derived variables were also constructed to make the trend variables comparable between the 2003 NAAL and the 1992 National Adult Literacy Survey (NALS) for trend analysis.

Appendix A contains the construction procedures of the derived variables that appear in each of the household and prison data files.

## 16.3 ANALYZING 2003 NAAL DATA USING AM SOFTWARE

The NAAL used a complex assessment design that allowed maximum coverage of the broad domain of literacy while minimizing the time burden on any one participant (see chapter 14). Under this design, the NAAL administered only a fraction of the assessment items on each scale to each participant. Although individual participants were required to take only a small portion of the entire pool of assessment questions, the aggregate results across the entire assessment allow broad reporting of literacy for the targeted population. However, because participants did not receive enough literacy tasks to provide reliable information about individual performance, traditional test scores for individual participants will result in misleading estimates of population characteristics and therefore are not appropriate to use for estimates of population statistics. Rather, statistical procedures based on the method of marginal maximum likelihood (MML) need to be used to provide consistent estimates of population statistics from data collected under such design. The usual statistical software packages such as SAS, SPSS or STATA cannot implement MML procedures; therefore, special analysis tools are needed. The AM software, which is available at http://am.air.org/, can be used to implement these procedures. The NAAL data that can be used in AM are available from NCES.

Appendix B has instructions to help first-time AM users analyze the NAAL data. The software also has an interactive help system that explains both how to use the software and the statistical procedures themselves.

## 16.4 USING THE ELECTRONIC CODEBOOK

The NAAL public use data for the Household Study and the Prison Study are available from NCES and can be accessed with an electronic codebook produced by NCES staff. The electronic codebook provides the option of producing SPSS, SAS, and STATA control statements that can be used to create SPSS, SAS, and STATA data files. These control statements also include statements for the following features:

| | |
|---|---|
| VARIABLE DEFINITION | The field names are listed in the order in which they appear on the file. The electronic codebook will produce control statements in SPSS, SAS, or STATA with column positions and input formats. |
| VARIABLE LABEL | This is a 40-character text description for each field. |
| VALUE LABEL | All numeric fields with discrete (or categorical) values have 20-character text descriptors for each value within the variable's range. |

In the electronic codebook, sections with variables for the household sample are arranged first, followed by variables for the prison sample. The data for the household and prison samples are stored in separate files. The electronic codebook also contains unweighted descriptive statistics of all variables. The electronic codebook is available in a Windows version. See appendix C for a more comprehensive user's manual for the electronic codebook.

# CHAPTER 17

# THE NAAL HEALTH LITERACY COMPONENT

Ying Jin and Eric Dunleavy, American Institutes for Research

Understanding the health literacy of America's adults is important because so many aspects of health care, from finding health information to maintaining health, depend on understanding written information. To determine the health literacy of the nation's adults and directly compare it with the measures of the general literacy of the population (i.e., prose, document, and quantitative), the 2003 National Assessment of Adult Literacy (NAAL) included a new component—a health literacy component. This chapter describes the definition of health literacy, construction of the health literacy scale, and reporting of the results on the health literacy scale.

## 17.1 DEFINING HEALTH LITERACY AND DETERMINING THE CONTENT OF THE HEALTH LITERACY COMPONENT

The NAAL health literacy scale and health literacy tasks were guided by the definition of health literacy used by the Institute of Medicine and *Healthy People 2010* (a set of national disease prevention and health promotion objectives led by the U.S. Department of Health and Human Services). This definition states that health literacy is:

> The degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions. (U.S. Department of Health and Human Services 2000 and Institute of Medicine 2004).

The content of the health literacy component was determined by the U.S. Department of Health and Human Services (HHS) in accordance with the public health priorities represented in *Healthy People 2010*, the disease prevention and health promotion agenda for the nation, and in consultation with HHS staff and external health literacy experts. The health literacy component of the NAAL offers a vehicle by which HHS can obtain a baseline measurement of the U.S. population's health literacy skills and project a target for improvement by the end of the decade (see section 2.4.3 in chapter 2).

## 17.2 IDENTIFYING HEALTH LITERACY TASKS

Twenty-eight tasks were included in the health literacy scale. These health literacy tasks represent a range of literacy activities that adults are likely to face in their daily lives. In the development of the health literacy tasks, the Office of Disease Prevention and Health Promotion (ODPHP) within the U.S.

Department of Health and Human Services suggested materials and questions based on input from other HHS agencies and stakeholders and experts, and on information from federal health materials and other health-related assessments.

### 17.2.1 Overlap of the health literacy tasks with the tasks on the prose, document and quantitative scales

The health literacy stimulus materials and the 28 tasks were designed to elicit respondents' skills for locating and understanding health-related information and services and to represent the three general literacy scales—prose, document, and quantitative—developed to report NAAL results. That is to say, the health literacy tasks were developed to fit into the NAAL's prose, document, or quantitative scales but were distinguished from the other tasks on those scales by their health content. The health literacy scale consisted of 12 prose, 12 document, and 4 quantitative NAAL tasks. As described in chapter 2, prose, document, and quantitative literacy were defined as follows:

- *Prose literacy*. The knowledge and skills needed to perform prose tasks (i.e., to search, comprehend, and use information from continuous texts). Prose examples include editorials, news stories, brochures, and instructional materials. Prose texts can be further broken down as expository, narrative, procedural, or persuasive.

- *Document literacy.* The knowledge and skills needed to perform document tasks (i.e., to search, comprehend, and use information from noncontinuous texts in various formats). Document examples include job applications, payroll forms, transportation schedules, maps, tables, and drug and food labels.

- *Quantitative literacy*. The knowledge and skills required to perform quantitative tasks (i.e., to identify and perform computations, either alone or sequentially, using numbers embedded in printed materials). Examples include balancing a checkbook, figuring out a tip, completing an order form, and determining the amount of interest on a loan from an advertisement.

### 17.2.2 Framework of the health literacy scale

Tasks used to measure health literacy were organized around three domains of health and health care information and services: *clinical, prevention,* and *navigation of the health care system*.

The materials were selected to be representative of real-world health-related information, including insurance information, medicine directions, and preventive care information.

Of the 28 health literacy tasks, 3 represented the *clinical* domain, 14 represented the *prevention* domain, and 11 items represented the *navigation of the health care system* domain. The domains are defined in the following way:

- The *clinical* domain encompasses those activities associated with the health care provider-patient interaction, clinical encounters, diagnosis and treatment of illness, and medication. Tasks from the clinical domain are filling out a patient information form for an office visit, understanding dosing instructions for medication, and following a health care provider's recommendation for a diagnostic test.

- The *prevention* domain encompasses those activities associated with maintaining and improving health, preventing disease, intervening early in emerging health problems, and engaging in self-care and self-management of illness. Examples are following guidelines for age-appropriate preventive health services, identifying signs and symptoms of health problems that should be addressed with a health professional, and understanding how eating and exercise habits decrease risks for developing serious illness.

- The *navigation of the health care system* domain encompasses those activities related to understanding how the health care system works and individual rights and responsibilities. Examples are understanding what a health insurance plan will and will not pay for, determining eligibility for public insurance or assistance programs, and being able to give informed consent for a health care service (U.S. Department of Health and Human Services 2003, p. 37).

The NAAL health literacy scale did not include tasks that did not fit the definitions of prose, document, or quantitative literacy even if they were consistent with the definition of health literacy used by *Healthy People 2010*. For example, none of the NAAL health tasks required knowledge of specialized health terminology. The assessment also did not measure the ability to obtain information from non-print sources, although questions about the use of all sources of health information—both written and oral—were included on the background questionnaire.

## 17.3 SCALING THE HEALTH ITEMS

Since health literacy is distinguished from the prose, document and quantitative literacy with its unique health content, a separate health literacy scale was established to represent this unique latent trait. Similar item calibration procedures used for the other three literacy scales (see chapter 14) were carried out for the health items.

### 17.3.1 The scaling model

As discussed in chapter 14, each respondent did not answer all of the NAAL items under the NAAL assessment design and a simple average percent correct would not allow for appropriate reporting of population characteristics. Item response theory (IRT) methods, which models the probability of answering a question correctly as a mathematical function of proficiency or skill, were used to scale the health items as well as the prose, document, and quantitative items. The IRT scaling procedures provide a common scale on which performance on some latent trait can be compared across groups.

IRT models assume that an examinee's performance on each item reflects characteristics of the item and characteristics of the examinee. All models assume that all items on a scale measure a common latent ability or proficiency dimension (e.g., health literacy) and that the probability of a correct response on an item is uncorrelated with the probability of a correct response on another item given fixed values of the latent trait. Items are measured in terms of their difficulty as well as their ability to discriminate among examinees of varying ability.

The same IRT models used for scaling the prose, document, and quantitative items were used for the health literacy scale: the two-parameter logistic (2PL) model and the Graded Response Logistic (GRL) model, depending on the item types (i.e., dichotomous or polytomous). Each model is a "latent variable" model, which expresses respondents' tendencies to achieve certain scores (such as correct/incorrect) on the items contributing to a scale as a function of a parameter that is not directly observed, called proficiency ($\theta$) on the scale. Details of the two scaling models were described in chapter 14.

The health literacy scale was established for the first time in the 2003 NAAL. Therefore, unlike the prose, document, and quantitative scales, there was no previous health assessment scale to link to. The reporting metric for the health literacy scale was set to have a mean of 245 and a standard deviation of 55, in order to resemble the reporting metrics of the prose, document and quantitative scales.

### 17.3.2 Item parameter estimation

As mentioned above, similar item calibration procedures for each of the prose, document and quantitative scales were performed for the health literacy scale. The IRT package of the AM software developed by Cohen et al. (2000) was used to fit the IRT models to the 2003 NAAL health assessment data. The two-parameter logistic item response theory model was adopted for dichotomous items and the Graded Response Logistic item response theory model was used for partial credit items. Model fit was evaluated at the item level by inspecting residuals from fitted item response curves from AM. The item response curves were visually examined by comparing the empirical item response functions (IRFs) with the theoretical curves. As in the scaling for the other NAAL literacy scales, preliminary sample weights were used during the calibration procedures.

One difference in item calibration procedures between the health literacy scale and the other NAAL scales is that, unlike the concurrent calibration procedures for the prose, document and quantitative scales that used the assessment data from both 1992 and 2003 (see chapter 14), the calibration of the health items was conducted only on the 2003 data. This was because the health scale was new to the 2003 assessment and there was no such previous scale in the 1992 assessment.

Estimated item parameters for the health literacy scale are presented in table D-4 in appendix D. As shown in appendix D, the slope or discrimination parameters (parameter a) range from 0.34 to 2.57 for the health literacy scale. The difficulty parameters (parameter b) for dichotomous items range from -7.11 to 1.52. The step parameters for polytomous items range from -2.07 to 1.64. Some items, e.g., CC001 in table D-2 and CC007 in table D-4, had extremely low values of difficulty parameters as such items were designed to be extremely easy to discriminate adults at the very low end of the literacy scale.

## 17.4 REPORTING RESULTS ON THE HEALTH LITERACY SCALE

In addition to average literacy scores, the NAAL results were also reported as the percentage of adults in the pre-defined performance levels. Performance levels are used to identify and characterize the relative strengths and weaknesses of adults falling within various ranges of literacy ability. Describing the adult population according to such levels allows analysts, policy-makers, and others to examine and discuss the typical performance and capabilities of specified groups within the adult population.[1]

### 17.4.1 Establishing health literacy performance levels

In response to the request of the Department of Education, the National Research Council's Board on Testing and Assessment (BOTA) Committee on Performance Levels for Adult Literacy recommended a new set of performance levels for the prose, document, and quantitative scales for the 2003 NAAL assessment, instead of using the same reporting levels used for the 1992 National Adult Literacy Survey. Hauser et al. (2005) described in detail the procedures followed by the BOTA Committee to determine the NAAL performance levels.

[1] For more information on NAAL performance levels, see White and Dillow (2005).

The new set of performance levels for each of the prose, document and quantitative scales are: *Below Basic, Basic, Intermediate,* and *Proficient*. Table 17-1 summarizes the knowledge, skills, and capabilities that adults needed to demonstrate to be classified into one of the four levels on the prose, document, and quantitative scales. However, The BOTA Committee was not asked to recommend performance levels for the health literacy scale. Because every health literacy task was included on the prose, document, or quantitative scale, it was assigned a performance level (*Below Basic, Basic, Intermediate,* and *Proficient*) corresponding to their position on one of those scales.

**Table 17-1. NAAL literacy levels, by definition and key abilities associated with each level: 2003**

| Level and definition | Key abilities associated with level |
|---|---|
| *Below Basic* indicates no more than the most simple and concrete literacy skills.<br><br>Score ranges for *Below Basic*:<br>Prose: less than or equal to 209<br>Document: less than or equal to 204<br>Quantitative: less than or equal to 234 | Adults at the *Below Basic* level range from being nonliterate in English to having the abilities listed below:<br><br>▪ locating easily identifiable information in short, commonplace **prose** texts<br>▪ locating easily identifiable information and following written instructions in simple **documents** (e.g., charts or forms)<br>▪ locating numbers and using them to perform simple **quantitative** operations (primarily addition) when the mathematical information is very concrete and familiar |
| *Basic* indicates skills necessary to perform simple and everyday literacy activities.<br><br>Score ranges for *Basic*:<br>Prose: 210–264<br>Document: 205–249<br>Quantitative: 235–289 | ▪ reading and understanding information in short, commonplace **prose** texts<br>▪ reading and understanding information in simple **documents**<br>▪ locating easily identifiable **quantitative** information and using it to solve simple, one-step problems when the arithmetic operation is specified or easily inferred |
| *Intermediate* indicates skills necessary to perform moderately challenging literacy activities.<br><br>Score ranges for Intermediate:<br>Prose: 265–339<br>Document: 250–334<br>Quantitative: 290–349 | ▪ reading and understanding moderately dense, less commonplace **prose** texts as well as summarizing, making simple inferences, determining cause and effect, and recognizing the author's purpose<br>▪ locating information in dense, complex **documents** and making simple inferences about the information<br>▪ locating less familiar **quantitative** information and using it to solve problems when the arithmetic operation is not specified or easily inferred |
| *Proficient* indicates skills necessary to per-form more complex and challenging literacy activities.<br><br>Score ranges for Proficient:<br>Prose: 340–500<br>Document: 335–500<br>Quantitative: 350–500 | ▪ reading lengthy, complex, abstract **prose** texts as well as synthesizing information and making complex inferences<br>▪ integrating, synthesizing, and analyzing multiple pieces of information located in complex **documents**<br>▪ locating more abstract **quantitative** information and using it to solve multi-step problems when the arithmetic operations are not easily inferred and the problems are more complex |

NOTE: Although the literacy levels share common names with the National Assessment of Educational Progress (NAEP) levels, they do not correspond to the NAEP levels.
SOURCE: Hauser, R.M, Edley, C.F. Jr., Koenig, J.A., and Elliott, S.W. (Eds.). (2005). Measuring Literacy: Performance Levels for Adults, Interim Report. Washington, DC: National Academies Press; White, S. and Dillow, S. (2005). Key Concepts and Features of the 2003 National Assessment of Adult Literacy (NCES 2006-471). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

To determine the performance levels and cut scores on the health literacy scale, “item mapping” procedure was used. With this procedure, the health items were “placed” at a specific point on the health literacy scale based on their level of difficulty. The level of difficulty for each item was estimated as the proficiency level (i.e., score) at which the probability of answering an item correctly was 67 percent. The response probability (RP) of 67 percent – abbreviated as RP67– was the convention used by the BOTA Committee to map items on the prose, document, and quantitative scales (see Hauser et al. 2005). Using a response probability convention, it is possible to use the IRT model to place the items at some specific level on the scale, which in turn allows one to make statements or predictions about the likelihood that a person who scores at the level will answer the questions correctly.

The scale score associated with a 67 percent probability of a correct response was calculated for each health item based on the item parameters obtained from the item calibration process. The items were then rank-ordered based on the RP67 scale scores and mapped to the health literacy scale. Cut-points for performance levels on the health scale were established so that each item was classified into the same level on the health scale as on the respective prose, document, or quantitative scale with which the item was associated.  Table 17-2 presents the item map for the heath literacy items. For dichotomous items, the RP67 scores are associated with a 67 percent probability of answering the item correctly. For partial credit items, the RP67 scores take into account the different score points and are associated with a 67 percent probability of a fully correct response and a partially correct response, respectively.

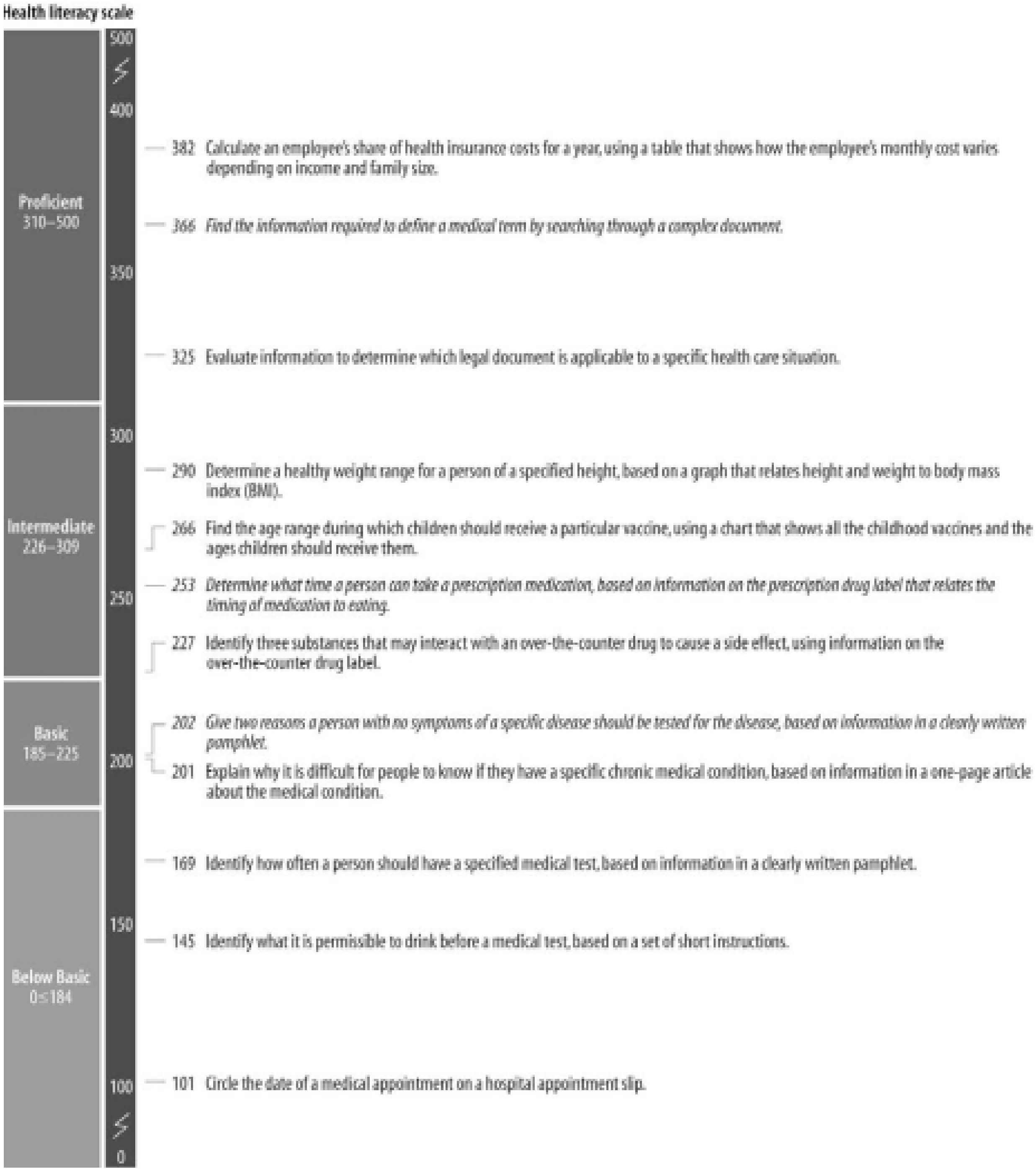
For further illustration, figure 17-1 shows the positions of selected health tasks on the health literacy scale. The position of a question on the scale represents the average scale score attained by adults who had a 67 percent probability of successfully answering the question. Next to each task is an indication of whether it was classified into the *Below Basic*, *Basic*, *Intermediate*, or *Proficient* category on the other scale on which the task was included (prose, document, or quantitative).

## Table 17-2. Item map for the NAAL health literacy items: 2003

| Item ID | Performance level on other scale | Item type | Score category | RP67 Scale score |
|---|---|---|---|---|
| CC007 | Below Basic Document | Dichotomous | 1 | -58.7 |
| CC002 | Below Basic Document | Dichotomous | 1 | 101.3 |
| C070101 | Below Basic Prose | Dichotomous | 1 | 145.3 |
| C071001 | Below Basic Document | Partial credit | 1 | 154.6 |
| C020901 | Below Basic Prose | Partial credit | 1 | 157.9 |
| C030201 | Below Basic Prose | Dichotomous | 1 | 168.8 |
| C030101 | Below Basic Prose | Partial credit | 1 | 169.7 |
| C070901 | Below Basic Prose | Dichotomous | 1 | 177.5 |
| C050801 | Below Basic Prose | Dichotomous | 1 | 178.7 |
| | | | | |
| C020901 | Basic Prose | Partial credit | 2 | 186.3 |
| C040501 | Basic Document | Partial credit | 1 | 191.3 |
| N110101 | Basic Prose | Dichotomous | 1 | 201.0 |
| C040504 | Basic Document | Dichotomous | 1 | 201.5 |
| C030101 | Basic Prose | Partial credit | 2 | 202.3 |
| C030301 | Basic Prose | Dichotomous | 1 | 202.4 |
| C021001 | Basic Document | Dichotomous | 1 | 203.5 |
| C071101 | Basic Prose | Dichotomous | 1 | 224.4 |
| C080101 | Basic Quantitative | Partial credit | 1 | 225.0 |
| | | | | |
| C080201 | Intermediate Document | Dichotomous | 1 | 227.1 |
| C060501 | Intermediate Document | Dichotomous | 1 | 238.8 |
| C040503 | Intermediate Document | Dichotomous | 1 | 250.4 |
| C080101 | Intermediate Quantitative | Partial credit | 2 | 253.1 |
| C040502 | Intermediate Document | Dichotomous | 1 | 258.1 |
| C060601 | Intermediate Document | Dichotomous | 1 | 265.6 |
| C051101 | Intermediate Prose | Dichotomous | 1 | 275.9 |
| C050901 | Intermediate Prose | Dichotomous | 1 | 277.2 |
| N110201 | Intermediate Quantitative | Dichotomous | 1 | 280.8 |
| C021101 | Intermediate Document | Dichotomous | 1 | 290.0 |
| C040601 | Intermediate Quantitative | Dichotomous | 1 | 290.3 |
| C040501 | Intermediate Document | Partial credit | 2 | 296.5 |
| | | | | |
| C051001 | Proficient Prose | Dichotomous | 1 | 324.7 |
| C071101 | Proficient Prose | Partial credit | 2 | 366.1 |
| C040801 | Proficient Quantitative | Dichotomous | 1 | 382.1 |

NOTE: For dichotomous items, the score category of 1 corresponds to the score point for correct responses. For partial credit items, the score category of 1 corresponds to the score point for partially correct responses, and the score category of 2 corresponds to the score point for fully correct responses. The RP67 scale score for item CC007 is extremely low (negative) because that item was designed to be extremely easy to discriminate adults at the very low end of the literacy scale.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Figure 17-1. Difficulty of selected NAAL health literacy tasks: 2003**




NOTE: The position of a question on the scale represents the average scale score attained by adults who had a 67 percent probability of successfully answering the question. Only selected questions are presented. Scale score ranges for performance levels are referenced on the figure. Regular type denotes a dichotomous item. Italic type denotes a partial credit item.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

For example, as shown in figure 17-1, a task that requires a respondent to “evaluate information to determine which legal document is applicable to a specific health care situation” mapped to 325 on the health scale. That task was also included on the prose scale and, on the basis of the performance levels set by the BOTA Committee on Performance Levels for Adult Literacy, it mapped to the *Proficient* category on the prose scale. On the health scale, the cut-point between *Proficient* and *Intermediate* was set so that this task would fall into the *Proficient* category. That is, a health literacy task that was mapped to the *Proficient* level on the prose scale was also mapped to the *Proficient* level on the health scale.

Similarly, as shown in figure 17-1, a task that requires a respondent to “determine a healthy weight range for a person of a specified height, based on a graph that relates height and weight to body mass index (BMI)” mapped to 290 on the health scale. This task was also included on the document scale, where it was classified as *Intermediate*. The cut-points for the health scale were set so that the task would also map to the *Intermediate* level on the health scale.

As shown in figure 17-1, *Below Basic* health tasks required locating straightforward pieces of information in short simple texts or documents. Adults with literacy near the top of the *Below Basic* health literacy level were more likely to accomplish these tasks than adults who placed at the bottom of the *Below Basic* level.

Health tasks that mapped to the *Basic* level generally required finding information in texts and documents that were somewhat longer than those in the *Below Basic* level, and the information to be found was usually more complex. For example, a task that mapped to the *Basic* level required giving two reasons a person with no symptoms of a specific disease should be tested for the disease by using information in a pamphlet, while a task that mapped to the *Below Basic* level required finding one piece of information–the date–on a medical appointment slip that was shorter and simpler than the text in the *Basic* task. Health tasks that mapped to the *Intermediate* level went beyond simply searching texts and documents to find information. Most health tasks that mapped to the *Intermediate* level required adults to interpret or apply information that was presented in complex graphs, tables, or other health-related texts or documents. Health tasks that mapped to the *Proficient* level required drawing abstract inferences, comparing or contrasting multiple pieces of information within complex texts or documents, or applying abstract or complicated information from texts or documents.

Based on the procedures discussed above, the score ranges for each of the health performance levels were as follows: *Below Basic* 0-184, *Basic* 185-225, *Intermediate* 226-309, and *Proficient* 310-500.

### 17.4.2 Relationship of health scale to prose, document, and quantitative scales

Given the overlap of health items with prose, document, and quantitative items, it was expected that the correlations between the health literacy scale and the prose, document, and quantitative scales would be quite high. For reference purposes, table 17-3 presents these correlations.

**Table 17-3. Correlations among NAAL literacy scales, by literacy scale: 2003**

| Literacy scale | Health | Prose | Document | Quantitative |
|---|---|---|---|---|
| Health | 1.000 | | | |
| Prose | 0.940 | 1.000 | | |
| Document | 0.998 | 0.857 | 1.000 | |
| Quantitative | 0.891 | 0.875 | 0.894 | 1.000 |

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

## REFERENCES


Allen, N.L., Carlson, J.E., and Zelenak, C.A. (1999). *The NAEP 1996 Technical Report* (NCES 1999-452). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Bader, L.A. (2002). *Bader Reading and Language Inventory.* Upper Saddle River, MN: Prentice Hall.

Baker, D.W., Gazmararian, J.A.,Williams, M.V., Scott, T., Parker, R.M., Green, D., Ren, J., and Peel, J. (2004). Health Literacy and Use of Outpatient Physician Services by Medicare Managed Care Enrollees. *Journal of General Internal Medicine, 19*(3): 215–220.

Barton, D., and Hamilton, M. (1998). *Local Literacies: Reading and Writing in One Community.* London: Routledge.

Berndt, E.K., Hall, B.H., Hall, R.E., and Hausman, J.A. (1974). Estimation and Inference in Nonlinear Structural Models. *Annals of Economic and Social Measurement, 4*: 653–665.

Binder, D.A. (1983). On the Variances of Asymptotically Normal Estimates for Complex Surveys. *International Statistical Review, 51*(3): 279–292.

Birnbaum, A. (1968). Some Latent Trait Models and Their Use in Inferring an Examinee's Ability. In F.M. Lord and M.R. Novick (Eds.), *Statistical Theories of Mental Test Scores.* Reading, MA: Addison-Wesley Publishing.

Burke, J., Mohadjer, L., Green, J., Waksburg, J., Kirsh, I., and Kolstad, A. (1994). Composite Estimation in National and State Surveys. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 873–878.

Burns, P., and Roe, B.D. (2002). *Informal Reading Inventory: Preprimer to Twelfth Grade* (sixth edition). Boston: Houghton Mifflin.

Campbell, A., Kirsch, I.S., and Kolstad, A. (1992). *Assessing Literacy: The Framework for the National Adult Literacy Survey (*NCES 92113). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Cochran, W.G. (1977). *Sampling Techniques* (third edition). New York: John Wiley & Sons.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (second edition). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J., Jiang, T., Gaidurguis, A., and Hollender, D. (2000). AM (Software for marginal maximum likelihood and other statistical analyses of data from complex samples), Washington, DC: American Institutes for Research. See website at http://am.air.org.

CTB/McGraw-Hill. (2000). *Fox in a Box: An Adventure in Literacy*. Monterey, CA: Author.

Cummins, J. (1979). Linguistic Interdependence and the Educational Development of Bilingual Children. *Review of Educational Research*, *49*(2): 222–251.

Ericsson, K., and Simon, H. (1980). Verbal Reports as Data. *Psychological Review, 87*(3): 215–251.

Fuchs, L., Fuchs, D., Hosp, M., and Jenkins, J. (2001). Oral Reading Fluency as an Indicator of Reading Competence: A Theoretical, Empirical, and Historical Analysis. *Scientific Studies of Reading*, *5*(3): 239–256.

Godambe, V.P. (1960). An Optimum Property of Regular Maximum Likelihood Estimation. *Annals of Mathematical Statistics, 31*(4): 1208–1212.

Godambe, V.P. (1991). *Estimating Functions*. Oxford, U.K.: Clarendon Press.

Godambe, V.P. and Kale, B.K. (1991). Estimating functions: An overview. In *Estimating Functions* (V.P. Godambe, Ed.), pp. 3-20, Oxford, U.K.: Clarendon Press.

Godambe, V.P., and Thompson, M.E. (1986). Parameters of Superpopulation and Survey Population: Their Relationships and Estimation. *International Statistical Review, 54*(2): 127–138.

Good, R.H., Kaminski, R.A., and Dill, S. (2002). DIBELS Oral Reading Fluency. In R.H. Good and R.A. Kaminski (Eds.), *Dynamic Indicators of Basic Early Literacy Skills* (sixth edition). Eugene, OR: Institute for the Development of Educational Achievement. From http://dibels.uoregon.edu.

Gordon, M.M., Hampson, R., Capell, H.A., and Madhok, R. (2002). Illiteracy in Rheumatoid Arthritis Patients as Determined by the Rapid Estimate of Adult Literacy in Medicine (REALM) Score. *Rheumatology (Oxford),41*(7): 750–754.

Groves, R.M., Dillman, D.A., Eltinge, J.L., and Little, R.J.A. (2001). *Survey Nonresponse*. New York: John Wiley & Sons, Inc.

Halliday, M.A.K. (2002–). *The Collected Works of M. A. K . Halliday*. Ed. by Jonathan Webster. London: Continuum.

Halliday, M.A.K., and Hasan, R. (1985). *Language, Context, and Text: Aspects of Language in a Social-Semiotic Perspective.* Oxford: Oxford University Press.

Hambleton, R.K., Swaminathan, H., and Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage.

Hambleton, R.K., and Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Kluwer Nijhoff Publishing.

Hamilton, M. (2000). Exploring Literacy as Social Practice through Media Photographs. In D. Barton, M. Hamilton, and R. Ivanic (Eds.), *Situated Literacies* (pp 16–34). London: Routledge.

Hauser, R.M, Edley, C.F. Jr., Koenig, J.A., and Elliott, S.W. (Eds.). (2005). Measuring Literacy: Performance Levels for Adults, Interim Report. Washington, DC: National Academies Press

Henning, G. (1987). *A Guide to Language Testing: Development, Evaluation, Research*. Boston: Heinle & Heinle Publishers.

Hill, C., and Parry, K. (1992). The Test at the Gate: Models of Literacy in Reading Assessment. *TESOL Quarterly, 26*(3): 433–461.

Holland, P.W., and Thayer, D.T. (1988). Differential Item Performance and Mantel-Haenszel. In H. Wainer, and H. Braun (Eds.). *Test Validity*. Hillsdale, NJ: Lawrence Erlbaum.

Institute of Medicine (2004). *Health literacy: A Prescription to End Confusion.* Washington, DC: Institute of Medicine, Board on Neuroscience and Behavioral Health, Committee on Health Literacy.

Johns, J. (2001). *Basic Reading Inventory: Pre-Primer through Grade 12 & Early Literacy Assessments* (eighth edition). New York: Kendall Hunt Publishing Company.

Kirsch, I., Yamamoto, K., Norris, N., Rock, D., Jungeblut, A., O'Reilly, P., Berlin, M., Mohadjer, L., Waksberg, J., Goksel, H., Burke, J., Rieger, S., Green, J., Klein, M., Campbell, A., Jenkins, L., Kolstad, A., Mosenthal, P., and Baldi, S. (2001). *Technical Report and Data File User's Manual for the 1992 National Adult Literacy Survey* (NCES 2001-457). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

Kucera, H., and Francis, W.N. (1967). *Computational Analysis of Present-Day American English.* Providence, RI: Brown University Press.

Leslie, L. and Caldwell, J. (2001). *Qualitative Reading Inventory-3*. New York: Addison Wesley Longman, Inc.

Lindau, S.T., Tomori, C., McCarville, M.A., and Bennett, C.L. (2001). Improving Rates of Cervical Cancer Screening and Pap Smear Follow-Up for Low-Income Women With Limited Health Literacy. *Cancer Investigation, 19*(3): 316–323.

Little, R.J.A., and Rubin, D.B. (1987). *Statistical Analysis with Missing Data*. New York: J. Wiley & Sons.

Lord, F.M. (1952). *A Theory of Test Scores.* Iowa City, IA: The Psychometric Society No. 7.

Marcus, M.P., Santorini, B., and Marcinkiewicz, M. (1993). Building a Large Annotated Corpus of English: The Penn Treebank. *Computational Linguistics, 19*(2): 313–330.

Mislevy, R.J. (1984). Estimating Latent Distributions. *Psychometrika*, *49*(3): 359–381.

Mislevy, R.J. (1985). Estimation of Latent Group Effects. *Journal of the American Statistical Association*, *80*(392): 993–997.

Mislevy, R.J. (1991). Randomization-Based Inference About Latent Variables from Complex Samples. *Psychometrika*, *56*(2): 177–196.

Mislevy, R.J., Beaton, A.E., Kaplan, B., and Sheehan, K.M. (1992). Estimating Population Characteristics From Sparse Matrix Samples of Item Responses. *Journal of Educational Measurement, 29*(2): 133–161.

Muraki, E., and Bock, R. D. (2003). *PARSCALE 4: IRT item analysis and test scoring for rating-scale data* [computer program]. Chicago, IL: Scientific Software.

Oh, H.L., and Scheuren, F.J. (1987). Modified Raking Ratio Estimation. *Survey Methodology, 13*(2): 209-219.

Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests*. Copenhagen: Denmarks Paedagogiske Institut.

Reder, S and Edmonston, B. (2000). *Demographic Changes and Literacy Development in a Decade* (NCES 2000-09). U.S. Department of Education. Washington, DC: National Center for Education Statistics Working Paper.

Sabatini, J.P., and Venezky, R.L. (2000b). *Study of Adult Reading Acquisition (SARA): Test Development and Technical Data.* Philadelphia: University of Pennsylvania, National Center on Adult Literacy.

Sabatini, J.P., Venezky, R.L., Kharik, P., and Jain, R. (2000a). *Cognitive Assessments for Low Literate Adults: An Analytic Review and New Framework* (Technical Report TR00-01). Philadelphia: University of Pennsylvania, National Center on Adult Literacy.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometric Monograph*, No. 17.

Samejima, F. (1972). A general model for free-response data. *Psychometric Monograph,* No. 18.

Shanker, J.L., and Ekwall, E.E. (2000). *Ekwall/Shanker Reading Inventory*. Boston: Allyn & Bacon.

Sheehan-Holt, J. K., and Smith, M. C. (2000). Does Basic Skills Education Affect Adults' Literacy Proficiencies and Reading Practices? *Reading Research Quarterly, 35*(2): 226–243.

Sherman, R., Condelli, L., and Koloski, J. (1999). *Developing the National Assessment of Adult Literacy: Recommendations from Stakeholders* (NCES 1998–17). U.S. Department of Education. Washington, DC: National Center for Education Statistics Working Paper.

Shinn, M. R. (Ed.). (1989). *Curriculum-Based Measurement: Assessing Special Children*. New York: Guildford Press.

Silvaroli, N., and Wheelock, W.H. (2001). *Classroom Reading Inventory, Form C* (tenth edition). Boston: McGraw Hill.

Street, B. (1998). New Literacies in Theory and Practice: What are the Implications for Language in Education? *Linguistics and Education, 10*(1): 1–24.

Street, B. (2001). Introduction. In B. Street (Ed.), *Literacy and Development: Ethnographic Perspectives* (pp. 21–26). London: Routledge.

Swanson, H.L., Trainin, G., Necoechea, D.M., and Hammill, D.D. (2003). Rapid Naming, Phonological Awareness, and Reading: A Meta-Analysis of the Correlation Evidence. *Review of Educational Research, 73*(2), 407-440.

Thomas, N. (1993). Asymptotic Corrections for Multivariate Posterior Moments with Factored Likelihood Functions. *Journal of Computational and Graphical Statistics*, *2*(3): 309–322.

Texas Education Agency. (2001-2002). *Texas Primary Reading Inventory (TPRI)*. Austin, TX: Texas Education Agency.

U.S. Department of Education, National Center for Education Statistics. (2003). *NCES Statistical Standards* (NCES 2003-601). Washington, DC: Author.

U.S. Department of Health and Human Services (HHS). (2000). *Healthy People 2010: Understanding and Improving Health.* Washington, DC: U.S. Government Printing Office. Retrieved May 23, 2007 from http://www.healthypeople.gov/Publications/.

U.S. Department of Health and Human Services (HHS). (2003). *Communicating Health: Priorities and Strategies for Progress*. Washington, DC. Retrieved June 15, 2006 from http://odphp.osophs.dhhs.gov/projects/HealthComm/.

U.S. Office of Management and Budget. (1997, October 30). Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity. *The Federal Register*. Washington, DC: Government Printing Office.

White, S. and Dillow, S. (2005). *Key Concepts and Features of the 2003 National Assessment of Adult Literacy* (NCES 2006-471). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

White, S., and McCloskey, M. (2006 forthcoming). *Framework for the 2003 National Assessment of Adult Literacy* (NCES 2006-473). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Widdowson, H.G. (1983). *Learning Purpose and Language Use*. Oxford: Oxford University Press.

Widdowson, H.G. (1990). *Aspects of Language Teaching*. Oxford: Oxford University Press.

Williams, M.V., Davis, T., Parker, R.M., and Weiss, B.D. (2002). The Role of Health Literacy in Patient-Physician Communication. *Family Medicine, 34*(5), 383–389.

Woods, M.L., and Moe, A.J. (2003). *Analytical Reading Inventory* (eighth edition). New Jersey: Merrill/Prentice Hall.

Yamamoto, K. (2001). Estimating Literacy Proficiencies With and Without Cognitive Data. In I. Kirsch, K. Yamamoto, N. Norris, D. Rock, A. Jungeblut, P. O'Reilly, M. Berlin, L. Mohadjer, J. Waksberg, H. Goksel, J. Burke, S. Rieger, J. Green, M. Klein, A. Campbell, L. Jenkins, A. Kolstad, P. Mosenthal, and S. Baldi (Eds.) *Technical Report and Data File User's Manual For the 1992 National Adult Literacy Survey* (NCES 2001–457) (pp.142-164). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

# Appendix A

# DERIVED VARIABLES

# Appendix A: Derived Variables

This appendix describes how the derived variables that appear in each of the NAAL household and prison public use data files were constructed. References such as A-3 in the description for DAGE and K-3 in the description for DRACE for the household sample correspond to the question ID on the 2003 NAAL household background questionnaire (BQ). Similarly, references such as A-3 in the description for DAGEC and J-1 in the description for DRACE for the prison sample correspond to the question ID on the 2003 NAAL prison background questionnaire. The 2003 NAAL household and prison background questionnaires can be found at http://nces.ed.gov/NAAL/index.asp?file=DesignDevelop/SInstruments/BackQuestion.asp&PageID=116.

## 1. Derived variables for the NAAL 2003 household sample

**DAGE (Age: 6 Categories)**

Recoded from a continuous age variable derived from date of interview and the BQ response for date of birth, A-3. Missing values were filled in by using the age information from the Screener.

**DAGEC (Age: 4 Categories)**

Recoded from a continuous age variable derived from date of interview and the BQ response for date of birth, A-3. Missing values were filled in by using the age information from the Screener.

**DRACE (Race/Ethnicity)**

Derived from K-3 and K-5. If K-3 = 1, then DRACE = 3. Otherwise, if K-5A = 1 and K-5B-E = missing, then DRACE = 1; if K-5B = 1 and (K-5A and K-5C-E = missing), then DRACE = 2. Otherwise, DRACE = 4.

Note:

1. White, Black and Other categories include no Hispanics.

2. When the Ethnicity question K-3 = 7/8, the ethnicity indicated in the Screener was substituted.

   Likewise, if K-5A-E = 7/8, the race indicated in the Screener was substituted. Note that the data on race and ethnicity from the Screener may have not been reported by the respondent. That is, the person answering the Screener could have answered the race and ethnicity questions for another person in the household who was selected as a respondent. In these cases, it is not completely certain that the respondent would agree with the race/ethnicity designations recorded during the screener.

**DARRIVE (Age Arrived In The U.S.)**

Derived from A-1 and A-1A. If A-1 = “UNITED STATES (50 STATES OR DC),” then DARRIVE = 1. Otherwise, recode continuous responses for A-1A to categorical. If A-1A = missing, then DARRIVE = missing.

**DENGAGE (Age Learned To Speak English)**

Derived from A-8. If A-8 = missing, then DENGAGE = 1. These are people who speak only English. Otherwise, if A-8 = 1/2, then DENGAGE = 2; if A-8 = 3/4/5, then DENGAGE = 3; if A-8 = 95, then DENGAGE = 4; if A-8 = 97/98, then DENGAGE = missing.

**DLIVEUS (Years Living In The U.S.)**

Derived from A-1 and A-2. If A-1 = “UNITED STATES (50 STATES OR DC),” then DLIVEUS = 1. Otherwise, if A-2 = 1, then DLIVEUS = 2; if A-2 = 2/3/4/5/6/7/8, then DLIVEUS = 3; if A-2 = missing, then DLIVEUS = missing.

**DCBIRTH (Country Of Birth)**

Derived from A-1. If A-1 = “UNITED STATES (50 STATES OR DC),” then DCBIRTH = 1. Otherwise, DCBIRTH = 2.

**DMCBIRTH (Mother's Country Of Birth)**

Derived from G-1. If G-1 = “UNITED STATES (50 STATES OR DC),” then DMCBIRTH = 1; if G-1 = "99999999999999999999999999999999999999999999999997/99999999999999999999999999999999 99999999999999998", then DMCBIRTH = missing. Otherwise, DMCBIRTH = 2.

**DFCBIRTH (Father's Country Of Birth)**

Derived from G-3. If G-3 = "UNITED STATES (50 STATES OR DC)", then DFCBIRTH = 1; if G-3 = “99999999999999999999999999999999999999999999999997/99999999999999999999999999999999 99999999999999998,” then DFCBIRTH = missing. Otherwise, DFCBIRTH = 2.

**D1STLAN (Language Spoken Before School: 5 Categories)**

Derived from A-6. Note: Languages that all apply were selected from lookup table and reported in 5 initial variables and 5 follow-up variables when “Other, Specify” was reported in the 5 initial variables. All 10 variables were cross-checked and recoded as follows: If the respondent speaks English only, then D1STLAN = 1; if the respondent speaks both English and Spanish, regardless whether he/she also speaks another language(s), D1STLAN = 2; if the respondent speaks English and another language(s) but not Spanish, D1STLAN = 3; if the respondent speaks Spanish only or Spanish plus another language(s) but not English, D1STLAN = 4; if the respondent speaks neither English nor Spanish, D1STLAN = 5.

**D1STLANC (Language Spoken Before School: 3 Categories)**

Derived from A-6. Note: Languages that all apply were selected from lookup table and reported in 5 initial variables and 5 follow-up variables when “Other, Specify” was reported in the 5 initial variables. All 10 variables were cross-checked and recoded as follows: If the respondent speaks English only, then D1STLANC = 1; if the respondent speaks English and another language(s), D1STLANC = 2; if the respondent does not speak English, D1STLANC = 3.

**DHMLANG (Language Spoken At Home When Growing Up: 5 Categories)**

Derived from A-5. Note: Languages that all apply were selected from lookup table and reported in 5 initial variables and 5 follow-up variables when “Other, Specify” was reported in the 5 initial variables. All 10 variables were cross-checked and recoded as follows: If the respondent speaks English only, then DHMLANG = 1; if the respondent speaks both English and Spanish, regardless whether he/she also speaks another language(s), DHMLANG = 2; if the respondent speaks English and another language(s) but not Spanish, DHMLANG = 3; if the respondent speaks Spanish only or Spanish plus another language(s) but not English, DHMLANG = 4; if the respondent speaks neither English nor Spanish, DHMLANG = 5.

**DHMLANGC (Language Spoken At Home When Growing Up: 3 Categories)**

Derived from A-5. Note: Languages that all apply were selected from lookup table and reported in 5 initial variables and 5 follow-up variables for “Other, Specify” in the 5 initial variables. All 10 variables were cross-checked and recoded as follows: If the respondent speaks English only, then DHMLANGC = 1; if the respondent speaks English and another language(s), DHMLANGC = 2; if the respondent does not speak English, DHMLANGC = 3.

**DLANGRW (Language First Learned To Read And Write: 3 Categories)**

Derived from A-7. If A-7 = “English,” then DLANGRW = 1; if A-7 = “Spanish,” then DLANGRW = 2; if A-7 = “9999999999999999999999999999999999999997/9999999999999999999999999999999999999998,” then DLANGRW = missing; if A-7 = missing, then DLANGRW = 1 (these are respondents who reported “English only” to A-5 and A-6). Otherwise, DLANGRW = 3.

**DLANGRWC (Language First Learned To Read And Write: 2 Categories)**

Derived from A-7. If A-7 = “English,” then DLANGRWC = 1; if A-7 = “9999999999999999999999999999999999999997/9999999999999999999999999999999999999998,” then DLANGRWC = missing; if A-7 = missing, then DLANGRWC = 1 (these are respondents who reported “English only” to A-5 and A-6). Otherwise, DLANGRWC = 2.

**DEDBFUS (Education Before Coming To The U.S.)**

Derived from A-4. If A-4 = 1/2, then DEDBFUS = 1; if A-4 = 3, then DEDBFUS = 2; if A-4 = 4/5/6, then DEDBFUS = 3; if A-4 = missing, then DEDBFUS = 98; if A-4 = 97/98, then DEDBFUS = missing.

**DEDATTN (Educational Attainment: 9 Categories)**

Derived from B-1. Recoded as follows: If B-1 = 1, then DEDATTN = 1; if B-1 = 2/3, then DEDATTN = 2; if B-1 = 4, then DEDATTN = 3; if B-1 = 5, then DEDATTN = 4; if B-1 = 6, then DEDATTN = 5; if B-1 = 7, then DEDATTN = 6; if B-1 = 8/9, then DEDATTN = 7; if B-1 = 10, then DEDATTN = 8; if B-1 = 11/12, then DEDATTN = 9; if B-1 = 97/98, then DEDATTN = missing.

**DEDATTNC (Educational Attainment: 6 Categories)**

Derived from B-1. Recoded as follows: If B-1 = 1, then DEDATTNC = 1; if B-1 = 2/3, then DEDATTNC = 2; if B-1 = 4, then DEDATTNC = 3; if B-1 = 5, then DEDATTNC = 4; if B-1 = 6/7/8/9, then DEDATTNC = 5; if B-1 = 10/11/12, then DEDATTNC = 6; if B-1 = 97/98, then DEDATTNC = missing.

**DHSAGE (Age Upon Graduation From High School Or Obtaining A Ged)**

Derived from A-3, B-1 and B-2. DHSAGE was assigned only to respondents who completed high school/obtained a GED: B-1 > 3. If B-1 < = 3, then DHSAGE = 98.

For respondents with B-1 >3:

- Date of graduation was set as June 30 in the year indicated in B-2.
- Respondent’s date of birth (A-3) was subtracted from date of graduation

- This yielded an age expressed in years and months (e.g. 18.5 = 18 years, 6 months). Ages not expressed in whole numbers were rounded down.

- Ages were then recoded into reporting categories 1 and 2. The lower boundary for category 1 was 12 years of age.

If B-1 = 97/98 or if A-3 = 99999997/99999998 or if B-2 = 9997/9998, then DHSAGE = missing.

**DMED (Mother's Educational Attainment: 8 Categories)**

Derived from G-2 as follows: if G-2 = 1/2, then DMED = 1; if G-2 = 3, then DMED = 2; if G-2 = 4, then DMED = 3; if G-2 = 5, then DMED = 4; if G-2 = 6, then DMED = 5; if G-2 = 7/8, then DMED = 6; if G-2 = 9, then DMED = 7; if G-2 = 10/11, then DMED = 8; if G-2 = 97/98, then DMED = missing.

**DMEDC (Mother's Educational Attainment: 5 Categories)**

Recoded from G-2 as follows: if G-2 = 1/2, then DMEDC = 1; if G-2 = 3, then DMEDC = 2; if G-2 = 4, then DMEDC = 3; if G-2 = 5/6/7/8, then DMEDC = 4; if G-2 = 9/10/11, then DMEDC = 5; if G-2 = 97/98, then DMEDC = missing.

**DFED (Father's Educational Attainment: 8 Categories)**

Derived from G-4 as follows: if G-4 = 1/2, then DFED = 1; if G-4 = 3, then DFED = 2; if G-4 = 4, then DFED = 3; if G-4 = 5, then DFED = 4; if G-4 = 6, then DFED = 5; if G-4 = 7/8, then DFED = 6; if G-4 = 9, then DFED = 7; if G-4 = 10/11, then DFED = 8; if G-4 = 97/98, then DFED = missing.

**DFEDC (Father's Educational Attainment: 5 Categories)**

Recoded from G-4 as follows: if G-4 = 1/2, then DFEDC = 1; if G-4 = 3, then DFEDC = 2; if G-4 = 4, then DFEDC = 3; if G-4 = 5/6/7/8, then DFEDC = 4; if G-4 = 9/10/11, then DFEDC = 5; if G-4 = 97/98, then DFEDC = missing.

**DWEEKWG (Weekly Wage: Previous Week)**

Derived from D-1, D-3, D-3A and D-4. Note: DWEEKWG was assigned to full-time employees only. If a respondent was not employed full time, DWEEKWG = 98. To determine if a respondent was employed full time, use responses to D-1A, D-1C, D-1E and D-1F. If a respondent indicated in D-1A, D-1C, D-1E and D-1F that she/he was employed, she/he was employed full time. For full time employees:

- If reported pay was gross pay, i.e., if D-3A = 2, then reported income was converted to weekly gross pay as follows:

  if D-3 (Unit) = 1, then weekly gross pay = D-3 (Dollar amount)*40;

  if D-3 (Unit) = 2, then weekly gross pay = D-3 (Dollar amount)*5;

if D-3 (Unit) = 3, then weekly gross pay = D-3 (Dollar amount);

if D-3 (Unit) = 4, then weekly gross pay = D-3 (Dollar amount)/2;

if D-3 (Unit) = 5, then weekly gross pay = D-3 (Dollar amount)/4.3;

if D-3 (Unit) = 6, then weekly gross pay = D-3 (Dollar amount)/52;

if D-3 (Unit) = 91, then do as follows:

If “Other, specify” of D-3 (Unit) = ‘EVERY 15TH AND 31’/‘1700 FOR 15 DAY PERIOD (GET PAID 2 A MTH’/‘TWICE PER MO’/‘TWICE A MONTH’/‘1ST AND 15$^{TH}$’/‘2XMO’/‘2XMO.’/ ‘BIMONTHLY’/‘BI MONTHLY’, then weekly gross pay = D-3 (Dollar amount)/(4.3*0.5).

- If reported pay was take home rather than gross, i.e., if D-3A = 1, reported income was converted to weekly gross pay in two steps.

Step 1: Convert reported income to weekly take home pay as follows:

if D-3 (Unit) = 1, then weekly take home pay = D-3 (Dollar amount)*40;

if D-3 (Unit) = 2, then weekly take home pay = D-3 (Dollar amount)*5;

if D-3 (Unit) = 3, then weekly take home pay = D-3 (Dollar amount);

if D-3 (Unit) = 4, then weekly take home pay = D-3 (Dollar amount)/2;

if D-3 (Unit) = 5, then weekly take home pay = D-3 (Dollar amount)/4.3;

if D-3 (Unit) = 6, then weekly take home pay = D-3 (Dollar amount)/52;

if D-3 (Unit) = 91, then do as follows:

If "Other, specify" of D-3 (Unit) = 'EVERY 15TH AND 31'/'1700 FOR 15 DAY PERIOD (GET PAID 2 A MTH'/'TWICE PER MO'/'TWICE A MONTH'/'1ST AND 15TH'/'2XMO'/'2XMO.'/'BIMONTHLY'/'BI MONTHLY', then weekly take home pay = D-3 (Dollar amount)/(4.3*0.5).

Step 2: Add the following tax withholding adjustments to weekly take home pay to estimate weekly gross pay:

1. Add FICA adjustment (Social and Medicare) at a flat rate of 7.65 percent.

2. Add adjustment based on IRS withholding tables for single taxpayers in 2003 (IRS Form Pub15-T, Table 1 (Weekly Payroll Period), (a) Single Person).

3. Add proxy adjustment for state taxes and miscellaneous deductions at a rate of 10 percent.

- If D-3A = missing or 7, then weekly gross pay = missing.

Continuous weekly gross pay was then rounded and recoded to categorical DWEEKWG.

**DINCOME (Income Adequacy)**

Derived for household sample only from CALCAGEA - CALCAGEY from the Screener, and H-1, H-2 and K-2 from the Background Questionnaire, using the table "Poverty Thresholds for 2003 by Size of Family and Number of Related Children Under 18 Years (Dollars)" published by the Census Bureau. Steps:

1. Use CALCAGEA - CALCAGEY to determine the total number of people in the home, the number of related children under age 18 in the home, and the number of people age 65 and over in the home. The combined information would determine the "Family type" in the Poverty Thresholds Table. Note: When there were two persons in the household, the following guidelines were used to determine who the householder was and hence the "family type":

    i. If neither of the 2 persons was a child, and both people were under 65, then family type = "Two persons, Householder under 65, No children."

    ii. If 1 person was a child and the other person was under 65, then family type = "Two persons, Householder under 65, 1 child."

    iii. If neither of the 2 persons was a child, and 1 person was over 65 and the other person was under 65, then family type = "Two persons, Householder 65 and over, No children."

    iv. If 1 of the 2 persons was a child, and the other person was over 65, then family type = "Two persons, Householder 65 and over, 1 child."

2. Compare the lower boundary of the income range reported in K-2 with the poverty thresholds in the Poverty Thresholds Table for the appropriate family type. If the lower boundary of K-2 was less than the corresponding poverty threshold, then DINCOME = 1. Otherwise, DINCOME = 2.

Note: If K-2 = 97/98, use the follow-up probes to identify the range of income by first recoding them into the following categories:

1 = <$10k
2 = $10k – $15k
3 = <$15k
4 = $15k – $30k
5 = <$20k
6 = $20k – $30k
7 = <$30k
8 = $30k – $40k
9 = $30k – $60k
10 = <$40k
11 = $40k – $60k
12 = <$60k
13 = $60k – $100k
14 = <$100k
15 = over $30k
16 = over $60k
17 = over $100k

Categories 5 (< 20k), 7 (< 30k), 10 (< 40k), 12 (< 60k), and 14 (< 100k) were treated as missing. Then the lower boundaries of the above income categories were compared to the poverty thresholds corresponding to the appropriate family type to create DINCOME.

**DMARITAL (Marital Status)**

Derived from I-1. If I-1 = 1, then DMARITAL = 1; if I-1 = 2/3/4, then DMARITAL = 2; if I-1 = 5/6, then DMARITAL = 3; if I-1 = 7/8, then DMARITAL = missing.

**DLFORCE (Labor Force Participation)**

Derived from D-1 and D-2. If D-1B = 2, then DLFORCE = 2. Otherwise, if D-1A = 1 or D-1C = 3, then DLFORCE = 1. Otherwise, if D-1E = 5 or D-1F = 6, then DLFORCE = 3. Otherwise, if D-1D = 4 and D-2 = 1, then DLFORCE = 4. Otherwise, if D-1D = 4 and D-2 = 2, then DLFORCE = 5. Otherwise, if D-1G = 7 or D-1H = 8 or D-1I = 9 or D-1J = 10 or D-1K = 91, then DLFORCE = 5; all else DLFORCE = missing.

**DWFTIME (Length Of Participation In Welfare Programs)**

Derived from I-3G, I-8D, I-6, and I-10. If I-8D = 2, then DWFTIME = 1. Otherwise, if I-3G = 2/7/8 and I-8D = 7/8, then DWFTIME = missing. Otherwise, if I-6 = 1/2/3 or I-10 = 1/2/3, then DWFTIME = 2; if I-6 = 4/5 or I-10 = 4/5, then DWFTIME = 3; if I-6 = 7/8 or I-10 = 7/8, then DWFTIME = missing.

**DVOTE (Voting In The Most Recent Presidential Election)**

Derived from C-8, C-10 and C-11. If C-8 = 2, DVOTE = 0. (Note: Only respondents who were not born in the U.S. were asked this question. Everyone born in the U.S. was assumed to be a citizen.) Otherwise, If C-10 = 3 (voted), DVOTE = 2. Otherwise, if C-10 = 4, then DVOTE = 1; if C-10 = 3, then DVOTE = 2; also, If C-10 = 1 and C-11 = 1, then DVOTE = 2; if C-10 = 2, then DVOTE = 3; if C-10 = 1 and C-11 = 2, then DVOTE = 1. Otherwise, DVOTE = missing.

**DEMPTYPC (Type Of Employer In The Past Three Years: 3 Categories)**

Recoded from D-13. If D-13 = 2, then DEMPTYPC = 1; if D-13 = 3, then DEMPTYPC = 2; if D-13 = 1/4, then DEMPTYPC = 3; if D-13 = missing, then DEMPTYPC = 98; if D-13 = 7/8, then DEMPTYPC = missing.

**DSPUDSTD (How Well Understand Spanish)**

Derived from A-14A and the associated listings of the non-English languages identified in A-6, A-11, A-12 and A-13. Note that respondents were allowed to select multiple non-English languages. A-14A was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DSPUDSTD, all non-English languages were checked and if the non-English language = "SPANISH," then DSPUDSTD = the A-14A that was linked to that SPANISH language. If A-14A = 7/8, then DSPUDSTD was recoded as missing. For respondents who did not speak SPANISH as their non-English language or who were identified as "English only" in A-6, A-11, A-12 and A-13, DSPUDSTD = 98.

**DOTUDSTD (How Well Understand Other Non-English Language)**

Derived from A-14A and the associated listings of the non-English languages identified in A-6, A-11, A-12 and A-13. Note that respondents were allowed to select multiple non-English languages. A-14A was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DOTUDSTD, all non-English languages were checked. If a respondent spoke SPANISH only as non-English language or was identified as "English only" in A-6, A-11, A-12 and A-13, DOTUDSTD = 98. Otherwise, if a respondent spoke only one non-English non-Spanish language, then DOTUDSTD = the A-14A that was linked to that language; if a respondent spoke more than one non-English non-Spanish languages, then DOTUDSTD = the A-14A that was linked to the language rated as best. For example, if a respondent spoke two non-English non-Spanish languages and if one language was rated as "Well" and the other as "Not well," then DOTUDSTD = "Well." If the linked A-14A = 7/8, then DOTUDSTD was recoded as missing.

**DSPSPEAK (How Well Speak Spanish)**

Derived from A-14B and the associated listings of the non-English languages identified in A-6, A-11, A-12 and A-13. Note that respondents were allowed to select multiple non-English languages. A-14B was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DSPSPEAK, all non-English languages were checked and if the non-English language = "SPANISH," then DSPSPEAK = the A-14B that was linked to that SPANISH language. If A-14B = 7/8, then DSPSPEAK was recoded as missing. For respondents who did not speak SPANISH as their non-English language or who were identified as "English only" in A-6, A-11, A-12, and A-13, DSPSPEAK = 98.

**DOTSPEAK (How Well Speak Other Non-English Language)**

Derived from A-14B and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. A-14B was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DOTSPEAK, all non-English languages were checked. If a respondent spoke SPANISH only as non-English language or was identified as "English only" in A-6, A-11, A-12, and A-13, DOTSPEAK = 98. Otherwise, if a respondent spoke only one non-English non-Spanish language, then DOTSPEAK = the A-14B that was linked to that language; if a respondent spoke more than one non-English non-Spanish languages, then DOTSPEAK = the A-14B that was linked to the language rated as best. For example, if a respondent spoke two non-English non-Spanish languages and if one language was rated as "Well" and the other as "Not well," then DOTSPEAK = "Well." If the linked A-14B = 7/8, then DOTSPEAK was recoded as missing.

**DSPREAD (How Well Read Spanish)**

Derived from A-14C and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. A-14C was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DSPREAD, all non-English languages were checked and if the non-English language = "SPANISH," then DSPREAD = the A-14C that was linked to that SPANISH language. If A-14C = 7/8, then DSPREAD was recoded as missing. For respondents who did not speak SPANISH as their non-English language or who were identified as "English only" in A-6, A-11, A-12, and A-13, DSPREAD = 98.

**DOTREAD (How Well Read Other Non-English Language)**

Derived from A-14C and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. A-14C was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DOTREAD, all non-English languages were checked. If a respondent spoke SPANISH only as non-English language or was identified as "English only" in A-6, A-11, A-12, and A-13, DOTREAD = 98. Otherwise, if a respondent spoke only one non-English non-Spanish language, then DOTREAD = the A-14C that was linked to that language; if a respondent spoke more than one non-English non-Spanish languages, then DOTREAD = the A-14C that was linked to the language rated as best. For example, if a respondent spoke two non-English non-Spanish languages and if one language was rated as "Well" and the other as "Not well," then DOTREAD = "Well." If the linked A-14C = 7/8, then DOTREAD was recoded as missing.

**DSPWRITE (How Well Write Spanish)**

Derived from A-14D and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. A-14D was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DSPWRITE, all non-English languages were checked and if the non-English language = "SPANISH," then DSPWRITE = the A-14D that was linked to that SPANISH language. If A-14D = 7/8, then DSPWRITE was recoded as missing. For respondents who did not speak SPANISH as their non-English language or who were identified as "English only" in A-6, A-11, A-12, and A-13, DSPWRITE = 98.

**DOTWRITE (How Well Write Other Non-English Language)**

Derived from A-14D and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. A-14D was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DOTWRITE, all non-English languages were checked. If a respondent spoke SPANISH only as non-English language or was identified as "English only" in A-6, A-11, A-12, and A-13, DOTWRITE = 98. Otherwise, if a respondent spoke only one non-English non-Spanish language, then DOTWRITE = the A-14D that was linked to that language; if a respondent spoke more than one non-English non-Spanish languages, then DOTWRITE = the A-14D that was linked to the language rated as best. For example, if a respondent spoke two non-English non-Spanish languages and if one language was rated as "Well" and the other as "Not well," then DOTWRITE = "Well." If the linked A-14D = 7/8, then DOTWRITE was recoded as missing.

**DSPINFO (How Much Info Got In Spanish About Current Events/Public Affairs/Government)**

Derived from C-2 and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. C-2 was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DSPINFO, all non-English languages were checked and if the non-English language = "SPANISH," then DSPINFO = the C-2 that was linked to that SPANISH language. If C-2 = 7/8, then DSPINFO was recoded as missing. For respondents who did not speak SPANISH as their non-English language or who were identified as "English only" in A-6, DSPINFO = 98.

**DOTINFO (How Much Info Got In Other Non-English Language About Current Events/Public Affairs/Government)**

Derived from C-2 and the associated listings of the non-English languages identified in A-6, A-11, A-12 and A-13. Note that respondents were allowed to select multiple non-English languages. C-2 was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DOTINFO, all non-English languages were checked. If a respondent spoke SPANISH only as non-English language or was identified as “English only” in A-6, DOTINFO = 98. Otherwise, if a respondent spoke only one non-English non-Spanish language, then DOTINFO = the C-2 that was linked to that language; if a respondent spoke more than one non-English non-Spanish languages, then DOTINFO = the C-2 that was linked to the language in which the respondent got the most information. For example, if a respondent spoke two non-English non-Spanish languages and he/she got “most” info in one language and “some” info in the other, then DOTINFO = “Most.” If the linked C-2 = 7/8, then DOTINFO was recoded as missing.

**DCLANGS (Language Usually Speak Now: 3 Categories)**

Derived from A-11. If A-11 = missing, then DCLANGS = 98; if A-11 = “9999999999999999999999999999999999999997 or 9999999999999999999999999999999999999998,” then DCLANGS = missing; if A-11 = “English,” then DCLANGS = 1; if A-11 = “Spanish,” then DCLANGS = 2; otherwise DCLANGS = 3.

**DCLANGSC (Language Usually Speak Now: 2 Categories)**

Derived from A-11. If A-11 = “English,” then DCLANGSC = 1; if A-11 = “9999999999999999999999999999999999999997 or 9999999999999999999999999999999999999998,” then DCLANGSC = missing; otherwise, DCLANGSC = 2.

**DOLANGSB (Other Language Speak Best)**

Derived from A-13. If A-13 = “Spanish,” then DOLANGSB = 1; if A-13 = missing, then DOLANGSB = 98; otherwise, DOLANGSB = 2.

**DDTYPE (Type Of High School Degree: 4 Categories)**

Derived from B-3S. If B-3S = missing, then DDTYPE = 98; if B-3S = 1/2, then DDTYPE = 1; if B-3S = 3, then DDTYPE = 2; if B-3S = 4, then DDTYPE = 3; otherwise, if B-3S = 6/97/98, then DDTYPE = 4.

**DDTYPEC (Type Of High School Degree: 3 Categories)**

Derived from B-3S. If B-3S = missing, then DDTYPEC = 98; if B-3S = 1/2, then DDTYPEC = 1; if B-3S = 4, then DDTYPEC = 2; otherwise, if B-3S = 3/6/99, then DDTYPEC = 3.

**DSGRDHS (What State Did You Live In When You Graduated High School)**

Derived from B-5. If B-5_state = missing, then DSGRDHS = 98; if B-5_state = “99999999999999999997 or 99999999999999999998,” then DSGRDHS = missing; if B-5_state = “state,” then DSGRDHS = 1; if B-5_state not equal to “state,” then DSGRDHS = 2; otherwise, if B-5_state = “outside US,” then DSGRDHS = 2.

**DSGRDCO (What State Did You Live In When You Graduated College)**

Derived from B-8. If B-8_state = missing, then DSGRDCO = 98; if B-8_state = “99999999999999999997 or 99999999999999999998,” then DSGRDCO = missing; if B-8_state = “state,” then DSGRDCO = 1; if B-8_state not equal to “state,” then DSGRDCO = 2; if B-8_STATE = “outside US,” then DSGRDCO = 2.

**DOLSOPT (Other Language Often Spoken Combined: 5 Categories)**

Derived from A-11 and A-12. Note: Languages that all apply were selected from lookup table and reported in 5 initial variables and 5 follow-up variables when “Other, Specify” was reported in the 5 initial variables. All 10 variables were cross-checked and recoded in combination with A-11 as follows: If the respondent speaks English only, then DOLSOPT = 1; If the respondent speaks both English and Spanish, regardless whether he/she also speaks another language(s), DOLSOPT = 2; If the respondent speaks English and another language(s) but not Spanish, DOLSOPT = 3; If the respondent speaks Spanish only or Spanish plus another language(s) but not English, DOLSOPT = 4; If the respondent speaks neither English nor Spanish, DOLSOPT = 5.

**DSPPAPER (How Often Read Newspapers/Magazines In Spanish)**

Derived from E-3A and the associated listings of the non-English languages identified in A-6, A-11, A-12 and A-13. Note that respondents were allowed to select multiple non-English languages. E-3A was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DSPPAPER, all non-English languages were checked and if the non-English language = “SPANISH,” then DSPPAPER = the E-3A that was linked to that SPANISH language. If E-3A = 7/8, then DSPPAPER was recoded as missing. For respondents who did not speak SPANISH as their non-English language or who were identified as “English only” in A-6, DSPPAPER = 98.

**DOTPAPER (How Often Read Newspapers/Magazines In Other Non-English Language)**

Derived from E-3A and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. E-3A was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DOTPAPER, all non-English languages were checked. If a respondent spoke SPANISH only as non-English language or was identified as "English only" in A-6, then DOTPAPER = 98. Otherwise, if a respondent spoke only one non-English non-Spanish language, then DOTPAPER = the E-3A that was linked to that language; if a respondent spoke more than one non-English non-Spanish languages, then DOTPAPER = the E-3A that was linked to the language in which the respondent read the most often. For example, if a respondent spoke two non-English non-Spanish languages and if one language was reported as "Everyday" and the other as "Once a week," then DOTPAPER = "Everyday," If the linked E-3B = 7/8, then DOTPAPER was recoded as missing.

**DSPBOOK (How Often Read Books In Spanish)**

Derived from E-3B and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. E-3B was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DSPBOOK, all non-English languages were checked and if the non-English language = "SPANISH," then DSPBOOK = the E-3B that was linked to that SPANISH language. If E-3B = 7/8, then DSPBOOK was recoded as missing. For respondents who did not speak SPANISH as their non-English language or who were identified as "English only" in A-6, DSPBOOK = 98.

**DOTBOOK (How Often Read Books In Other Non-English Language)**

Derived from E-3B and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. E-3B was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DOTBOOK, all non-English languages were checked. If a respondent spoke SPANISH only as non-English language or was identified as "English only" in A-6, then DOTBOOK = 98. Otherwise, if a respondent spoke only one non-English non-Spanish language, then DOTBOOK = the E-3B that was linked to that language; if a respondent spoke more than one non-English non-Spanish languages, then DOTBOOK = the E-3B that was linked to the language in which the respondent read the most often. For example, if a respondent spoke two non-English non-Spanish languages and if one language was reported as "Everyday" and the other as "Once a week," then DOTBOOK = "Everyday." If the linked E-3B = 7/8, then DOTBOOK was recoded as missing.

**DSPNOTES (How Often Read Notes In Spanish)**

Derived from E-3C and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. E-3C was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DSPNOTES, all non-English languages were checked and if the non-English language = "SPANISH," then DSPNOTES = the E-3C that was linked to that SPANISH language. If E-3C = 7/8, then DSPNOTES was recoded as missing. For respondents who did not speak SPANISH as their non-English language or who were identified as "English only" in A-6, DSPNOTES = 98.

**DOTNOTES (How Often Read Notes In Other Non-English Language)**

Derived from E-3C and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. E-3C was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DOTNOTES, all non-English languages were checked. If a respondent spoke SPANISH only as non-English language or was identified as "English only" in A-6, then DOTNOTES = 98. Otherwise, if a respondent spoke only one non-English non-Spanish language, then DOTNOTES = the E-3C that was linked to that language; if a respondent spoke more than one non-English non-Spanish languages, then DOTNOTES = the E-3C that was linked to the language in which the respondent read the most often. For example, if a respondent spoke two non-English non-Spanish languages and if one language was reported as "Everyday" and the other as "Once a week," then DOTNOTES = "Everyday." If the linked E-3C = 7/8, then DOTNOTES was recoded as missing.

**DRFSSCHC (Reason For Stopping School Before College Degree)**

Derived from B-4. If B-4 = missing, then DRFSSCHC = 98; if B-4 = 1, then DRFSSCHC = 98; if B-4 = 97/98, then DRFSSCHC = missing; if B-4 = 2, then DRFSSCHC = 1; if B-4 = 3, then DRFSSCHC = 2; if B-4 = 4, then DRFSSCHC = 3; if B-4 = 5, then DRFSSCHC = 4; if B-4 = 6/7, then DRFSSCHC = 5; if B-4 = 8/9/10/11, then DRFSSCHC = 6; if B-4 = 13, then DRFSSCHC = 7; if B-4 = 12, then DRFSSCHC = 8; otherwise, if B-4 = 14, then DRFSSCHC = 8.

**DBQ1435 (Frequency Of Volunteering)**

Derived from C-3 and C-4. If C-3 = 2, then DBQ1435 = 5. Otherwise, if C-4 = missing, then DBQ1435 = 98; otherwise, if C-4 = 7/8, then DBQ1435 = missing.

**DBQ1530B (Not Employed Because Retired)**

Derived from D-6A and D-6B. If D-6A = 7/8, then DBQ1530B = missing; if D-6B = missing, then DBQ1530B = 98; otherwise, DBQ1530B = D-6B.

**DBQ1530C (Not Employed Because Taking Care Of Home Or Family)**

Derived from D-6A and D-6C. If D-6A = 7/8, then DBQ1530C = missing; if D-6C = missing, then DBQ1530C = 98; otherwise, DBQ1530C = D-6C.

**DBQ1530D (Not Employed Because Going To School)**

Derived from D-6A and D-6D. If D-6A = 7/8, then DBQ1530D = missing; if D-6D = missing; then DBQ1530D = 98; otherwise, DBQ1530D = D-6D.

**DBQ1530E (Not Employed Because Could Not Find Work)**

Derived from D-6A and D-6E. If D-6A = 7/8, then DBQ1530E = missing; if D-6E = missing, then DBQ1530E = 98; otherwise, DBQ1530E = D-6E.

**DBQ1530F (Not Employed Because Of Other Reason)**

Derived from D-6A and D-6F. If D-6A = 7/8, then DBQ1530F = missing; if D-6F = missing, then DBQ1530F = 98; otherwise, DBQ1530F = D-6F.

**DUSECOMP (Ever Use A Computer)**

Derived from C-1C and E-1. If C-1C = 1/2/3, then DUSECOMP = 1. Otherwise, if E-1 = 7/8/missing, then DUSECOMP = missing. Otherwise, DUSECOMP = E-1.

**DREDCHLD (How Many Times Read To Child: Past Week)**

Derived from H-5 and H-6. If H-5 = 2, then DREDCHLD = 5. Otherwise, if H-6 = 7, then DREDCHLD = missing; if H-6 = missing, then DREDCHLD = 98. Otherwise, DREDCHLD = H-6.

**DBQ2165 (Ever Received Supplemental Security Income)**

Derived from I-8A and I-3B. If I-8A = 1 or I-3B = 1/3, then DBQ2165 = 1; if I-8A = 2, then DBQ2165 = 2. Otherwise, if I-8A = 7/8/missing, then DBQ2165 = missing.

**DBQ2170 (Ever Received Food Stamps)**

Derived from I-8B and I-3D. If I-8B = 1 or I-3D = 1/3, then DBQ2170 = 1; if I-8B = 2, then DBQ2170 = 2. Otherwise, if I-8B = 7/8/missing, then DBQ2170 = missing.

**DBQ2175 (Ever Received Wic Supplemental Nutrition Benefits)**

Derived from I-8C and I-3E. If I-8C = 1 or I-3E = 1/3, then DBQ2175 = 1; if I-8C = 2, then DBQ2175 = 2. Otherwise, if I-8C = 7/8/missing, then DBQ2175 = missing.

**DBQ2180 (Ever Received Tanf Public Assistance Or Public Welfare  Payments)**

Derived from I-8D and I-3G. If I-8D = 1 or I-3G = 1/3, then DBQ2180 = 1; if I-8D = 2, then DBQ2180 = 2; if I-8D = 7/8, then DBQ2180 = missing.

**DWLFLST (Last Received Welfare Payments: 4 Categories)**

Derived from I-3G and I-9. If I-3G = 1/3, then DWLFLST = 1; if I-9 = 1, then DWLFLST = 2; if I-9 = 2, then DWLFLST = 3; if I-9 = 3, then DWLFLST = 4; if I-9 = missing, then DWLFLST = 98; if I-9 = 7/8, then DWLFLST = missing.

**DWLFLSTC (Last Received Welfare Payments: 2 Categories)**

Derived from I-3G and I-9. If I-3G = 1/3, then DWLFLSTC = 1; also, if I-9 = 1/2, then DWLFLSTC = 1; if I-9 = 3, then DWLFLSTC = 2; if I-9 = missing, then DWLFLSTC = 98; if I-9 = 7/8, then DWLFLSTC = missing.

**DBQ2421 (Approximate Personal Income: 8 Categories)**

Derived from K-1. if K-1 = 1/14, then DBQ2421 = 1; if K-1 = 2/3, then DBQ2421 = 2; if K-1 = 4/5, then DBQ2421 = 3; if K-2 = 6, then DBQ2421 = 4; if K-2 = 7 then DBQ2421 = 5; if K-2 = 8, then DBQ2421 = 6; if K-1 = 9/10, then DBQ2421 = 7; if K-1 = 11/12/13, then DBQ2421 = 8. Otherwise, DBQ2421 = missing.

**DBQ2430 (Approximate Household Income: 8 Categories)**

Derived K-2. If K-2 = 1/2/3/15, then DBQ2430 = 1; if K-2 = 4/5, then DBQ2430 = 2; if K-2 = 6, then DBQ2430 = 3; if K-2 = 7, then DBQ2430 = 4; if K-2 = 8, then DBQ2430 = 5; if K-2 = 9/10, then DBQ2430 = 6; if K-2 = 11/12, then DBQ2430 = 7; if K-2 = 13/14, then DBQ2430 = 8. Otherwise, DBQ2430 = missing.

**ICODE_C (Industry)**

Derived from D-9 and D-10. If D-9 = 3, then ICODE_C = 21. Otherwise, use D-10. Responses to D-10 were categorized into the standard 4-digit classifications used by the U.S. Census. These 4-digit classifications were further categorized into the Census standard combinations of 2-digit classifications with one exception: “Military” was combined into “Unknown (Missing).”

**OCODE_C (Occupation)**

Derived from D-9 and D-11. If D-9 = 3, then OCODE_C = 31. Otherwise, use D-11. Responses to D-11 were categorized into the standard 4-digit classifications used by the U.S. Census. These 4-digit classifications were further categorized into the Census standard combinations of 2-digit classifications with two exceptions: “Military” was combined into “Unknown (Missing)” and “Funeral workers” was combined into “Personal Care and Service.”

## 2. Derived variables for the NAAL 2003 prison sample

**DAGEC (Age: 4 Categories)**

Recoded from a continuous age variable derived from date of interview and the BQ response for date of birth, A-3.

**DRACE (Race/Ethnicity)**

Derived from J-1 and J-3. If J-1 = 1, then DRACE = 3. Otherwise, if J-3A = 1 and J-3B-E = missing, then DRACE = 1; if J-3B = 1 and (J-3A and J-3C-E = missing), then DRACE = 2. Otherwise, DRACE = 4.

Note: White, Black and Other categories include no Hispanics.

**DARRIVE (Age Arrived In The U.S.)**

Derived from A-1 and A-1A. If A-1 = “UNITED STATES (50 STATES OR DC),” then DARRIVE = 1. Otherwise, recode continuous responses for A-1A to categorical. If A-1A = missing, then DARRIVE = missing.

**DENGAGE (Age Learned To Speak English)**

Derived from A-8. If A-8 = missing, then DENGAGE = 1. These are people who speak only English. Otherwise, if A-8 = 1/2, then DENGAGE = 2; if A-8 = 3/4/5, then DENGAGE = 3; if A-8 = 95, then DENGAGE = 4; if A-8 = 99, then DENGAGE = missing.

**DCBIRTH (Country Of Birth)**

Derived from A-1. If A-1 = “UNITED STATES (50 STATES OR DC),” then DCBIRTH = 1. Otherwise, DCBIRTH = 2.

**DMCBIRTH (Mother's Country Of Birth)**

Derived from G-1. If G-1 = “UNITED STATES (50 STATES OR DC),” then DMCBIRTH = 1; if G-1 = “99999999999999999999999999999999999999999999999999,” then DMCBIRTH = missing. Otherwise, DMCBIRTH = 2.

**DFCBIRTH (Father's Country Of Birth)**

Derived from G-3. If G-3 = “UNITED STATES (50 STATES OR DC),” then DFCBIRTH = 1; if G-3 = “99999999999999999999999999999999999999999999999999,” then DFCBIRTH = missing. Otherwise, DFCBIRTH = 2.

**D1STLANC (Language Spoken Before School: 3 Categories)**

Derived from A-6. Note: Languages that all apply were selected from lookup table and reported in 5 initial variables and 5 follow-up variables when "Other, Specify" was reported in the 5 initial variables. All 10 variables were cross-checked and recoded as follows: If the respondent speaks English only, then D1STLANC = 1; if the respondent speaks English and another language(s), D1STLANC = 2; if the respondent does not speak English, D1STLANC = 3.

**DHMLANGC (Language Spoken At Home When Growing Up: 3 Categories)**

Derived from A-5. Note: Languages that all apply were selected from lookup table and reported in 5 initial variables and 5 follow-up variables for "Other, Specify" in the 5 initial variables. All 10 variables were cross-checked and recoded as follows: If the respondent speaks English only, then DHMLANGC = 1; if the respondent speaks English and another language(s), DHMLANGC = 2; if the respondent does not speak English, DHMLANGC = 3.

**DLANGRWC (Language First Learned To Read And Write: 2 Categories)**

Derived from A-7. If A-7 = "English," then DLANGRWC = 1; if A-7 = missing, then DLANGRWC = 1 (these are respondents who reported "English only" to A-5 and A-6). Otherwise, DLANGRWC = 2.

**DEDBFUS (Education Before Coming To The U.S.)**

Derived from A-4. If A-4 = 1/2, then DEDBFUS = 1; if A-4 = 3, then DEDBFUS = 2; if A-4 = 4/5/6, then DEDBFUS = 3; if A-4 = missing, then DEDBFUS = 98.

**DEDATTNC (Educational Attainment: 6 Categories)**

Derived from B-1, B-2 and B-5. Note: Education was recorded before (B-1) and while (B-5) in prison. If B-2 = 1 and B-5 >B-1, then use B-5 for DEDATTNC. If B-2 ~ = 1, then use B-1 for DEDATTNC. B-5/B-1 was recoded into corresponding categories of DEDATTNC.

**DHSAGE (Age Upon Graduation From High School Or Obtaining A Ged)**

Derived from A-3, B-1 and B-6. DHSAGE was assigned only to respondents who completed high school/obtained a GED: B-1 >3. If B-1 < = 3, then DHSAGE = 98.

For respondents with B-1 >3:

- Date of graduation was set as June 30 in the year indicated in B-6.
- Respondent's date of birth (A-3) was subtracted from date of graduation.

- This yielded an age expressed in years and months (e.g. 18.5 = 18 years, 6 months). Ages not expressed in whole numbers were rounded down.

- Ages were then recoded into reporting categories 1 and 2. The lower boundary for category 1 was 12 years of age.

**DMEDC (Mother's Educational Attainment: 5 Categories)**

Recoded from G-2 as follows: if G-2 = 1/2, then DMEDC = 1; if G-2 = 3, then DMEDC = 2; if G-2 = 4, then DMEDC = 3; if G-2 = 5/6/7/8, then DMEDC = 4; if G-2 = 9/10/11, then DMEDC = 5; if G-2 = 99, then DMEDC = missing.

**DFEDC (Father's Educational Attainment: 5 Categories)**

Recoded from G-4 as follows: if G-4 = 1/2, then DFEDC = 1; if G-4 = 3, then DFEDC = 2; if G-4 = 4, then DFEDC = 3; if G-4 = 5/6/7/8, then DFEDC = 4; if G-4 = 9/10/11, then DFEDC = 5; if G-4 = 99, then DFEDC = missing.

**DMARITAL (Marital Status)**

Derived from H-1. If H-1 = 1, then DMARITAL = 1; if H-1 = 2/3, then DMARITAL = 2; if H-1 = 4/5, then DMARITAL = 3; if H-1 = 7/8, then DMARITAL = missing.

**DWFTIME (Length Of Participation In Welfare Programs)**

Derived from H-2D and H-4. If H-2D = 2, then DWFTIME = 1. Otherwise, if H-4 = 1/2/3, then DWFTIME = 2; if H-4 = 4/5, then DWFTIME = 3. Otherwise, DWFTIME = missing.

**DVOTE (Voting In The Most Recent Presidential Election)**

Derived from E-9, E-11 and E-12. If E-9 = 2, DVOTE = 0. (Note: Only respondents who were not born in the U.S. were asked this question. Everyone born in the U.S. was assumed to be a citizen.) Otherwise, if E-11 = 3 (voted), DVOTE = 2.

Otherwise, if both E-11 and E-12 = missing, then DVOTE = 98. These were prisoners who were in prison for current offense in November 2000 and skipped E-11 and E-12.

Otherwise, if E-11 = 4, then DVOTE = 1; if E-11 = 3, then DVOTE = 2; also, If E-11 = 1 and E-12 = 1, then DVOTE = 2; if E-11 = 2, then DVOTE = 3; if E-11 = 1 and E-12 = 2, then DVOTE = 1. Otherwise, DVOTE = missing.

**DEMPTYPC (Type Of Employer In The Past Three Years: 3 Categories)**

Derived from D-12. If D-12 = 2, then DEMPTYPC = 1; if D-12 = 3, then DEMPTYPC = 2; if D-12 = 1/4, then DEMPTYPC = 3; if D-12 = missing, then DEMPTYPC = 98; if D-12 = 8, then DEMPTYPC = missing.

**DSPUDSTD (How Well Understand Spanish)**

Derived from A-14A and the associated listings of the non-English languages identified in A-6, A-11, A-12 and A-13. Note that respondents were allowed to select multiple non-English languages. A-14A was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DSPUDSTD, all non-English languages were checked and if the non-English language = "SPANISH," then DSPUDSTD = the A-14A that was linked to that SPANISH language. For respondents who did not speak SPANISH as their non-English language or who were identified as "English only" in A-6, A-11, A-12, and A-13, DSPUDSTD = 98.

**DOTUDSTD (How Well Understand Other Non-English Language)**

Derived from A-14A and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. A-14A was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DOTUDSTD, all non-English languages were checked. If a respondent spoke SPANISH only as non-English language or was identified as "English only" in A-6, A-11, A-12, and A-13, DOTUDSTD = 98. Otherwise, if a respondent spoke only one non-English non-Spanish language, then DOTUDSTD = the A-14A that was linked to that language; if a respondent spoke more than one non-English non-Spanish languages, then DOTUDSTD = the A-14A that was linked to the language rated as best. For example, if a respondent spoke two non-English non-Spanish languages and if one language was rated as "Well" and the other as "Not well," then DOTUDSTD = "Well."

**DSPSPEAK (How Well Speak Spanish)**

Derived from A-14B and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. A-14B was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DSPSPEAK, all non-English languages were checked and if the non-English language = "SPANISH," then DSPSPEAK = the A-14B that was linked to that SPANISH language. For respondents who did not speak SPANISH as their non-English language or who were identified as "English only" in A-6, A-11, A-12, and A-13, DSPSPEAK = 98.

**DOTSPEAK (How Well Speak Other Non-English Language)**

Derived from A-14B and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. A-14B was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DOTSPEAK, all non-English languages were checked. If a respondent spoke SPANISH only as non-English language or was identified as "English only" in A-6, A-11, A-12, and A-13, DOTSPEAK = 98. Otherwise, if a respondent spoke only one non-English non-Spanish language, then DOTSPEAK = the A-14B that was linked to that language; if a respondent spoke more than one non-English non-Spanish languages, then DOTSPEAK = the A-14B that was linked to the language rated as best. For example, if a respondent spoke two non-English non-Spanish languages and if one language was rated as "Well" and the other as "Not well," then DOTSPEAK = "Well."

**DSPREAD (How Well Read Spanish)**

Derived from A-14C and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. A-14C was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DSPREAD, all non-English languages were checked and if the non-English language = "SPANISH," then DSPREAD = the A-14C that was linked to that SPANISH language. For respondents who did not speak SPANISH as their non-English language or who were identified as "English only" in A-6, A-11, A-12, and A-13, DSPREAD = 98.

**DOTREAD (How Well Read Other Non-English Language)**

Derived from A-14C and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. A-14C was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DOTREAD, all non-English languages were checked. If a respondent spoke SPANISH only as non-English language or was identified as "English only" in A-6, A-11, A-12, and A-13, DOTREAD = 98. Otherwise, if a respondent spoke only one non-English non-Spanish language, then DOTREAD = the A-14C that was linked to that language; if a respondent spoke more than one non-English non-Spanish languages, then DOTREAD = the A-14C that was linked to the language rated as best. For example, if a respondent spoke two non-English non-Spanish languages and if one language was rated as "Well" and the other as "Not well," then DOTREAD = "Well."

**DSPWRITE (How Well Write Spanish)**

Derived from A-14D and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. A-14D was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DSPWRITE, all non-English languages were checked and if the non-English language = "SPANISH," then DSPWRITE = the A-14D that was linked to that SPANISH language. For respondents who did not speak SPANISH as their non-English language or who were identified as "English only" in A-6, A-11, A-12, and A-13, DSPWRITE = 98.

**DOTWRITE (How Well Write Other Non-English Language)**

Derived from A-14D and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. A-14D was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DOTWRITE, all non-English languages were checked. If a respondent spoke SPANISH only as non-English language or was identified as "English only" in A-6, A-11, A-12, and A-13, DOTWRITE = 98. Otherwise, if a respondent spoke only one non-English non-Spanish language, then DOTWRITE = the A-14D that was linked to that language; if a respondent spoke more than one non-English non-Spanish languages, then DOTWRITE = the A-14D that was linked to the language rated as best. For example, if a respondent spoke two non-English non-Spanish languages and if one language was rated as "Well" and the other as "Not well," then DOTWRITE = "Well."

**DSPINFO (How Much Info Got In Spanish About Current Events/Public Affairs/Government)**

Derived from E-2 and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. E-2 was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DSPINFO, all non-English languages were checked and if the non-English language = "SPANISH," then DSPINFO = the E-2 that was linked to that SPANISH language. For respondents who did not speak SPANISH as their non-English language or who were identified as "English only" in A-6, DSPINFO = 98.

**DOTINFO (How Much Info Got In Other Non-English Language About Current Events/Public Affairs/Government)**

Derived from E-2 and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. E-2 was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DOTINFO, all non-English languages were checked. If a respondent spoke SPANISH only as non-English language or was identified as "English only" in A-6, DOTINFO = 98. Otherwise, if a respondent spoke only one non-English non-Spanish language, then DOTINFO = the E-2 that was linked to that language; if a respondent spoke more than one non-English non-Spanish languages, then DOTINFO = the E-2 that was linked to the language in which the respondent got the most information. For example, if a respondent spoke two non-English non-Spanish languages and he/she got "most" info in one language and "some" info in the other, then DOTINFO = "Most,"

**DCLANGSC (Language Usually Speak Now: 2 Categories)**

Derived from A-11. If A-11 = "English," then DCLANGSC = 1; otherwise, DCLANGSC = 2.

**DOLANGSB (Other Language Speak Best)**

Derived from A-13. If A-13 = "Spanish," then DOLANGSB = 1; if A-13 = missing, then DOLANGSB = 98; otherwise, DOLANGSB = 2.

**DDTYPEC (Type Of High School Degree: 3 Categories)**

Derived from B-3S. If B-3S = missing, then DDTYPEC = 98; if B-3S = 1/2, then DDTYPEC = 1; if B-3S = 4, then DDTYPEC = 2; otherwise, if B-3S = 3/6/99, then DDTYPEC = 3.

**DSPPAPER (How Often Read Newspapers/Magazines In Spanish)**

Derived from F-2A and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. F-2A was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DSPPAPER, all non-English languages were checked and if the non-English language = "SPANISH," then DSPPAPER = the F-2A that was linked to the SPANISH language. For respondents who did not speak SPANISH as their non-English language or who were identified as "English only" in A-6, DSPPAPER = 98.

**DOTPAPER (How Often Read Newspapers/Magazines In Other Non-English Language)**

Derived from F-2A and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. F-2A was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DOTPAPER, all non-English languages were checked. If a respondent spoke SPANISH only as non-English language or was identified as "English only" in A-6, then DOTPAPER = 98. Otherwise, if a respondent spoke only one non-English non-Spanish language, then DOTPAPER = the F-2A that was linked to that language; if a respondent spoke more than one non-English non-Spanish languages, then DOTPAPER = the F-2A that was linked to the language in which the respondent read the most often. For example, if a respondent spoke two non-English non-Spanish languages and if one language was reported as "Everyday" and the other as "Once a week," then DOTPAPER = "Everyday."

**DSPBOOK (How Often Read Books In Spanish)**

Derived from F-2B and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. F-2B was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DSPBOOK, all non-English languages were checked and if the non-English language = "SPANISH," then DSPBOOK = the F-2B that was linked to that SPANISH language. For respondents who did not speak SPANISH as their non-English language or who were identified as "English only" in A-6, DSPBOOK = 98.

**DOTBOOK (How Often Read Books In Other Non-English Language)**

Derived from F-2B and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. F-2B was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DOTBOOK, all non-English languages were checked. If a respondent spoke SPANISH only as non-English language or was identified as "English only" in A-6, then DOTBOOK = 98. Otherwise, if a respondent spoke only one non-English non-Spanish language, then DOTBOOK = the F-2B that was linked to that language; if a respondent spoke more than one non-English non-Spanish languages, then DOTBOOK = the F-2B that was linked to the language in which the respondent read the most often. For example, if a respondent spoke two non-English non-Spanish languages and if one language was reported as "Everyday" and the other as "Once a week," then DOTBOOK = "Everyday."

**DSPNOTES (How Often Read Notes In Spanish)**

Derived from F-2C and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. F-2C was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DSPNOTES, all non-English languages were checked and if the non-English language = "SPANISH," then DSPNOTES = the F-2C that was linked to that SPANISH language. For respondents who did not speak SPANISH as their non-English language or who were identified as "English only" in A-6, DSPNOTES = 98.

**DOTNOTES (How Often Read Notes In Other Non-English Language)**

Derived from F-2C and the associated listings of the non-English languages identified in A-6, A-11, A-12, and A-13. Note that respondents were allowed to select multiple non-English languages. F-2C was repeated for each of the non-English languages and therefore comprised an array of responses linked to each of the non-English languages. To create DOTNOTES, all non-English languages were checked. If a respondent spoke SPANISH only as non-English language or was identified as "English only" in A-6, then DOTNOTES = 98. Otherwise, if a respondent spoke only one non-English non-Spanish language, then DOTNOTES = the F-2C that was linked to that language; if a respondent spoke more than one non-English non-Spanish languages, then DOTNOTES = the F-2C that was linked to the language in which the respondent read the most often. For example, if a respondent spoke two non-English non-Spanish languages and if one language was reported as "Everyday" and the other as "Once a week," then DOTNOTES = "Everyday,"

**DRFSSCHC (Reason For Stopping School Before College Degree)**

Derived from B-4. If B-4 = missing, then DRFSSCHC = 98; if B-4 = 1, then DRFSSCHC = 98; if B-4 = 2, then DRFSSCHC = 1; if B-4 = 3, then DRFSSCHC = 2; if B-4 = 4, then DRFSSCHC = 3; if B-4 = 5, then DRFSSCHC = 4; if B-4 = 6/7, then DRFSSCHC = 5; if B-4 = 8/9/10/11, then DRFSSCHC = 6; if B-4 = 13, then DRFSSCHC = 7; if B-4 = 12, then DRFSSCHC = 8; otherwise, if B-4 = 14, then DRFSSCHC = 8.

**DWLFLSTC (Last Received Welfare Payments: 2 Categories)**

Derived from H-3. If H-3 = 1/2/3, then DWLFLSTC = 1; if H-3 = 4, then DWLFLSTC = 2; if H-3 = missing, then DWLFLSTC = 98.

**DPGED (Ged Earned While In Prison)**

Derived from B-1, B-3, B-5, and B-12. If B-1 > = 4 and (B-5 > = 4 or B-5 = missing or B-5 = 9), then DPGED = 1; if B-1 < 4 and B-5 > = 4, or B-1 > = 4 and B-5 = 3, then DPGED = 2. Otherwise, if B-1 <4 and (B-3 = 1 or B-12 = 4), then DPGED = 3. Otherwise, DPGED = 4. Note: There were 5 cases where B-1 > = 4 and (B-5 = 1/2). All 5 cases indicated that they were not currently enrolled in academic (B-3) or basic skills (B-12) classes. For all 5 cases, DPGED = 4.

**DPBQ1209 (Currently Enrolled In Academic Classes)**

Derived from B-3. If B-3 = 1, then DPBQ1209 = 1; if B-3 = 2, then DPBQ1209 = 2; if B-3 = missing, then DPBQ1209 = 2. Note: Respondents who answered "No" to B-2 skipped B-3 and were coded as missing in B-3.

**DPBLOCAT (Where Took Basic Skills Class)**

Derived from B-10 and B-11. If B-10 = 2, then DPBLOCAT = 4. If B-10 = 1, then do as follows: if B-11A = 1, then DPBLOCAT = 1; if B-11B = 2, then DPBLOCAT = 2; if B-11C = 3, then DPBLOCAT = 3. Note: Respondents were allowed to select more than 1 answer for B-11. In this situation, because education completed in prison was of most interest, the answer that was closest to the respondent's most recent incarceration was taken for DPBLOCAT. For example: if B-11A = 1 and B-11C = 3, then DPBLOCAT = 1; if B-11A = 1 and B-11B = 2, then DPBLOCAT = 1; if B-11B = 2 and B-11C = 3, then DPBLOCAT = 2.

**DPITCERT (Completion Of It Certification While In Prison)**

Derived from B-16. If B-16 = 1, then DPITCERT = 1; if B-16 = 2, then DPITCERT = 2; if B-16 = missing, then DPITCERT = 3.

**DPOTCERT (Completion Of Other Job Certification While In Prison)**

Derived from B-19. If B-19 = 1, then DPOTCERT = 1; if B-19 = 2, then DPOTCERT = 2; if B-19 = missing, then DPOTCERT = 3.

**DPVOC (Length Of Time In Prison Vocational Training Program)**

Derived from C-2 and C-4. If C-2 = 2, then DPVOC = 0. Otherwise, if C-4 = 1 or 2, then DPVOC = 1; if C-4 = 3, then DPVOC = 2; if C-4 = 9 or missing, then DPVOC = missing.

**DPCLSHR (How Many Hours Spent In Prison Classes Last Week)**

Derived from C-7. If C-7 = 0, then DPCLSHR = 0; if 1< = C-7 < = 19, then DPCLSHR = 1; if 20 < = C-7< = 49, then DPCLSHR = 2; if C-7 > = 50, then DPCLSHR = 3; if C-7 = 98/99, then DPCLSHR = missing.

**DOFFENS1 (Offense 1 For Which Inmate Is In Prison)**

Derived from C-8A. The text responses of inmate's offenses to C-8A were recoded into 1 of 5 major offense classifications.

**DOFFENS2 (Offense 2 For Which Inmate Is In Prison)**

Derived from C-8B. The text responses of inmate's offenses to C-8B were recoded into 1 of 5 major offense classifications. If C-8B = missing, DOFFENS2 = 98.

**DOFFENS3 (Offense 3 For Which Inmate Is In Prison)**

Derived from C-8C. The text responses of inmate's offenses to C-8C were recoded into 1 of 5 major offense classifications. If C-8C = missing, DOFFENS3 = 98.

**DOFFENSE (Type Of Offense For Which Inmate Received Longest Sentence)**

Derived from C-8 and C-9. DOFFENSE captures the offense for which the inmate received the longest sentence. Each of C-8A through C-8E was first recoded into 5 major offense classifications. If only 1 offense was listed in C-8, then DOFFENSE = C-8A (recoded). If multiple offenses were listed in C-8, then C-9 was used to determine the offense for which the inmate received the longest sentence. If C-9 = 1, then DOFFENSE = C-8A (recoded); if C-9 = 2, then DOFFENSE = C-8B (recoded); if C-9 = 3, then DOFFENSE = C-8C (recoded); if C-9 = 4, then DOFFENSE = C-8D (recoded); if C-9 = 5, then DOFFENSE = C-8E (recoded). If C-9 = 95/98/99, then DOFFENSE = missing.

**DCRIMHIS (Previous Criminal History)**

Derived from C-11 and C-12. If C-11 = 1 and C-12 = 1, then DCRIMHIS = 4; if C-11 = 1 and C-12 = 2, then DCRIMHIS = 3; if C-11 = 2 and C-12 = 1, then DCRIMHIS = 2; if C-11 = 2/8 and C-12 = 2/8, then DCRIMHIS = 1.

**DRELEASE (Expected Date Of Release)**

Derived from C-13, C-14, C-15, and C-16. There were two steps for calculating this variable: 1) determining the respondent's expected year of release and 2) subtracting the year the assessment was administered (2004) from the respondent's expected year of release.

Calculating expected year of release:

1. If C-13 = 1, then calculate expected year of release from the year in C-14. If the year in C-14 = 9998, then DRELEASE = missing.

2. If C-13 = 2 or C-15 = 1, then calculate expected year of release from the year in C-16. If the year in C-16 = 9998 or missing, then DRELEASE = missing.

3. If C-13 = 2 and C-15 = 2, then DRELEASE = 2. These are prisoners who did not expect to be released.

4. If C-13 = missing and C-15 = missing, then DRELEASE = missing. These are prisoners who had not been sentenced when the BQ was administered.

For respondents not classified for DRELEASE in steps 1—4, calculate DRELEASE by subtracting 2004 from expected year of release. Recode the difference to either DRELEASE = 1 or DRELEASE = 2.

**DLENGTHC (Length Of Sentence: Collapsed)**

Derived from C-10, C-13, C-14, C-15, and C-16. DLENGTHC was recoded from a detailed derived variable, DLENGTHD. DLENGTHD was derived as follows: If C-13 = missing and C-15 = missing, then DLENGTHD = 999999999 (Not sentenced yet). Otherwise, there were two steps for calculating this variable: 1) determining the respondent's expected month and year of release and 2) subtracting the respondent's date of admission to prison from the expected date of release and recoding the date into months.

Note: If the month in C-10/C-14/C-16 equals 98 or missing while the year in C-10/C-14/C-16 is non-missing (i.e., unequal to missing or 9998), set June as the month for the month variables.

Calculating expected month and year of release:

If C-13 = 1, then calculate expected month and year of release from C-14.

If C-13 = 2 or C-15 = 1, then calculate expected month and year of release from C-16.

Once expected month and year of release were calculated:

Use C-10 to get the month and year of admission. Subtract the date of admission from the respondent's expected date of release and recode the date into months.

Note: If the year in C-14/C-16 = missing/9998 or the year in C-10 = missing/9998, then DLENGTHD = missing. If expected date of release is earlier than date of admission, then DLENGTHD = missing.

If C-13 = 2 and C-15 = 2, then DLENGTHD = 99999999997 (Do not expect to be released).

If C-13 = 8/9 and C-15 = 8/9, then DLENGTHD = missing.

DLENGTHC was then recoded from DLENGTHD as follows: If DLENGTHD = 99999999997, then DLENGTHC = 3; if DLENGTHD = 99999999999, then DLENGTHC = missing. Otherwise, if 0< = DLENGTHD< = 60 then DLENGTHC = 1; if 61< = DLENGTHD< = 120 then DLENGTHC = 2; if DLENGTHD> = 121 then DLENGTHC = 3.

**DPJOBHR (How Many Hours Worked At Current Job In Prison In The Last Week)**

Derived from D-4. If D-4 = 0, then DPJOBHR = 1; if 1 < = D-4 < = 9, then DPJOBHR = 2; if 10 < = D-4< = 19, then DPJOBHR = 3; if 20 < = D-4< = 29, then DPJOBHR = 4; if D-4 > = 30, then DPJOBHR = 5; if D-4 = missing, then DPJOBHR = 6; if D-4 = 98/99, then DPJOBHR = missing.

**DPBQ1615 (Worked Full Time Or Not In The Past Three Years While Not In Prison)**

Derived from D-11. If D-11 = 1, then DPBQ1615 = 1; if D-11 = 2/3, then DPBQ1615 = 2; if D-11 = missing, then DPBQ1615 = 98.

**DPLIBACS (Length Of Time To Access Prison Library)**

Derived from E-4, E-5, E-6, and E-7. If E-4 = 1 then DPLIBACS = 1. Otherwise, if E-5 = 2 or E-6 = 2, then DPLIBACS = 5. Otherwise, use E-7 to code DPLIBACS as follows: If E-7 = 8/9, then DPLIBACS = missing; if E-7 = missing, then DPLIBACS = 98; otherwise, DPLIBACS = E-7.

# Appendix B

# INSTRUCTIONS ON USING AM SOFTWARE TO ANALYZE THE 2003 NAAL DATA

# Appendix B: Instructions on Using AM Software to Analyze the 2003 NAAL Data

As indicated in Chapter 16, NAAL used a complex assessment design that allowed maximum coverage of the broad domain of literacy, while minimizing the time burden on any one participant (see Chapter 14). Under this design, participants did not receive enough literacy tasks to provide reliable information about individual performance and therefore traditional test scores for individual participants were not appropriate to use for estimates of population statistics. To obtain consistent estimates of population statistics from data collected under such design, statistical procedures based on the method of marginal maximum likelihood (MML) need to be used. The usual statistical software packages such as SAS, SPSS or STATA can not implement MML procedures and therefore special analysis tools, such as the AM software, are needed. This appendix describes how to use AM software to analyze the 2003 NAAL data.

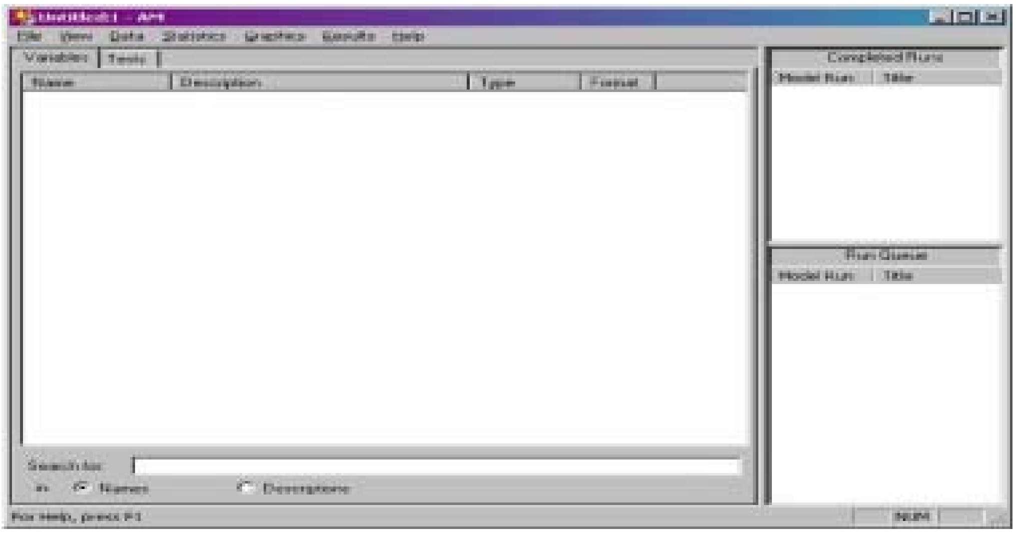
## Starting AM

To open the program, simply double-click the AM software icon. The AM program has a main menu bar and three windows:

- The Variables List appears on the left-hand side of the screen. Use the Variables List to view and manage the variables in your data file. You can use the AM View menu to select the types of variables to be displayed in the Variables List.
- The Run Queue appears in the bottom-right quadrant. Statistical models that you are preparing appear here until it is their turn to run.
- The Completed Runs Queue holds statistical runs once they are complete. Most statistical models allow you to conduct post-hoc analysis by right-clicking icons in the completed run queue.

The program looks like figure B-1 when it first opens:

**Figure B-1. AM program starting window: 2003**




SOURCE: Cohen, J., Jiang, T. Gaidurguis, A., & Hollender, D. (2000). AM (Software for marginal maximum likelihood and other statistical analyses of data from complex samples), Washington: American Institutes for Research. See website at http://am.air.org.

### Getting Data Into The Program

The first step to using AM is getting data into the program. There are two ways to getting data into AM, depending on the type of the data files.

1. If your data file is an AM file with an extension of .am in the file name, then:
    a. Under the AM File menu, select “Open Database”.
    b. Locate the 2003 NAAL AM data file and click “Open”.
2. If your data file is an SPSS file, then:
    a. Under the AM File menu, select “Import” and then in the pull-down menu, select “SPSS.sav File”.
    b. Locate the SPSS data file and click “Open”.
    c. In addition to importing the SPSS data file, you will also need to upload the AM dictionary file that contains information on item parameters (see 16.1.3). To upload the AM dictionary file:
        i. Under the AM File menu, select “Update Metadata”.
        ii. Locate the dictionary file and click “Open”.

3. If your data file is in other file types such as SAS or STATA, then:

    a. Under the AM File menu, select "Import" and then in the pull-down menu, select "General Import".

    b. Select the appropriate file type using the pull down menu for "Files of type". Locate the data file and click "Open".

    c. Repeat 2.c above.

**Analyzing 2003 NAAL Data In AM**

Literacy scores from the NAAL are reported in two formats: 1) as means and 2) as the percentage of adults within each of four literacy levels. Two procedures in AM generate mean literacy scores and the percentage of adults within the literacy levels.

For both types of analyses, you will need to select an **independent variable(s)** and a **dependent variable.**

***Independent variables:*** These variables describe attributes or characteristics of people, such as gender, race/ethnicity, or educational attainment. Use the 2003 NAAL Electronic Codebook to identify independent variables of interest.

***Dependent variable:*** The dependent variable is the literacy scale for which you want results. Select prose, document, quantitative, or health.

**Assigning Sampling Design Variables**

If you use the 2003 NAAL AM data file, AM will automatically account for the NAAL sampling design and includes strata, cluster, and weight variables necessary to obtain more precise estimates for analyses. If you use SPSS or other types of the 2003 NAAL data file, you need to specify the strata, cluster, and weight variables. The strata, cluster, and weight variables are: VARSTRAT, VARUNIT and WEIGHT, respectively. To specify these sampling design variables:

1. Find one of the above variables from the "Variables" window on the left-hand side of the screen and right click the variable.

2. Select "Edit Metadata".

3. In the pop-up window, select the appropriate "Design Role" for the variable.

These sampling design variables will be automatically included in any analyses that follow.

**Calculating Mean Literacy Scores**

1. Once you have opened the 2003 NAAL data file in AM, go to the “Statistics” menu, select “MML Procedures for Test Data,” and then select “MML Means (Separate Variances).”

2. Select the independent variable you would like to analyze (e.g., race/ethnicity, educational attainment, gender). You can enter independent variables in two ways:

   a. Type the variable name in the small box beneath the Independent Variables box. Click “Enter” after you type the name to move it to the list of Independent Variables.

   b. Drag the variable from the Variables List window to the Independent Variables box.

3. Select the literacy scale for which you would like results (prose, document, quantitative, or health). You can select the scale in two ways:

   a. Type the scale name in the small box beneath the Dependent Variables box. Click “Enter” after you type the name to move it to the list of Dependent Variables.

   b. Click the Tests window on the left side of the screen. Expand the “NAAL2003” icon and drag the scale of interest to the Dependent Variables box.

4. Select the format in which you would like the results to appear. The default setting is “Web browser.” If you would like the results to be outputted as a text file or a spreadsheet, select “Plain Text” or “Spreadsheet” output.

5. Click “OK.” AM will execute the command and display the results in the output format you specified.

**Figure B-2. Example of calculating mean prose literacy scores for men and women: 2003**

The following example generates mean prose literacy scores for men and women. The results match the literacy estimates presented in Figure 4 of the first NAAL report, *A First Look at the Literacy of America's Adults in the 21st Century.* [1]

1. From the "Statistics" menu, select "MML Procedures for Test Data."

2. Select "MML Means (Separate variances)."

3. As noted in the 2003 NAAL codebook, the variable capturing gender is *DSEX.* To estimate the literacy of men and women, drag the variable *DSEX* from the Variables List window to the Independent Variables box (you can also type *DSEX* in the small box beneath the Independent Variables box and click enter).

4. To estimate prose literacy scores, drag the Prose test from the Tests window to the Dependent Variables box (you can also type "Prose" in the small box beneath the Dependent Variables box and click enter).

5. For the *First Look Report,* some of the advanced specifications were changed to get more precise estimates. For example, the number of iterations was increased to 1,000 and the convergence criterion was decreased to .0000001. You can change these settings by clicking the "Advanced Parameters" tab. However, for general secondary analyses, the default settings should suffice.

6. Click "OK" and the mean prose literacy scores for men and women will appear in the output format you selected.

To estimate means for document, quantitative and health literacy for men and women, replace Prose in step 4 with the Document, Quantitative, or Health test variables. To estimate means for a different population group (e.g., country of birth), replace *DSEX* in the Independent Variables box with the variable of interest.

SOURCE: U.S. Department of Education. National Center for Education Statistics. Technical Report and Data File User's Manual For the 2003 National Assessment of Adult Literacy.

### Calculating the Percentage of Adults Within Literacy Levels

The 2003 NAAL also reports results by using four literacy levels: *Below Basic, Basic, Intermediate,* and *Proficient.* As shown in Figure B-3, each of the literacy scales (prose, document, quantitative, and health) has unique cutpoints for the literacy levels.

---

[1] Kutner, M., Greenberg, E., and Baer, J. (2005). *A First Look at the Literacy of America's Adults in the 21st Century* (NCES 2006-470). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

**Figure B-3. NAAL cutpoints, by literacy scale and literacy level: 2003**

| Prose | | Document | | Quantitative | | Health | |
|---|---|---|---|---|---|---|---|
| Below Basic: | 0–209 | Below Basic: | 0–204 | Below Basic: | 0–234 | Below Basic: | 0–184 |
| Basic: | 210–264 | Basic: | 205–249 | Basic: | 235–289 | Basic: | 185–225 |
| Intermediate: | 265–339 | Intermediate: | 250–334 | Intermediate: | 290–349 | Intermediate: | 226–309 |
| Proficient: | 340–500 | Proficient: | 335–500 | Proficient: | 350–500 | Proficient: | 310–500 |

SOURCE: Kutner, M., Greenberg, E., Jin,Y., Boyle, B., Hsu,Y., Paulsen, C. (2006). *Literacy in Everyday Life: Results From the 2003 National Assessment of Adult Literacy* (NCES 2006–477). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

1. Once you have opened the 2003 NAAL data file in AM, go to the "Statistics" menu, select "MML Procedures for Test Data," and then select "NALS Table."[2]

2. Select the independent variable you would like to analyze (e.g., race/ethnicity, educational attainment, gender). You can enter independent variables in two ways:

   a. Type the variable name in the small box beneath the Independent Variables box. Click "Enter" after you type the name to move it to the list of Independent Variables.

   b. Drag the variable from the Variables List window to the Independent Variables box.

3. Select the literacy scale for which you would like results (prose, document, quantitative or health). You can select the scale in two ways:

   a. Type the scale name in the small box beneath the Dependent Variables box. Click "Enter" after you type the name to move it to the list of Dependent Variables.

   b. Click the Tests window on the left side of the screen. Expand the "NAAL2003" icon and the scale of interest to the Dependent Variables box.

4. Enter the appropriate cut scores for the literacy scale you have selected. The cut scores for the scales are as follows:

[2] The procedure is called "NALS Table" after the 1992 National Adult Literacy Survey (NALS), the precursor to the 2003 National Assessment of Adult Literacy (NAAL).

**Figure B-4. NAAL cut scores, by literacy scale: 2003**

| Prose | Document | Quantitative | Health |
|---|---|---|---|
| Cut 1: 210<br>Cut 2: 265<br>Cut 3: 340 | Cut 1: 205<br>Cut 2: 250<br>Cut 3: 335 | Cut 1: 235<br>Cut 2: 290<br>Cut 3: 350 | Cut 1: 185<br>Cut 2: 226<br>Cut 3: 310 |

SOURCE: Kutner,M., Greenberg, E., Jin,Y., Boyle, B., Hsu,Y., Paulsen, C. (2006). Literacy in Everyday Life: Results From the 2003 National Assessment of Adult Literacy (NCES 2006–477). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

a. To enter cut scores, double-click "Cut 1" in the Cut Scores box. Type the first cut score for the appropriate scale in the box. For example, if you select prose as your dependent variable, type "210" for Cut 1. Click "OK" after entering the cut score.

b. Enter the remaining two cut scores for the scale you have selected, following the same steps used to set the first cut score. Double-click "Cut 2," enter the appropriate score (e.g., 265 for prose), and click "OK." Enter the last cut score by double-clicking "Cut 3," enter the cut score (e.g., 340 for prose), and then click "OK."

5. Select the format in which you would like the results to appear. The default setting is "Web browser." You can also select "Spreadsheet" or "Plain Text" output.

6. Click "OK." AM will execute the command and display the results in the output format you specified.

7. In the output file, you'll see that the column headings correspond to the cut scores you specified. These labels correspond to the 2003 NAAL literacy levels. For example, if you entered cut scores for the prose scale, the column headings will be the following:

**Figure B-5. Column headings in AM output file, by NAAL prose scale cut scores: 2003**

| Weighted N | Percent in 210.000 | (Standard Error) | Percent in 265.000 | (Standard Error) | Percent in 340.000 | (Standard Error) | Percent above 340.000 | (Standard Error) |
|---|---|---|---|---|---|---|---|---|

SOURCE: Cohen, J., Jiang, T. Gaidurguis, A., & Hollender, D. (2000). AM (Software for marginal maximum likelihood and other statistical analyses of data from complex samples), Washington: American Institutes for Research. See website at http://am.air.org.

The labels correspond to the 2003 NAAL literacy levels as in figure B-6:

**Figure B-6. Column headings in AM output file, by NAAL literacy levels: 2003**

| Weighted N | Percent *Below Basic* | (Standard Error) | Percent *Basic* | (Standard Error) | Percent *Intermediate* | (Standard Error) | Percent *Proficient* | (Standard Error) |
|---|---|---|---|---|---|---|---|---|

SOURCE: Cohen, J., Jiang, T. Gaidurguis, A., & Hollender, D. (2000). AM (Software for marginal maximum likelihood and other statistical analyses of data from complex samples), Washington: American Institutes for Research. See website at http://am.air.org.

**Figure B-7. Example of calculating the percentage of men and women within document literacy levels: 2003**

The following example calculates the percentage of men and women in each of the four NAAL literacy levels *(Below Basic, Basic, Intermediate, and Proficient).* The results match the literacy estimates presented in Figure 6 of the first NAAL report, *A First Look at the Literacy of America's Adults in the 21st Century.*

From the "Statistics" menu, select "MML Procedures for Test Data."

1. Select "NALS Table."
2. As noted in the 2003 NAAL codebook, the variable capturing gender is *DSEX.* To estimate the literacy of men and women, drag the variable *DSEX* from the Variables List window to the Independent Variables box (you can also type *DSEX* in the small box beneath the Independent Variables box).
3. To estimate Document literacy scores, drag the Document test from the Tests window to the Dependent Variables box (you can also type "Document" in the small box beneath the Dependent Variables box).
4. In the Cut Scores window, enter the appropriate cut scores for the Document scale. Double-click "Cut 1," type "205," and then click "OK." Enter "250" and "335" as the second and third cut scores, respectively.
5. Click "OK" and the weighted percentage of men and women in each Document literacy level will appear in the output format you selected. The results are presented as proportions, so you must multiply each proportion and its corresponding standard error by 100 to convert the numbers to percentages.

To estimate the percentage of men and women in each prose, quantitative and health literacy level, replace Document in step 4 with the Prose, Quantitative or Health test variables. You will also need to change the cut scores to match the cut scores for the literacy scale you have selected. To estimate means for a different population group (e.g., country of birth), replace *DSEX* in the Independent Variables box with the variable of interest.

SOURCE: U.S. Department of Education. National Center for Education Statistics. *Technical Report and Data File User's Manual For the 2003 National Assessment of Adult Literacy.*

## TIPS AND SHORTCUTS

### Modifying and Rerunning Models in Current Session

After you execute a procedure in AM, the model appears in the Completed Runs window. You can rerun a model by right-clicking a model and selecting "Copy model to modify and rerun." The specifications for the model will appear. You can change any of the specifications (e.g., change the dependent and independent variables) and then rerun the model by clicking "OK."

### Saving and Rerunning Model Specifications

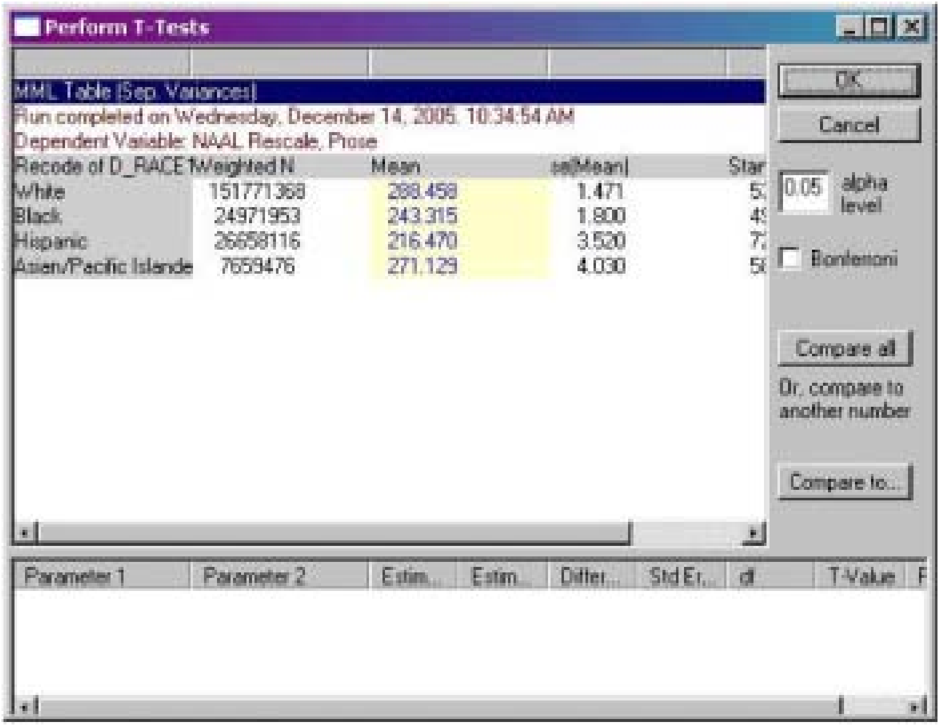
AM allows you to save the specifications for a statistical run to a file and to reload and run the models specified. To save a file, right-click the model in the Completed Runs Queue and select "Save specification to disk." You can either create a new file or add the specification to an existing specification file. We recommend that you store procedures in files with other related models that were originally run on the same data sets. When you rerun them, they will run properly only on data files with variables of the same name.

If you use this feature, it is beneficial to give all your runs informative titles.

### *t* Tests

To access the *t* test dialog box, right-click an icon next to the menu of the Model Run for which you want *t* tests in the Completed Runs Queue and select "T-Tests." That will bring up a dialog box much like figure B-8:

**Figure B-8. AM t test dialog box: 2003**




SOURCE: Cohen, J., Jiang, T. Gaidurguis, A., & Hollender, D. (2000). AM (Software for marginal maximum likelihood and other statistical analyses of data from complex samples), Washington: American Institutes for Research. See website at http://am.air.org.

Notice that some parameter estimates are highlighted in yellow. These are available for *t* testing.

To conduct a *t* test:

1. Move the cursor over the estimate for the first item in the comparison. The cursor will turn into a hand.
2. Click the highlighted item. A blue outline should appear around that cell. This is now the "anchored parameter." Any other cells you click will be compared with the outlined cell.
3. Move the cursor over the parameter to be compared. Again, it should turn into a hand.
4. Click the cell to be compared. The results of the t test will appear in the window at the bottom of the dialog box. Significant results will be highlighted in yellow.
5. To unanchor the anchored parameter (i.e., to conduct tests not involving the specified parameter), simply click the anchored parameter.
6. When you have completed your t tests, click "OK" and the results will be sent to the output format you have selected (e.g., web browser or spreadsheet file).

**Flip Table (Reversing the Independent and Dependent Variables)**

Most analyses using literacy levels present the percentage of adults in a certain group (e.g., gender, race/ethnicity) within each of the four literacy levels. For example, the NAAL *First Look* report shows the percentage of men in each of the four literacy levels as well as the percentage of women in each of the levels.

As noted in the instructions above, these analyses use a population group (e.g., gender, race/ethnicity) as the independent variable and the literacy levels for a particular scale (prose, document, or quantitative) as the dependent variable.

AM can reverse, or "flip," these analyses to show the percentage of adults in a certain literacy level (e.g., *Below Basic* prose literacy) by population groups. For example, the results for employment status summarized in figure 16 (page 10) of the NAAL *First Look* report show the percentage of adults with *Below Basic* prose literacy who were employed full time, employed part time, unemployed, or not in the labor force. The pool of adults for this analysis is the group of adults with *Below Basic* prose literacy; this group is divided across the four employment categories on the basis of their performance on the prose items.

To "flip" results for literacy levels, follow these steps:

1. Run the NALS Table procedure, following the usual procedures. For example, if you wanted to "flip" prose results for employment status, start by selecting prose as the dependent variable and DLFORCE as the independent variable.

2. Once AM has executed the procedure, right-click the completed model in the Completed Runs window. Select "Flip Table."

3. AM will "flip" the results and write the output to the same file.

# Appendix C

# ELECTRONIC CODEBOOK FOR WINDOWS USER'S MANUAL

## Appendix C: Electronic Codebook for Windows User's Manual

The NAAL public use data for the Household Study and the Prison Study are available from the National Center for Education Statistics (NCES) and can be accessed with the Electronic Codebook for Windows (ECBW) produced by NCES staff. The electronic codebook provides the option of producing SPSS, SAS and STATA control statements that can be used to create SPSS, SAS and STATA data files. This appendix describes how to use the ECBW.

When the ECBW is opened, the user will see the main window containing variable names and label names for the 2003 National Assessment of Adult Literacy. At the top of the table are a menu and several buttons, which allow users to move through the ECBW and obtain the information needed. All the menu options are described below. Shortcuts and toolbar buttons are underlined and described within the menu text.

### File Menu

The File menu options give the user the ability to move and export information about the variables selected from the main window.

- **Output.** The Output option of the File menu allows the user to export the codes that have been selected from the table. The user may export this information to one of the following types of files:

    1. *SAS-PC Code*—Allows the user to create a file that tells a SAS program how to read the data.

    2. *SPSS*—Allows the user to create a file that tells a SPSS program how to read in the data. Data in this file are compressed.

    3. *STATA*—Allows the user to create a file that tells a STATA program how to read in the data.

    4. *Codebook Text*—Allows the user to create a file that contains information about the variables selected (frequency, labels, etc.).

    5. *Tag File*—Allows the user to create a file that contains the variables selected so that items do not have to be re-selected each time the user enters the program.

    6. *AccessDB*—Allows the user to create an Access database file that contains the variables selected.

- **View Output.** The View Output option allows the user to look at the output file created.

- **Import Tag Files.** The Import Tag Files option allows the user to recall a previously created tag file.

- **Set Up.** The Set Up option tells the user what directory the files are in and where the files are located.

- **Exit.** This allows the user to exit the ECBW.

## Move Menu

The Move menu allows the user to move between variable names and labels within the ECBW. The menu options follow:

- Top of List
- Prev Section (previous section)
- Next Section
- Bottom of List
- Prev File (previous file)
- Next File

The six black arrow buttons on the toolbar can also be used to move within ECBW without accessing the Move menu.

**Move Button Options:**

| **Arrow Button** | **Function** |
|---|---|
| Left/Right arrows with double lines: | Move to top of list/bottom of list |
| Left/Right arrows with single line: | Move to previous file/next file |
| Left/Right single arrows: | Move to previous section/next section |

## Tag Menu

The Tag menu provides options for selecting and deselecting items.

- **Tag/Untag Items.** The Tag Items option in the Tag menu allows the user to select variables within the ECBW once they are highlighted with the mouse. The Untag Items option allows the user to deselect individual items. The user may also select a variable by simply clicking the box next to each variable. To deselect, the user must click the box a second time.

- **Clear All Tags.** The Clear All Tags option erases all the checked boxes so that the user can make a new selection of items.
- **Previous Tag/Next Tag.** The Previous Tag and Next Tag options allow the user to move back and forth between the selected variables. The user may also move between tags by clicking the blue and red arrow buttons on the toolbar.

## View Menu

Once a variable is selected, the user may choose to obtain a more detailed description of the variable. The View menu provides this information.

- **Description.** When chosen, this option produces a Description/Frequency Window. The user may also reach the Description/Frequency Window by double-clicking the selected variable.
- **Description/Frequency Window.** This window has two options.

    1. *View Description*—The header over the window provides the user with information about the survey and where the information came from. The text inside the window describes the variable (parameters, how it was derived, etc.).
    2. *View Frequency*—This provides the user with the code, frequency, and percent category label of the variable selected.

    To exit the window, the user must click the X button in the top right corner of the Description/Frequency Window.

- **Tagged Items.** This option in the View menu allows the user to create a list of the items that have been tagged/checked from the main window. The user may also create a list of tagged items by clicking the toolbar button that looks like a sheet of paper.

## Search Menu

When the Search menu is selected, a search window will appear. This window allows the user to search the Codebook by variable, label, or description in a forward or backward direction. To exit this window, the user must click the X button in the top right corner of the search window. This window can also be accessed by clicking the magnifying glass button on the toolbar.

## Help Menu

The Help menu (in the far right corner of the main window) provides the user with information pertaining to the ECBW.

- **Contents.** The user may search for information on a certain topic by selecting the Contents option. Once selected, a new window will appear, which allows the user to select a topic of interest.

**Help topics include these:**

  - System overview
  - The ECB Main Window
  - System Requirements
  - How to Tag/Untag a Variable
  - How to View Descriptions/Frequencies
  - How to Create SAS/SPSS Program Code, Codebook Text, Tag and Access Database Files
  - How to Navigate Through Variables
  - How to View Tagged Variables
  - How to Import Tag Files
  - How to Search for Text
  - How to Change the System Setup

- **Search.** If the user is unsure of which topic to select, or would prefer to search for specific words and phrases in the Help feature, the user must click the Search button at the top of the Help screen and follow the given instructions. The user may choose to perform either an Index search or a Find search by clicking the tabs at the top left of the Search Window.

- **Back.** The Back button brings the user back to the main topics page, once a topic has been selected and viewed.

- **Print.** The user can print selections of the Help Manual in hard copy.

- **Other Menu Options: File, Edit, Bookmark, Options, Help.** These features allow the user to further manipulate the information in the Help Manual. Certain pages may be copied, saved to another file, bookmarked, or annotated, if needed.

To exit the Help window, the user must click the X button in the top right corner of the Help window. The user can keep this window open while working within the ECBW by shrinking it on the desktop. To minimize the window, the user must click the dashed line button, also in the top right corner of the Help window (to the left of the X). To enlarge the window, the user must click the view window button, which is to the left of the close window (X) button when minimized at the bottom of the screen.

# Appendix D

## NAAL ITEM DESCRIPTIONS AND PARAMETERS

**Table D-1. Prose literacy item description and parameters for the NAAL, by ID: 2003**

| ID | Item description | A | SE | B1 | SE | B2 | SE |
|---|---|---|---|---|---|---|---|
| CC003 | Chinese New Year: How many celebrate? | 0.41 | 0.023 | -1.73 | 0.102 | -- | -- |
| CC004 | Chinese New Year: Underline what Nian was afraid of | 0.90 | 0.038 | -1.85 | 0.062 | -- | -- |
| N010101 | Marketing: List two facts | 0.84 | 0.041 | 0.36 | 0.033 | -- | -- |
| N010201 | Marketing: Underline sentence explaining action | 1.02 | 0.043 | -0.76 | 0.036 | -- | -- |
| N010301 | Marketing: Give purpose of event | 0.76 | 0.039 | 1.64 | 0.068 | -- | -- |
| N011001 | Trend chart: Determine least # of points needed | 0.65 | 0.033 | -0.67 | 0.044 | -- | -- |
| C020401 | How did that get there? Explain why toy was thrown away | 1.90 | 0.172 | -1.15 | 0.047 | -- | -- |
| C020501 | How did that get there? Determine meaning of metaphor | 1.72 | 0.155 | -0.92 | 0.041 | -- | -- |
| C020901 | BMI Chart: List three health risks | 1.41 | 0.120 | -1.88 | 0.101 | -1.39 | 0.067 |
| C030101 | Colon Cancer: List two reasons to get tested | 1.26 | 0.108 | -1.90 | 0.107 | -1.23 | 0.064 |
| C030201 | Colon Cancer: How often tested? | 1.12 | 0.092 | -2.04 | 0.116 | -- | -- |
| C030301 | Colon Cancer: Describe purpose of pamphlet | 0.76 | 0.061 | -1.47 | 0.089 | -- | -- |
| C040101 | Program for at-risk students: Who provides funding? | 2.55 | 0.280 | -1.19 | 0.065 | -- | -- |
| C040201 | Program for at-risk students: How do parents/teacher communicate? | 1.79 | 0.155 | -0.88 | 0.050 | -- | -- |
| C040301 | Program for at-risk students: List benefits of program | 1.49 | 0.117 | -0.90 | 0.055 | -- | -- |
| C040701 | Olympics article: Describe how women paved way for future athletes | 0.53 | 0.042 | 0.15 | 0.065 | -- | -- |
| C050401 | First Farmer’s Bank: Give required work experience | 0.87 | 0.048 | -1.92 | 0.104 | -0.42 | 0.041 |
| C050801 | Advance directives: Find two types | 0.85 | 0.082 | -1.73 | 0.116 | -- | -- |
| C050901 | Advance directives: Explain benefits of advance directive | 0.92 | 0.066 | 0.07 | 0.046 | -- | -- |
| C051001 | Advance directives: Compare two types of advance directives | 0.66 | 0.055 | 0.75 | 0.070 | -- | -- |
| C051101 | Advance directives: Why redo living will? | 0.42 | 0.047 | -0.48 | 0.100 | -- | -- |
| C060101 | Teen saves dog: How many minutes was dog in lake? | 1.15 | 0.136 | -1.91 | 0.128 | -- | -- |
| C060201 | Teen saves dog: Explain why dog fell | 1.37 | 0.080 | -1.25 | 0.055 | -- | -- |

See notes at end of table.

**Table D-1. Prose literacy item description and parameters for the NAAL, by ID: 2003—Continued**

| ID | Item description | A | SE | B1 | SE | B2 | SE |
|---|---|---|---|---|---|---|---|
| C061001 | What determination can buy: How to build true wealth | 1.10 | 0.124 | -1.78 | 0.122 | -- | -- |
| C061101 | What determination can buy: How would author define true wealth? | 0.53 | 0.070 | 1.67 | 0.196 | -- | -- |
| C070101 | X-ray preparation: What can you drink before your X-ray? | 0.94 | 0.086 | -2.50 | 0.153 | -- | -- |
| C070201 | Making a web page: List one way computers are being used | 1.42 | 0.122 | -1.16 | 0.051 | -- | -- |
| C070401 | Making a web page: How can parents benefit from the service? | 1.24 | 0.082 | -0.60 | 0.041 | -- | -- |
| C070701 | Orioles tickets: Describe how to pick up tickets | 0.94 | 0.060 | -1.40 | 0.068 | -0.31 | 0.040 |
| C070901 | Depression symptoms: Explain why person would fill out checklist | 1.03 | 0.098 | -1.76 | 0.111 | -- | -- |
| C071101 | Medicare & You: Explain meaning of term | 0.72 | 0.056 | -0.93 | 0.088 | 1.52 | 0.123 |
| C080301 | Found Gloves: Explain why author changed what she collects | 0.75 | 0.056 | -0.80 | 0.058 | -- | -- |
| C080401 | Found Gloves: What does collection symbolize to the author? | 1.36 | 0.117 | -1.33 | 0.064 | -- | -- |
| C080601 | Democracy article: Describe how senators were formerly elected | 1.07 | 0.081 | -0.57 | 0.053 | -- | -- |
| C080701 | Democracy article: Compare election systems | 1.52 | 0.120 | 0.17 | 0.035 | 0.61 | 0.041 |
| N090601 | Face off: What group will mandate safe cars? | 1.77 | 0.097 | -1.02 | 0.037 | -- | -- |
| N090701 | Face off: Find correct information in article | 1.70 | 0.107 | -1.04 | 0.039 | -- | -- |
| N090801 | Contrast views on fuel-efficiency vs. size of car | 1.32 | 0.058 | 0.78 | 0.033 | -- | -- |
| N100101 | “Growing up”: Find first buyer’s name | 1.60 | 0.092 | -1.43 | 0.045 | -- | -- |
| N100201 | “Growing up”: Determine correct day of delivery | 1.38 | 0.065 | -0.65 | 0.030 | -- | -- |
| N100301 | “Growing up”: What reason given to stop selling? | 1.26 | 0.052 | -0.67 | 0.032 | -- | -- |
| N100401 | “Growing up”: Compare approaches to selling mags | 0.88 | 0.039 | -0.34 | 0.033 | 1.05 | 0.051 |
| N110101 | Blood pressure: Why difficult to know if high | 1.03 | 0.049 | -1.27 | 0.045 | -- | -- |
| N110401 | Jury: Length of time served by a juror | 0.89 | 0.039 | -0.44 | 0.030 | -- | -- |
| N110501 | Jury: Underline sentence explaining action | 1.11 | 0.055 | -0.83 | 0.040 | -- | -- |
| N110601 | Two challenges attorneys use to jurors | 1.05 | 0.049 | 0.90 | 0.035 | 1.62 | 0.055 |
| N120301 | Ida Chen: What experience turned Ida toward law? | 1.32 | 0.062 | -0.17 | 0.025 | -- | -- |

See notes at end of table.

**Table D-1. Prose literacy item description and parameters for the NAAL, by ID: 2003—Continued**

| ID | Item description | A | SE | B1 | SE | B2 | SE |
|---|---|---|---|---|---|---|---|
| N120401 | Two things Chen did to resolve discrimination conflicts | 1.23 | 0.055 | -0.69 | 0.032 | -0.03 | 0.027 |
| N120501 | Ida Chen: Interpret phrase from article | 0.89 | 0.049 | 1.71 | 0.062 | -- | -- |
| N120901 | Susan Butcher: Find number of wins of sled race | 0.97 | 0.060 | -2.13 | 0.089 | -- | -- |
| N130201 | Fueled: Determine phrase meaning | 1.18 | 0.049 | 0.04 | 0.027 | -- | -- |
| N130301 | Fueled: Give diff and similarity between events | 1.21 | 0.049 | -0.05 | 0.025 | 0.76 | 0.035 |
| N130401 | Fueled: Give suggestion about good value change | 1.96 | 0.105 | 0.57 | 0.026 | -- | -- |
| N130801 | Cost to raise child: Find information from article | 0.83 | 0.039 | -1.34 | 0.055 | -- | -- |

NOTE: For items fitting two-parameter logistic (2PL) model, the A parameter represents the slope or discrimination parameter and the B parameter represents the location or difficulty parameter. For items fitting Graded Response Logistic (GRL) model, the A parameter represents the slope or discrimination parameter; the B1 parameter represents the step parameter from score point 0 to score point 1; and the B2 parameter represents the step parameter from score point 1 to score point 2. SE means standard error.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table D-2. Document literacy item description and parameters for the NAAL, by ID: 2003**

| ID | Item description | A | SE | B1 | SE | B2 | SE |
|---|---|---|---|---|---|---|---|
| CC001 | Snider's Super Foods: Sign name on line | 0.58 | 0.181 | -6.34 | 1.596 | -- | -- |
| CC002 | Appointment slip: Circle date of appointment | 0.62 | 0.048 | -3.36 | 0.213 | -- | -- |
| CC007 | Medicare pamphlet: Find frequency of vaccination | 0.41 | 0.065 | -6.07 | 0.882 | -- | -- |
| N010401 | Vehicle chart: Find correct information | 1.00 | 0.053 | -2.08 | 0.076 | -- | -- |
| N010801 | Trend chart: Mark information on chart | 0.90 | 0.048 | -1.26 | 0.054 | -- | -- |
| N010901 | Trend chart: Put information on chart | 0.70 | 0.038 | 0.77 | 0.047 | -- | -- |
| C020101 | Smithsonian map: Mark correct intersection | 1.13 | 0.097 | -1.55 | 0.092 | -- | -- |
| C020201 | Smithsonian map: Follow directions | 0.90 | 0.075 | -0.37 | 0.050 | -- | -- |
| C021001 | BMI chart: Categorize height and weight | 0.92 | 0.081 | -1.26 | 0.088 | -- | -- |
| C021101 | BMI chart: Find range for weight | 0.98 | 0.073 | 0.34 | 0.046 | -- | -- |
| C030501 | Almanac: Locate page with information | 1.05 | 0.070 | -0.66 | 0.060 | 0.01 | 0.044 |
| C030601 | Almanac: Find table with information | 1.15 | 0.081 | -0.97 | 0.061 | -- | -- |
| C030701 | Catalogue order form: Enter quantities of products | 1.77 | 0.179 | -1.72 | 0.105 | -0.47 | 0.045 |
| C030702 | Catalogue order form: Enter product descriptions | 0.76 | 0.057 | -2.00 | 0.123 | 0.89 | 0.077 |
| C030703 | Catalogue order form: Enter quantity of product sold in multiple units | 2.02 | 0.180 | -1.81 | 0.083 | -1.60 | 0.074 |
| C030705 | Catalogue order form: Enter prices for each product | 1.75 | 0.172 | -2.01 | 0.129 | -0.44 | 0.041 |
| C030708 | Catalogue order form: Enter shipping and handling price | 0.99 | 0.096 | -1.60 | 0.126 | -0.85 | 0.068 |
| C040501 | CHIP application: Enter information | 1.34 | 0.084 | -1.33 | 0.079 | 0.64 | 0.046 |
| C040502 | CHIP application: Check boxes in second column | 1.07 | 0.077 | -0.18 | 0.044 | -- | -- |
| C040503 | CHIP application: Check boxes in third column | 1.17 | 0.100 | -0.30 | 0.044 | -- | -- |
| C040504 | CHIP application: Enter information | 1.94 | 0.186 | -1.03 | 0.063 | -- | -- |
| C050101 | Television schedule: Find program | 0.80 | 0.076 | -1.20 | 0.089 | -- | -- |
| C050201 | Television schedule: What time does program end? | 0.99 | 0.064 | -1.44 | 0.083 | -0.53 | 0.048 |
| C050501 | First Farmer's Bank: Find phone number to get directions | 1.05 | 0.101 | -1.82 | 0.110 | -- | -- |

See notes at end of table.

**Table D-2. Document literacy item description and parameters for the NAAL, by ID: 2003—Continued**

| ID | Item description | A | SE | B1 | SE | B2 | SE |
|---|---|---|---|---|---|---|---|
| C060301 | Ohio weather: Find two cities with likely flight delays | 0.87 | 0.062 | -2.03 | 0.125 | -0.48 | 0.051 |
| C060501 | Vaccination schedule: Find number of vaccines for age | 0.78 | 0.073 | -0.67 | 0.070 | -- | -- |
| C060601 | Vaccination schedule: Find age to receive specific vaccine | 0.76 | 0.065 | -0.17 | 0.060 | -- | -- |
| C060901 | Gas stove ratings: Compare stoves | 0.94 | 0.086 | -0.69 | 0.057 | -- | -- |
| C070501 | Wal-marting of music: Find percentage of market share | 0.75 | 0.080 | -2.55 | 0.195 | -- | -- |
| C071001 | Depression symptoms: What to do | 1.11 | 0.111 | -2.31 | 0.147 | -- | -- |
| C080201 | Alladryl label: Underline substances associated with side effect | 0.75 | 0.056 | -0.91 | 0.071 | -- | -- |
| C080501 | Electric bill: Write amount of check in numbers | 1.84 | 0.260 | -1.49 | 0.094 | -- | -- |
| C080502 | Electric bill: Written amount of check in words | 1.33 | 0.159 | -1.19 | 0.077 | -- | -- |
| C080503 | Electric bill: Write date on check | 0.80 | 0.060 | -1.55 | 0.099 | -- | -- |
| C080504 | Electric bill: Include name of electric company | 1.59 | 0.200 | -1.64 | 0.108 | -- | -- |
| N090301 | Essence: Determine page certain article begins on | 1.17 | 0.066 | -1.93 | 0.065 | -- | -- |
| N090401 | Essence: Determine topic of given article | 0.99 | 0.048 | -1.21 | 0.042 | -- | -- |
| N090501 | Essence: Determine topic of section of magazine | 0.69 | 0.038 | -1.04 | 0.056 | -- | -- |
| N100501 | Opinions table: Mark sentence explaining action | 1.11 | 0.050 | -0.40 | 0.029 | -- | -- |
| N100601 | Opinions table: Find correct group for given info | 1.18 | 0.057 | 0.35 | 0.028 | -- | -- |
| N100701 | Summarize views of parents and teachers | 1.04 | 0.058 | 1.52 | 0.064 | -- | -- |
| N110301 | Certified mail rec't: Enter name and address | 0.75 | 0.041 | -1.32 | 0.054 | -- | -- |
| N110302 | Certified mail rec't: Enter postage and fee | 0.85 | 0.046 | -1.65 | 0.067 | -- | -- |
| N110701 | Credit card table: Find correct bank | 0.43 | 0.029 | -0.97 | 0.078 | -- | -- |
| N110901 | Credit card table: Give 2 differences | 0.86 | 0.038 | -0.63 | 0.048 | 1.16 | 0.045 |
| N120101 | Campus map: Mark map for given info | 0.97 | 0.062 | -1.59 | 0.069 | -- | -- |
| N120201 | Campus map: Find correct room for given dean | 0.95 | 0.047 | -1.23 | 0.048 | -- | -- |
| N120601 | Middle class: Find projected percent | 0.79 | 0.055 | -2.34 | 0.112 | -- | -- |
| N130101 | S.S. card application: Identify and enter info(1) | 1.61 | 0.090 | -0.73 | 0.030 | -- | -- |

See notes at end of table.

**Table D-2. Document literacy item description and parameters for the NAAL, by ID: 2003—Continued**

| ID | Item description | A | SE | B1 | SE | B2 | SE |
|---|---|---|---|---|---|---|---|
| N130102 | S.S. card application: Identify and enter info(3) | 1.40 | 0.075 | -1.27 | 0.046 | -- | -- |
| N130103 | S.S. card application: Identify and enter info(2) | 2.63 | 0.249 | -1.01 | 0.042 | -- | -- |
| N130104 | S.S. card application: Identify and enter info(4) | 2.56 | 0.240 | -0.86 | 0.034 | -- | -- |

NOTE: For items fitting two-parameter logistic (2PL) model, the A parameter represents the slope or discrimination parameter and the B parameter represents the location or difficulty parameter. For items fitting Graded Response Logistic (GRL) model, the A parameter represents the slope or discrimination parameter; the B1 parameter represents the step parameter from score point 0 to score point 1; and the B2 parameter represents the step parameter from score point 1 to score point 2. SE means standard error.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table D-3. Quantitative literacy item description and parameters for the 2003 NAAL, by ID:2003**

| ID | Item description | A | SE | B1 | SE | B2 | SE |
|---|---|---|---|---|---|---|---|
| CC005 | ATM deposit: Calculate total amount | 0.72 | 0.053 | -2.82 | 0.166 | -- | -- |
| CC006 | Magazine subscription: Calculate difference | 0.49 | 0.028 | -1.62 | 0.086 | -- | -- |
| N010501 | Vehicle chart: Find sum of percentages | 0.91 | 0.044 | -0.88 | 0.037 | -- | -- |
| N010601 | Vehicle chart: Describe solution to percent problem | 1.21 | 0.060 | 0.64 | 0.032 | -- | -- |
| N010701 | Vehicle chart: Find magnitude of difference | 1.05 | 0.045 | 0.27 | 0.028 | -- | -- |
| N011101 | Gas gauge: Use info to answer question-show calcs | 1.05 | 0.054 | 0.06 | 0.029 | -- | -- |
| C020301 | Washington Post subscription: Calc total amt for subscription | 0.85 | 0.067 | -1.08 | 0.065 | -- | -- |
| C020601 | Texas BBQ menu: Calc total cost of two items | 1.00 | 0.084 | -1.10 | 0.070 | -- | -- |
| C020701 | Texas BBQ menu: Calc cost for 40 people | 0.74 | 0.064 | -0.90 | 0.085 | -- | -- |
| C020801 | Life insurance: Calc yearly cost of insurance | 0.91 | 0.053 | 0.12 | 0.046 | 0.74 | 0.053 |
| C030401 | Taxicab receipt: Calc change from $20 bill | 0.75 | 0.090 | -2.17 | 0.210 | -- | -- |
| C030704 | Catalogue order form: Calculate quantity | 1.50 | 0.092 | -1.02 | 0.052 | 0.17 | 0.028 |
| C030706 | Catalogue order form: Calculate total price for each product | 4.60 | 0.879 | -1.35 | 0.046 | 0.20 | 0.022 |
| C030707 | Catalogue order form: Calculate subtotal | 1.32 | 0.120 | -1.05 | 0.063 | -- | -- |
| C030709 | Catalogue order form: Calculate grand total | 4.53 | 0.744 | -1.12 | 0.061 | 0.23 | 0.020 |
| C040401 | Long distance rates: Calc difference between two countries | 1.00 | 0.069 | -0.61 | 0.051 | -- | -- |
| C040601 | Fitness center: Find least expensive way to join | 0.76 | 0.044 | 0.16 | 0.044 | -- | -- |
| C040801 | Health insurance table: Calc insurance costs | 0.34 | 0.038 | 1.77 | 0.197 | -- | -- |
| C050301 | First Farmer’s Bank: Calculate weekly salary | 0.74 | 0.053 | -1.76 | 0.115 | -0.93 | 0.073 |
| C050601 | Staples order form: Calculate total cost of two items | 0.85 | 0.074 | -1.42 | 0.093 | -- | -- |
| C050701 | Bus schedule: Find latest bus to arrive on time | 0.91 | 0.074 | 0.96 | 0.070 | -- | -- |
| C060701 | Pillow coupon: Find price with discount coupon | 0.79 | 0.059 | -0.78 | 0.058 | -- | -- |
| C060801 | Gas stove ratings: Calc price difference between two stoves | 0.85 | 0.076 | -1.49 | 0.099 | -- | -- |
| C070301 | Making a web page: Calculate total number with terminals | 0.93 | 0.054 | -1.33 | 0.081 | -0.28 | 0.045 |

See notes at end of table.

**Table D-3. Quantitative literacy item description and parameters for the 2003 NAAL, by ID: 2003—Continued**

| ID | Item description | A | SE | B1 | SE | B2 | SE |
|---|---|---|---|---|---|---|---|
| C070601 | Wal-marting of music: Calculate difference in sales between stores | 0.91 | 0.061 | -0.52 | 0.050 | -- | -- |
| C070801 | Orioles tickets: Calculate total cost of three tickets | 0.79 | 0.064 | -0.47 | 0.054 | -- | -- |
| C080101 | Medicine label: How early can patient take medicine? | 0.85 | 0.060 | -0.82 | 0.066 | -0.32 | 0.040 |
| C080801 | Boat schedule: Find latest boat to arrive on time | 1.00 | 0.070 | -0.19 | 0.049 | -- | -- |
| N090101 | Get discount if oil bill paid in 10 days | 1.49 | 0.077 | -0.20 | 0.024 | -- | -- |
| N090201 | Get net total owed after deduction | 1.74 | 0.111 | -0.48 | 0.025 | -- | -- |
| N090901 | Carpet ad: Get diff in reg and sale price | 0.83 | 0.039 | -1.21 | 0.054 | -- | -- |
| N091001 | Carpet ad: Get total cost to carpet room | 0.62 | 0.036 | 1.29 | 0.064 | -- | -- |
| N100801 | Salt River: Determine difference in costs | 0.68 | 0.041 | -1.64 | 0.082 | -- | -- |
| N100901 | Salt River: Determine miles between stops | 0.62 | 0.038 | -0.46 | 0.040 | -- | -- |
| N101001 | Salt River: Determine hours between points | 1.09 | 0.054 | -0.91 | 0.040 | -- | -- |
| N110201 | Blood pressure: Calculate death rate from info | 1.08 | 0.044 | 0.40 | 0.032 | -- | -- |
| N110303 | Certified mail rec’t: Calculate postage and fees | 0.74 | 0.041 | -1.95 | 0.087 | -- | -- |
| N110801 | Credit card table: Determine difference in rates | 0.98 | 0.058 | -0.73 | 0.043 | -- | -- |
| N120701 | Calc percent diff black & white middle class-1980 | 0.97 | 0.041 | -0.99 | 0.035 | -- | -- |
| N120801 | Middle class: Find difference in magnitude of pct | 1.14 | 0.055 | 0.53 | 0.031 | -- | -- |
| N121001 | Calc miles/day Butcher went in this year’s race | 1.02 | 0.044 | -0.05 | 0.033 | -- | -- |
| N121101 | Susan Butcher: Calc diff in times for completion | 0.95 | 0.056 | 1.40 | 0.055 | -- | -- |
| N130501 | Rec room: Calculate feet of molding needed | 0.70 | 0.036 | 0.61 | 0.039 | -- | -- |
| N130601 | Rec room: Calculate number of wall panels needed | 1.17 | 0.059 | -0.29 | 0.026 | -- | -- |
| N130701 | Rec room: Describe solution of calculation needed | 0.94 | 0.051 | 1.51 | 0.056 | -- | -- |
| CN130901 | Raise child: Calc money needed to raise child | 1.09 | 0.077 | -0.16 | 0.041 | -- | -- |

NOTE: For items fitting two-parameter logistic (2PL) model, the A parameter represents the slope or discrimination parameter and the B parameter represents the location or difficulty parameter. For items fitting Graded Response Logistic (GRL) model, the A parameter represents the slope or discrimination parameter; the B1 parameter represents the step parameter from score point 0 to score point 1; and the B2 parameter represents the step parameter from score point 1 to score point 2. SE means standard error.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

**Table D-4. Health literacy item description and parameters for the 2003 NAAL, by ID: 2003**

| ID | Item description | A | SE | B1 | SE | B2 | SE |
|---|---|---|---|---|---|---|---|
| CC002 | Appointment slip: Circle date of appointment | 0.56 | 0.049 | -3.55 | 0.236 | -- | -- |
| CC007 | Medicare pamphlet: Find frequency of vaccination | 0.34 | 0.060 | -7.11 | 1.139 | -- | -- |
| C020901 | BMI Chart: List three health risks | 1.21 | 0.122 | -2.07 | 0.129 | -1.52 | 0.081 |
| C021001 | BMI chart: Categorize height and weight | 0.93 | 0.080 | -1.29 | 0.082 | -- | -- |
| C021101 | BMI chart: Find range for weight | 0.94 | 0.085 | 0.38 | 0.045 | -- | -- |
| C030101 | Colon Cancer: List two reasons to get tested | 1.56 | 0.153 | -1.76 | 0.100 | -1.13 | 0.060 |
| C030201 | Colon Cancer: How often tested? | 1.65 | 0.203 | -1.76 | 0.102 | -- | -- |
| C030301 | Colon Cancer: Describe purpose of pamphlet | 0.84 | 0.074 | -1.36 | 0.094 | -- | -- |
| C040501 | CHIP application: Enter information | 1.48 | 0.100 | -1.36 | 0.069 | 0.66 | 0.041 |
| C040502 | CHIP application: Check boxes in second column | 1.12 | 0.084 | -0.16 | 0.044 | -- | -- |
| C040503 | CHIP application: Check boxes in third column | 1.14 | 0.093 | -0.31 | 0.043 | -- | -- |
| C040504 | CHIP application: Enter information | 2.57 | 0.333 | -1.05 | 0.045 | -- | -- |
| C040601 | Fitness center: Find least expensive way to join | 0.56 | 0.043 | 0.09 | 0.063 | -- | -- |
| C040801 | Health insurance table: Calc insurance costs | 0.39 | 0.039 | 1.52 | 0.144 | -- | -- |
| C050801 | Advance directives: Find two types | 0.84 | 0.089 | -1.82 | 0.131 | -- | -- |
| C050901 | Advance directives: Explain benefits of advance directive | 0.78 | 0.076 | 0.04 | 0.054 | -- | -- |
| C051001 | Advance directives: Compare two types of advance directives | 0.59 | 0.057 | 0.78 | 0.087 | -- | -- |
| C051101 | Advance directives: Why redo living will? | 0.34 | 0.050 | -0.67 | 0.137 | -- | -- |
| C060501 | Vaccination schedule: Find number of vaccines for age | 0.84 | 0.078 | -0.66 | 0.065 | -- | -- |
| C060601 | Vaccination schedule: Find age to receive specific vaccine | 0.82 | 0.079 | -0.15 | 0.052 | -- | -- |
| C070101 | X-ray preparation: What can you drink before your X-ray? | 1.01 | 0.100 | -2.38 | 0.147 | -- | -- |
| C070901 | Depression symptoms: Explain why person would fill out checklist | 1.15 | 0.126 | -1.71 | 0.113 | -- | -- |
| C071001 | Depression symptoms: What to do | 1.61 | 0.251 | -2.04 | 0.130 | -- | -- |
| C071101 | Medicare & You: Explain meaning of term | 0.64 | 0.063 | -1.09 | 0.109 | 1.64 | 0.153 |

See notes at end of table.

**Table D-4. Health literacy item description and parameters for the 2003 NAAL, by ID: 2003—Continued**

| ID | Item description | A | SE | B1 | SE | B2 | SE |
|---|---|---|---|---|---|---|---|
| C080101 | Medicine label: How early can patient take medicine? | 0.77 | 0.059 | -0.97 | 0.066 | -0.43 | 0.049 |
| C080201 | Alladryl label: Underline substances associated with side effect | 0.71 | 0.061 | -0.97 | 0.071 | -- | -- |
| N110101 | Blood pressure: Why difficult to know if high | 1.01 | 0.092 | -1.31 | 0.079 | -- | -- |
| N110201 | Blood pressure: Calculate death rate from info | 0.95 | 0.091 | 0.21 | 0.050 | -- | -- |

NOTE: For items fitting two-parameter logistic (2PL) model, the A parameter represents the slope or discrimination parameter and the B parameter represents the location or difficulty parameter. For items fitting Graded Response Logistic (GRL) model, the A parameter represents the slope or discrimination parameter; the B1 parameter represents the step parameter from score point 0 to score point 1; and the B2 parameter represents the step parameter from score point 1 to score point 2. SE means standard error.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

*This page is intentionally left blank.*

## Appendix E

## P-VALUES OF NAAL 2003 ITEMS

**Table E-1. P-values of NAAL items: 2003**

| Item | P-value |
|---|---|
| CC001 | 0.956 |
| CC002 | 0.903 |
| CC003 | 0.690 |
| CC004 | 0.847 |
| CC005 | 0.930 |
| CC006 | 0.725 |
| CC007 | 0.945 |
| N010101 | 0.348 |
| N010201 | 0.636 |
| N010301 | 0.162 |
| N010401 | 0.826 |
| N010501 | 0.660 |
| N010601 | 0.267 |
| N010701 | 0.396 |
| N010801 | 0.711 |
| N010901 | 0.301 |
| N011001 | 0.567 |
| N011101 | 0.435 |
| C020101 | 0.791 |
| C020201 | 0.534 |
| C020301 | 0.695 |
| C020401 | 0.745 |
| C020501 | 0.691 |
| C020601 | 0.711 |
| C020701 | 0.643 |
| C020801 | 0.357 |
| C020901 | 0.768 |
| C021001 | 0.678 |
| C021101 | 0.340 |
| C030101 | 0.792 |
| C030201 | 0.836 |
| C030301 | 0.726 |
| C030401 | 0.820 |
| C030501 | 0.524 |
| C030601 | 0.682 |
| C030701 | 0.691 |
| C030702 | 0.496 |

See notes at end of table.

**Table E-1. P-values of NAAL items: 2003—Continued**

| Item | P-value |
|---|---|
| C030703 | 0.802 |
| C030704 | 0.555 |
| C030705 | 0.691 |
| C030706 | 0.588 |
| C030707 | 0.675 |
| C030708 | 0.646 |
| C030709 | 0.540 |
| C040101 | 0.778 |
| C040201 | 0.696 |
| C040301 | 0.687 |
| C040401 | 0.625 |
| C040501 | 0.515 |
| C040502 | 0.495 |
| C040503 | 0.530 |
| C040504 | 0.742 |
| C040601 | 0.417 |
| C040701 | 0.402 |
| C040801 | 0.248 |
| C050101 | 0.692 |
| C050201 | 0.669 |
| C050301 | 0.712 |
| C050401 | 0.672 |
| C050501 | 0.811 |
| C050601 | 0.749 |
| C050701 | 0.237 |
| C050801 | 0.734 |
| C050901 | 0.412 |
| C051001 | 0.280 |
| C051101 | 0.441 |
| C060101 | 0.818 |
| C060201 | 0.751 |
| C060301 | 0.681 |
| C060501 | 0.580 |
| C060601 | 0.477 |
| C060701 | 0.628 |
| C060801 | 0.737 |
| C060901 | 0.586 |

See notes at end of table.

**Table E-1. P-values of NAAL items: 2003—Continued**

| Item | P-value |
|---|---|
| C061001 | 0.749 |
| C061101 | 0.173 |
| C070101 | 0.851 |
| C070201 | 0.736 |
| C070301 | 0.634 |
| C070401 | 0.601 |
| C070501 | 0.824 |
| C070601 | 0.585 |
| C070701 | 0.624 |
| C070801 | 0.558 |
| C070901 | 0.780 |
| C071001 | 0.825 |
| C071101 | 0.365 |
| C080101 | 0.596 |
| C080201 | 0.641 |
| C080301 | 0.609 |
| C080401 | 0.764 |
| C080501 | 0.811 |
| C080502 | 0.743 |
| C080503 | 0.742 |
| C080504 | 0.814 |
| C080601 | 0.565 |
| C080701 | 0.314 |
| C080801 | 0.467 |
| N090101 | 0.522 |
| N090201 | 0.600 |
| N090301 | 0.812 |
| N090401 | 0.692 |
| N090501 | 0.624 |
| N090601 | 0.695 |
| N090701 | 0.701 |
| N090801 | 0.210 |
| N090901 | 0.629 |
| N091001 | 0.176 |
| N100101 | 0.800 |
| N100201 | 0.635 |
| N100301 | 0.639 |

See notes at end of table.

**Table E-1. P-values of NAAL items: 2003—Continued**

| Item | P-value |
|---|---|
| N100401 | 0.340 |
| N100501 | 0.561 |
| N100601 | 0.357 |
| N100701 | 0.106 |
| N100801 | 0.698 |
| N100901 | 0.523 |
| N101001 | 0.656 |
| N110101 | 0.735 |
| N110201 | 0.400 |
| N110301 | 0.654 |
| N110302 | 0.757 |
| N110303 | 0.780 |
| N110401 | 0.536 |
| N110501 | 0.617 |
| N110601 | 0.149 |
| N110701 | 0.567 |
| N110801 | 0.627 |
| N110901 | 0.312 |
| N120101 | 0.773 |
| N120201 | 0.737 |
| N120301 | 0.499 |
| N120401 | 0.533 |
| N120501 | 0.113 |
| N120601 | 0.820 |
| N120701 | 0.693 |
| N120801 | 0.343 |
| N120901 | 0.775 |
| N121001 | 0.462 |
| N121101 | 0.136 |
| N130101 | 0.618 |
| N130102 | 0.764 |
| N130103 | 0.752 |
| N130104 | 0.714 |
| N130201 | 0.420 |
| N130301 | 0.347 |
| N130401 | 0.268 |
| N130501 | 0.322 |

See notes at end of table.

**Table E-1. P-values of NAAL items: 2003—Continued**

| Item | P-value |
|---|---|
| N130601 | 0.522 |
| N130701 | 0.155 |
| N130801 | 0.694 |
| N130901 | 0.458 |

NOTE: P-value for a dichotomous item was calculated as the ratio of the number of respondents who answered the item correctly to the total number of respondents who were administered the item. P-value for a partial-credit item was calculated as the ratio of half of the number of respondents with partially correct response plus the number of respondents with fully correct response to the total number of respondents who were administered the item. All numbers were weighted.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy.

# Appendix F

## NAAL HOUSEHOLD BACKGROUND QUESTIONNAIRE

## Instructions for Reading NAAL English Background Questionnaire

The NAAL background questionnaire is designed to be administered using a Computer Assisted Personal Interviewing (CAPI) system. Interviewers read the questions aloud to respondents from the screens of laptop computers, and they record respondents' answers directly on the computers. The interviewers read aloud response options that are written in lower-case letters, but do not read aloud the response options that are written in upper-case letters.

Administering the background questionnaire using a CAPI system allows for the inclusion of complex skip patterns that target questions only at respondents in a particular subgroup. However, these complex skip patterns make the background questionnaire difficult to follow on paper. Skip patterns are indicated on the paper version of the background questionnaire in two different ways.

For some items where a respondent's choice indicates that he or she should skip to a particular item, the skip pattern is indicated in parenthesis following the response item. For example, if a respondent answers "No" to question A-9, "Have you ever taken a class in the United States to learn English-as-a-second-language, sometimes called an ESL class?," the respondent skips to question A-11, since the question is about how long ago he or she took the class is not relevant for this respondent. If the respondent answers "Yes" to question A-9, the respondent continues to question A-10.

When skip patterns are based on responses to one or more earlier questions in the survey, the skip patterns are indicated in a box that appears before the item that some respondents may skip. For example, question A-4 is skipped by some respondents based on their answer to A-1. The box inserted before question A-4 describes the skip pattern.

For all questions, interviewers had the option to code responses "Don't Know" or "Refused" when appropriate.

# National Study Of America's Adults

## BACKGROUND QUESTIONNAIRE

| IF R IS NOT SCREENER R: | IF R IS SCREENER R: |
|---|---|
| Hello, I am (NAME) from Westat. My organization is helping the United States Department of Education with a very important survey about how adults use printed materials. Recently, another member of your household told me who lives here. Based on this information, you were selected at random to take part in the survey. | You have been selected to participate in the survey. |

I will ask you a short set of questions about your background, education, and work experiences. Then, I will give you a booklet containing exercises based on printed materials, such as newspapers, maps, stories, brochures, forms, and advertisements. Others who have completed these exercises found them interesting and fun. The entire survey will take approximately 90 minutes to complete, and you will be paid $30 for your participation.

Your participation in this survey is voluntary and very important. All of your answers will be kept strictly confidential. All information will be reported for a group as a whole and your responses will not be linked to your name. You do not have to answer any questions you do not want to answer.

IF R REQUESTS MORE INFORMATION ABOUT PURPOSE OF SURVEY:
Today, adults increasingly are expected to use printed information in our society, but there is very little information available on whether or not they are well prepared. This survey will provide information about the reading and writing experiences, activities, and skills of adults in the United States. Information will be used by educators, policymakers, and business leaders to design programs in order to improve the literacy skills of adults.

According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless such collection displays a valid OMB control number. The valid OMB control number for this information collection is 1850-0654 and will expire on **06/30/04**. The time required to complete these forms is estimated to average .45 hours per respondent, including the time to review instructions and complete the survey. **If you have any comments concerning the accuracy of the time estimate or suggestions for improving this form, please write to:** U.S Department of Education, Washington, DC 20202-4651. **If you have any comments or concerns regarding the status of your individual submission of this form, write directly to:** NCES, U.S. Department of Education, 1990 K Street, NW, Washington, DC 20006.

## Section A. General and Language Background

**Interviewer: Code gender of respondent. 1 = Male; 2 = Female**

A-1. In what country were you born?

Look up table ....................................................................................................

**If answered U.S.A. for A-1, go to A-3; otherwise continue.**

A-1A. How old were you when you moved to the United States?

______________
Age

A-2. How many years have you lived in the United States?

| | |
|---|---|
| A. 1 TO 5 .................... | 1 |
| B. 6 TO 10 .................... | 2 |
| C. 11 TO 15 .................... | 3 |
| D. 16 TO 20 .................... | 4 |
| E. 21 TO 30 .................... | 5 |
| F. 31 TO 40 .................... | 6 |
| G. 41 TO 50 .................... | 7 |
| H. 51 OR MORE .................... | 8 |

A-3. What is your date of birth?

______________ __________ ____________
Month Day Year

**If answered U.S.A. for A-1, go to A-5; otherwise continue.**

A-4. What was the highest level of education you completed before coming to the United States? (If response does not fit categories, probe for equivalent.)

| | |
|---|---|
| A. DID NOT ATTEND SCHOOL BEFORE COMING TO U.S. .................... | 1 |
| B. PRIMARY (GRADES K-3) .................... | 2 |
| C. ELEMENTARY (GRADES 4-8) .................... | 3 |
| D. SECONDARY (GRADES 9-12) .................... | 4 |
| E. VOCATIONAL (POSTSECONDARY) .................... | 5 |
| F. COLLEGE/UNIVERSITY .................... | 6 |
| G. OTHER (SPECIFY) .................... | 910 |

A-5. When you were growing up, what language or languages were usually spoken in your home? (Select all that apply.)

Look up table ....................................................................................................

A-6. What language or languages did you learn to speak before you started school? (Select all that apply.)

Look up table ..........................................................................................

**If English only for both A-5 and A-6, go to A-11; otherwise continue.**

A-7. What language did you first learn to read and write?

Look up table ..........................................................................................

A-8. How old were you when you learned to speak English?

| | | |
|---|---|---|
| A. | 1-4 YEARS OLD ........................................................................... | 1 |
| B. | 5-10 YEARS OLD ......................................................................... | 2 |
| C. | 11-15 YEARS OLD ....................................................................... | 3 |
| D. | 16-20 YEARS OLD ....................................................................... | 4 |
| E. | 21 YEARS OR OLDER ................................................................. | 5 |
| F. | DOES NOT SPEAK ENGLISH ....................................................... | 95 |

**If English only in A-6, go to A-11; otherwise continue.**

A-9. Have you ever taken a class in the United States to learn English-as-a-second-language, sometimes called an ESL class?

| | | | |
|---|---|---|---|
| A. | YES ............................................................................................. | 1 | (A-10) |
| B. | NO ............................................................................................... | 2 | (A-11) |

A-10. How long ago did you last take an English-as-a-second-language or ESL class in the United States? Was it…

| | | |
|---|---|---|
| A. | Within the last two years, ............................................................. | 1 |
| B. | 2 to 5 years ago, ........................................................................... | 2 |
| C. | More than 5 years ago, or .............................................................. | 3 |
| D. | Are you taking an ESL class now? ................................................ | 4 |

A-11. Which language do you usually speak now?

Look up table ..........................................................................................

A-12. What other language do you often speak now?

| | | |
|---|---|---|
| A. | Look up table ............................................................................... | |
| B. | No other language spoken .............................................................. | 95 |

A-13. Other than English, what language do you speak best?

| | | |
|---|---|---|
| A. | None ............................................................................................ | 1 |
| B. | Look up table ............................................................................... | |

**If English only in A-6, A-11, A-12, and A-13, go to A-15; otherwise, repeat item A-14 for each non-English language identified in questions A-6, A-11, A-12, and A-13.**

A-14. With regard to **(non-English language in A-6, A-11, A-12, and A-13),** how well do you…

Would you say…

| | Very well | Well | Not well | or | Not at all |
|---|---|---|---|---|---|
| A. Understand it when it is spoken to you? .................... | 1 | 2 | 3 | | 4 |
| B. Speak it?.............................. | 1 | 2 | 3 | | 4 |
| C. Read it? ............................... | 1 | 2 | 3 | | 4 |
| D. Write it? ............................... | 1 | 2 | 3 | | 4 |

A-15. With regard to the English language, how well do you…

Would you say…

| | Very well | Well | Not well | or | Not at all |
|---|---|---|---|---|---|
| A. Understand it when it is spoken to you? .................... | 1 | 2 | 3 | | 4 |
| B. Speak it?.............................. | 1 | 2 | 3 | | 4 |
| C. Read it? ............................... | 1 | 2 | 3 | | 4 |
| D. Write it? ............................... | 1 | 2 | 3 | | 4 |

**If English only for A-6, go to B-1; otherwise continue.**

A-16. **[HAND CARD.]** How difficult is it for you to **(item)** in English?

**HAND CARD 1**

Would you say you have...

| | No difficulty | Some difficulty | Moderate difficulty | or a | Great deal of difficulty | NEVER TRIED |
|---|---|---|---|---|---|---|
| A. Understand people having a conversation with you? ..... | 1 | 2 | 3 | | 4 | 5 |
| B. Understand television, movies, or videos? ............... | 1 | 2 | 3 | | 4 | 5 |
| C. Understand a telephone conversation?...................... | 1 | 2 | 3 | | 4 | 5 |

**If answered 1 to A-15C and A-15D, go to B-1. If answered 1 to A-15C, but A-15D was not answered 1, go to A-18; otherwise continue.**

A-17. **[HAND CARD.]** How difficult is it for you to **(item)** written in English?

Would you say you have...

**HAND CARD 1**

| | No difficulty | Some difficulty | Moderate difficulty | or a | Great deal of difficulty | NEVER TRIED |
|---|---|---|---|---|---|---|
| A. Understand a utility bill, such as telephone or electric? ...... | 1 | 2 | 3 | | 4 | 5 |
| B. Understand the dosage information on over-the-counter medicines? .............. | 1 | 2 | 3 | | 4 | 5 |
| C. Look up information in dictionaries, encyclopedias, phone books, or other reference books? .................. | 1 | 2 | 3 | | 4 | 5 |

A-18. **[HAND CARD.]** How difficult is it for you to fill out forms in English, such as at the doctor's office or at school? Would you say you have...

**HAND CARD 1**

A. No difficulty, .......................................... 1
B. Some difficulty, .......................................... 2
C. Moderate difficulty, or a.......................................... 3
D. Great deal of difficulty? .......................................... 4
E. NEVER TRIED .......................................... 5

## Section B. Educational Background and Experiences

B-1. **[HAND CARD.]** I'd like to ask you about your educational background and experiences. What is the highest level of public or private education you completed? [If respondent went to school outside United States, probe for equivalent.]

**HAND CARD 2**

| | | | |
|---|---|---|---|
| A. | STILL IN HIGH SCHOOL | 1 | (B-9) |
| B. | LESS THAN HIGH SCHOOL (0-8 YEARS) (SPECIFY GRADE) | 2 | (B-2) |
| C. | SOME HIGH SCHOOL (9-12 YEARS BUT DID NOT GRADUATE) (SPECIFY GRADE) | 3 | (B-2) |
| D. | GED OR HIGH SCHOOL EQUIVALENCY | 4 | (B-2) |
| E. | HIGH SCHOOL GRADUATE (12 YEARS; ACCELERATED OR EARLY GRADUATE PROGRAM) | 5 | (B-2) |
| F. | ATTENDED A VOCATIONAL, TRADE, OR BUSINESS SCHOOL AFTER HIGH SCHOOL | 6 | (B-2) |
| G. | COLLEGE: LESS THAN TWO YEARS | 7 | (B-2) |
| H. | COLLEGE: ASSOCIATE'S DEGREE (A.A.) | 8 | (B-2) |
| I. | COLLEGE: TWO YEARS OR MORE, NO DEGREE | 9 | (B-2) |
| J. | COLLEGE GRADUATE (B.S. OR B.A.) | 10 | (B-2) |
| K. | POSTGRADUATE/NO DEGREE | 11 | (B-2) |
| L. | POSTGRADUATE/DEGREE (M.S., M.A., PH.D., M.D., ETC.) | 12 | (B-2) |

B-2. What year did you **(graduate from high school/receive your GED/attend your last year of school)**?

Year ___________________________________

**If answered 2, 3 or 4 for B-1, go to B-4; otherwise continue.**

B-3. What type of high school diploma did you receive? Was it a regular high school diploma from a school in the United States; a regular high school diploma from a school outside the United States run by the United States government, such as a Department of Defense school; a regular high school diploma from a school outside the United States, not run by the United States government; a GED or high school equivalency degree; a certificate of completion that was different from a regular high school diploma; or something else?

| | | |
|---|---|---|
| A. | REGULAR HIGH SCHOOL DIPLOMA FROM A SCHOOL IN THE UNITED STATES | 1 |
| B. | REGULAR HIGH SCHOOL DIPLOMA FROM A SCHOOL OUTSIDE THE UNITED STATES RUN BY THE UNITED STATES GOVERNMENT, SUCH AS A DEPARTMENT OF DEFENSE SCHOOL | 2 |
| C. | REGULAR HIGH SCHOOL DIPLOMA FROM A SCHOOL OUTSIDE THE UNITED STATES, NOT RUN BY THE UNITED STATES GOVERNMENT | 3 |
| D. | GED | 4 |
| E. | CERTIFICATE OF COMPLETION | 5 |
| F. | DID NOT RECEIVE HIGH SCHOOL DIPLOMA | 6 |
| G. | OTHER (SPECIFY) | 910 |

**If answered 10, 11, or 12 for B-1, go to box before B-5; otherwise continue.**

B-4. What was the main reason you stopped your public or private schooling when you did? Please listen to me read all the answer categories, and then tell me which one best describes the reason you stopped your schooling. Was it…

| | |
|---|---|
| A. You are currently in school, | 1 |
| B. Financial problems, | 2 |
| C. Did not do well in school, | 3 |
| D. Did not like school or was bored in school, | 4 |
| E. Expelled from school or asked to leave, | 5 |
| F. Wanted to work, | 6 |
| G. Wanted to go into the military, | 7 |
| H. Personal illness, disability, or pregnancy, | 8 |
| I. Family reasons such as the illness or death of one of your parents, | 9 |
| J. School not available or not accessible, | 10 |
| K. Did not feel safe in school, or | 11 |
| L. Other? | 12 |

**If answered 2 or 3 for B-3, go to box before B-6; otherwise continue. If answered 2 or 3 for B-1, read "attended your last year of school" for B-5; if answered 4 for B-1 or 4 for B-3 read "received your GED" for B-5; otherwise read "graduated from high school" for B-5.**

B-5. When you **(graduated from high school/received your GED/attended your last year of school)**, what state did you live in?

Look up table ............

**If answered 1, 2, 3, 4, 5, or 6 for B-1, go to B-9; otherwise continue. If answered 7 or 9 for B-1, read "attend your last year of college" for B-6. If answered 11 or 12 for B-1, read "receive your undergraduate degree" for B-6. If answered 8 for B-1, read "receive your associate's degree" for B-6. If answered 10 for B-1, read "graduate from college" for B-6.**

B-6. What year did you **(attend your last year of college/receive your undergraduate degree/receive your associate's degree/graduate from college)?**

Year ________________________________________

**If answered 7 or 9 to B-1, go to B-9; otherwise continue.**

B-7. Did you receive your degree from a college in the United States?

| | | |
|---|---|---|
| A. YES | 1 | (B-8) |
| B. NO | 2 | (B-9) |

B-8. In what state was the college where you received your college degree located?

Look up table ............

B-9. How long have you lived in **(this state)**? Would you say…

| | |
|---|---|
| A. Since birth | 1 |
| B. Less than one year | 2 |
| C. 1 to 5 years | 3 |
| D. 6 to 10 years | 4 |
| E. 11 to 15 years | 5 |
| F. 16 to 20 years | 6 |
| G. More than 20 years | 7 |

**If answered 1 for B-1, go to B-11; otherwise continue.**

B-10. Are you currently enrolled in school or college, either full-time or part-time?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |

B-11. Are you currently enrolled in or have you ever taken part in a program other than in regular school in order to improve your *basic skills,* that is, basic reading, writing and arithmetic skills?

| | | |
|---|---|---|
| A. YES | 1 | (B-12) |
| B. NO | 2 | (B-14) |

B-12. How long ago did you last take a class to improve your basic skills? Was it...

| | |
|---|---|
| A. Within the last two years, | 1 |
| B. 2 to 5 years ago, | 2 |
| C. More than 5 years ago, or | 3 |
| D. Are you currently taking a basic skills class? | 4 |

**If answered 2 to A-9, go to B-14; otherwise continue.**

B-13. Was the basic skills class part of the English-as-a-second-language or ESL class you took, or was it a separate class?

| | |
|---|---|
| A. PART OF ESL CLASS | 1 |
| B. SEPARATE CLASS | 2 |

B-14. Have you received any type of information technology skill certification sponsored by a hardware or software manufacturer or an industry or professional association?

| | | |
|---|---|---|
| A. YES | 1 | (B-15) |
| B. NO | 2 | (B-17) |

B-15. Did you have to pass a test to get the certification?

| | | |
|---|---|---|
| A. YES | 1 | (B-16) |
| B. NO | 2 | (B-17) |

B-16. How did you prepare for the test? Was it...

| | YES | NO |
|---|---|---|
| A. A class offered by a four-year college or university? | 1 | 2 |
| B. A class offered by a community college? | 1 | 2 |
| C. A class offered by a technical school or private vendor? | 1 | 2 |
| D. A class offered by a high school or vocational secondary school? | 1 | 2 |
| E. A class offered directly by a hardware or software manufacturer, such as Microsoft, Oracle, Novell, or Cisco? | 1 | 2 |
| F. On the job training or apprenticeship? | 1 | 2 |
| G. Independent study? | 1 | 2 |
| H. Other? (Specify) | 1 | 2 |

B-17. Other than information technology, have you ever received any type of job-related skill certification recognized by a licensing board or an industry or professional association?

| | | |
|---|---|---|
| A. YES | 1 | (B-18) |
| B. NO | 2 | (C-1) |

B-18. Did you have to pass a test to get the certification?

| | | |
|---|---|---|
| A. YES | 1 | (B-19) |
| B. NO | 2 | (C-1) |

B-19. How did you prepare for the test? Was it...

| | YES | NO |
|---|---|---|
| A. A class offered by a four-year college or university? | 1 | 2 |
| B. A class offered by a community college? | 1 | 2 |
| C. A class offered by a technical school or private vendor? | 1 | 2 |
| D. A class offered by a high school or vocational secondary school? | 1 | 2 |
| E. On the job training or apprenticeship? | 1 | 2 |
| F. Independent study? | 1 | 2 |
| G. Other? (Specify) | 1 | 2 |

## Section C. Political and Social Participation

C-1. **[HAND CARD.]** I'd like to find out how you usually get information about current events, public affairs, and the government. How much information about current events, public affairs and the government do you get from…

**HAND CARD 3**

| | Would you say… A lot | Some | A little | or | None |
|---|---|---|---|---|---|
| A. Newspapers? | 1 | 2 | 3 | | 4 |
| B. Magazines? | 1 | 2 | 3 | | 4 |
| C. Internet? | 1 | 2 | 3 | | 4 |
| D. Radio and television? | 1 | 2 | 3 | | 4 |
| E. Books or brochures? | 1 | 2 | 3 | | 4 |
| F. Family members, friends, or co-workers? | 1 | 2 | 3 | | 4 |

**If English only for A-6, go to C-3; otherwise continue.**

C-2. How much of the information you get about current events, public affairs, and the government is presented in **(non-English language in A-6, A-11, A-12, or A-13)**? Would you say...

| | | |
|---|---|---|
| A. | All, | 1 |
| B. | Most, | 2 |
| C. | Some, or | 3 |
| D. | None? | 4 |

C-3. During the past year, did you give any **UNPAID** time as a volunteer to a group or organization?

| | | | |
|---|---|---|---|
| A. | YES | 1 | (C-4) |
| B. | NO | 2 | (C-5) |

C-4. How often do you volunteer? Would you say...

| | | |
|---|---|---|
| A. | Most days, | 1 |
| B. | A few days a week, | 2 |
| C. | About once a week, or | 3 |
| D. | Less than once a week? | 4 |

C-5. How many hours do you usually watch television, videotapes, or DVDs each day?

| | | |
|---|---|---|
| A. NONE | 0 | |
| B. 1 HOUR OR LESS | 1 | |
| C. 2 HOURS | 2 | |
| D. 3 HOURS | 3 | |
| E. 4 HOURS | 4 | |
| F. 5 HOURS | 5 | |
| G. 6 OR MORE HOURS | 6 | |

C-6. How often do you use the services of a library for any reason? Would you say...

| | | |
|---|---|---|
| A. Daily, | 1 | (C-7) |
| B. Weekly, | 2 | (C-7) |
| C. Monthly, | 3 | (C-7) |
| D. Once or twice a year, or | 4 | (C-7) |
| E. Never? | 5 | (Box before C-8) |

C-7. During the past month, did you borrow any materials from a library?

| | | |
|---|---|---|
| A. YES | 1 | |
| B. NO | 2 | |

**If answered (born in U.S.A.) or (U.S. territory) to A-1, go to C-9; otherwise continue.**

C-8. Are you a citizen of the United States?

| | | |
|---|---|---|
| A. YES | 1 | |
| B. NO | 2 | |

C-9. Did you ever serve on active duty in the U.S. Armed Forces?

| | | |
|---|---|---|
| A. YES | 1 | |
| B. NO | 2 | |

**If answered 2 to C-8, go to D-1; otherwise continue.**

C-10. In 2000, Al Gore ran on the Democratic ticket against George W. Bush for the Republicans. Do you remember for sure whether or not you voted in that election?

| | | |
|---|---|---|
| A. YES, I REMEMBER FOR SURE | 1 | (C-11) |
| B. NO, DON'T REMEMBER FOR SURE | 2 | (C-12) |
| C. YES, I VOTED | 3 | (C-12) |
| D. NO, I DIDN'T VOTE | 4 | (C-12) |

C-11. Did you vote in that election?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |

C-12. Are you currently registered to vote?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |

## Section D. Labor Force Participation

D-1. **[HAND CARD.]** Now I'd like to ask you some questions about what you were doing last week. Last week were you…(Select all that apply.)

**HAND CARD 4**

| | |
|---|---|
| A. Working a full-time job for pay or profit, that is, 35 hours or more? | 1 |
| B. Working for pay or profit part-time, that is, 1 to 34 hours? | 2 |
| C. Working two or more part-time jobs for pay, totaling 35 or more hours? | 3 |
| D. Unemployed, laid off, or looking for work? | 4 |
| E. With a job but not at work because of temporary illness, vacation, or work stoppage? | 5 |
| F. With a job but on family leave (maternity or paternity leave)? | 6 |
| G. In school? | 7 |
| H. Keeping house? | 8 |
| I. Retired? | 9 |
| J. Doing volunteer work? | 10 |
| K. OTHER (SPECIFY) | 910 |

**If answered 1, 5 or 6 to D-1, go to D-3; otherwise continue.**

D-2. Have you looked for a job at any time during the past four weeks?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |

**If answered 4, 7, 8, 9, 10 or 910 only to D-1, go to D-5; otherwise continue.**

D-3. Last week, what was your total weekly wage or salary from all jobs *before* any deductions? Include tips and commissions. (Write in dollar amount *and* select appropriate code.)

$___________ . ___________

| | |
|---|---|
| A. PER HOUR | 1 |
| B. PER DAY | 2 |
| C. PER WEEK | 3 |
| D. PER TWO-WEEK PERIOD | 4 |
| E. PER MONTH | 5 |
| F. PER YEAR | 6 |
| G. OTHER (SPECIFY) | 910 |

**If answered "$0" for D-3, go to D-4; otherwise continue.**

D-3A. Was that take-home pay or gross pay?

| | |
|---|---|
| A. TAKE-HOME PAY | 1 |
| B. GROSS PAY | 2 |

D-4. How many hours or days did you work last week?

HOURS:____________________

DAYS: ____________________

D-5. Now I'd like to ask you some questions about your work during the past 12 months. Including weeks of paid leave, such as vacation and sick leave, how many weeks did you work for pay or for profit during the past 12 months?

| | | |
|---|---|---|
| A. NONE (0) | 1 | (D-6) |
| B. LESS THAN 52 WEEKS (SPECIFY NUMBER OF WEEKS): | 2 | (D-6) |
| C. 52 WEEKS (FOR THE LAST 12 MONTHS) | 3 | (D-7) |

D-6. Of the weeks you were not employed, what were you doing? Were you…
(Select all that apply.)

| | |
|---|---|
| A. Ill, or disabled and unable to work, | 1 |
| B. Retired, | 2 |
| C. Taking care of home or family, | 3 |
| D. Going to school, | 4 |
| E. Could not find work, or | 5 |
| F. Doing something else? (Specify) | 6 |

**If answered 1 to D-5, go to D-9; otherwise continue.**

D-7. For the past 12 months, what was your average weekly wage or salary *before* any deductions? Include tips and commissions. (Write in dollar amount *and* select appropriate code.)

$___________ . ___________

| | |
|---|---|
| A. PER HOUR | 1 |
| B. PER DAY | 2 |
| C. PER WEEK | 3 |
| D. PER TWO-WEEK PERIOD | 4 |
| E. PER MONTH | 5 |
| F. PER YEAR | 6 |
| G. OTHER (SPECIFY) | 910 |

**If answered 3 for D-5, go to D-8; otherwise continue.**

D-7A. Is that your average wage or salary for the entire year, or just for the weeks you worked?

| | |
|---|---|
| A. ENTIRE YEAR | 1 |
| B. JUST WEEKS WORKED | 2 |

D-8. On average, how many hours or days did you work each week during the past 12 months?

HOURS:____________________
DAYS: ____________________

**If answered 3 for D5, go to D9; otherwise continue.**

D-8A. Is that your average for the entire year or just for the weeks you worked?

| | |
|---|---|
| A. ENTIRE YEAR | 1 |
| B. JUST WEEKS WORKED | 2 |

D-9. Which of the following describes your work history? Have you…

| | | |
|---|---|---|
| A. Held a paying job within the last three years, .......................................... | 1 | (D-10) |
| B. Held a paying job, but not within the last three years, or............................ | 2 | (Box before E-1) |
| C. Never been employed for pay full-time or part-time?............................... | 3 | (Box before E-1) |

D-10. Now I'd like to ask you some questions about your current full-time or part-time job or your most recent full-time or part-time job. For what kind of business or industry (do/did) you work? (For example, television and radio manufacturing, retail shoe store, state labor department, farm, etc.) (If R is working two or more jobs, probe: Tell me about the job you work the most hours or the job you consider your primary employment.)

BUSINESS OR INDUSTRY: ____________________________________________

____________________________________________

____________________________________________

____________________________________________

D-11. What (is/was) your occupation, that is, what (is/was) your job called? (For example, electrical engineer, stock clerk, typist, farmer, etc.)

OCCUPATION: ____________________________________________

____________________________________________

____________________________________________

____________________________________________

D-12. What (are/were) the most important activities or duties at this job? (For example, typing, keeping account books, filing, selling cars, operating a printing press, finishing concrete, etc.)

ACTIVITIES OR DUTIES: ____________________________________________

____________________________________________

____________________________________________

____________________________________________

____________________________________________

____________________________________________

D-13. (Are/were) you employed by government, by a **PRIVATE** employer, or (are/were) you self-employed or working in a family business?

| | |
|---|---|
| A. GOVERNMENT.......................................................................................... | 1 |
| B. PRIVATE EMPLOYER................................................................................ | 2 |
| C. SELF-EMPLOYED....................................................................................... | 3 |
| D. WORKING IN A FAMILY BUSINESS ........................................................... | 4 |

**If answered 1 to D-5, go to box before E-1. Otherwise, if answered 3 to D-13, continue to D-14; if answered 1, 2, or 4 to D-13, go to D-15.**

D-14. Were you self-employed for all of the past 12 months or did you have any other jobs?

| | | |
|---|---|---|
| A. SELF-EMPLOYED ALL YEAR ................................................................ | 1 | (Box before E-1) |
| B. HAD OTHER JOBS .................................................................................. | 2 | (D-15) |

D-15. For how many employers did you work during the past 12 months?

| | |
|---|---|
| A. ONE EMPLOYER | 1 |
| B. TWO EMPLOYERS | 2 |
| C. THREE EMPLOYERS | 3 |
| D. FOUR EMPLOYERS | 4 |
| E. FIVE OR MORE EMPLOYERS | 5 |

## Section E. Literacy Practices

**If answered 1, 2, or 3 to C-1C, go to E-2; otherwise continue.**

E-1. Do you ever use a computer?

A. YES .................................................. 1
B. NO .................................................. 2

E-2. **[HAND CARD.]** Now I'd like to talk to you about what you read in English. How often do you read **(item)** in English?

Would you say...

**HAND CARD 5**

| | Every day | A few times a week | Once a week | Less than once a week | or | Never |
|---|---|---|---|---|---|---|
| A. Newspapers or magazines ... | 1 | 2 | 3 | 4 | | 5 |
| B. Books.................................... | 1 | 2 | 3 | 4 | | 5 |
| C. Letters and notes .................. | 1 | 2 | 3 | 4 | | 5 |

**If English only in A-6, go to E-4A; otherwise continue.**

E-3. **[HAND CARD.]** How often do you read **(item)** in **(non-English language in A-6, A-11, A-12, or A-13)**?

Would you say...

**HAND CARD 5**

| | Every day | A few times a week | Once a week | Less than once a week | or | Never |
|---|---|---|---|---|---|---|
| A. Newspapers or magazines ... | 1 | 2 | 3 | 4 | | 5 |
| B. Books.................................... | 1 | 2 | 3 | 4 | | 5 |
| C. Letters and notes .................. | 1 | 2 | 3 | 4 | | 5 |

E-4A. How often do you read the nutritional information on food labels written in English? Would you say...

A. Every time I buy a food I never bought before, ........................................ 1
B. Most of the time when I buy a food I never bought before, .......................... 2
C. Sometimes when I buy a food I never bought before, or............................. 3
D. Never?.................................................. 4

E-4B. **[HAND CARD.]** How often do you look up a schedule in a movie or TV guide written in English? Would you say...

**HAND CARD**
**5**

| | | |
|---|---|---|
| A. | Every day, | 1 |
| B. | A few times a week, | 2 |
| C. | Once a week, | 3 |
| D. | Less than once a week, or | 4 |
| E. | Never? | 5 |

**If answered 2 or 3 to D-9, go to box before E-6; otherwise continue.**

**If answered 1 to E-1, display "other than email" for E-5A.**

E-5. **[HAND CARD.]** Now, I'd like to ask you some questions about what you read at work. How often (do/did) you read or use information from **(Item)** as part of your (current/most recent) job?

Would you say...

**HAND CARD**
**5**

| | Every day | A few times a week | Once a week | Less than once a week | or | Never |
|---|---|---|---|---|---|---|
| A. Letters or memos (other than e-mail) | 1 | 2 | 3 | 4 | | 5 |
| B. Reports, articles, magazines, or journals | 1 | 2 | 3 | 4 | | 5 |
| C. Manuals or reference books, including catalogs or parts lists | 1 | 2 | 3 | 4 | | 5 |
| D. Directions or instructions for medicines, recipes, or other products | 1 | 2 | 3 | 4 | | 5 |
| E. Diagrams or schematics | 1 | 2 | 3 | 4 | | 5 |
| F. Bills, invoices, spreadsheets, or budget tables | 1 | 2 | 3 | 4 | | 5 |
| G. Health and safety information in postings or booklets | 1 | 2 | 3 | 4 | | 5 |

**If answered 2 to E-1, go to E-7; otherwise continue.**

E-6. **[HAND CARD.]** Now I'd like to ask you about how you use the computer. How often do you **(item)**?

**HAND CARD 5**

Would you say...

| | Every day | A few times a week | Once a week | Less than once a week | or | Never |
|---|---|---|---|---|---|---|
| A. Send or receive an e-mail message? | 1 | 2 | 3 | 4 | | 5 |
| B. Write using a word processing program? | 1 | 2 | 3 | 4 | | 5 |
| C. Use a spreadsheet program or use a financial program, such as an electronic check-book, money management, or tax program? | 1 | 2 | 3 | 4 | | 5 |
| D. Look up information on a CD-ROM? | 1 | 2 | 3 | 4 | | 5 |
| E. Find information on the Internet? | 1 | 2 | 3 | 4 | | 5 |
| F. Talk in chat groups or with other people who are logged onto the Internet at the same time you are? | 1 | 2 | 3 | 4 | | 5 |

**If answered 1 to E-1, display "and email" for E-7C.**

E-7. **[HAND CARD.]** How much help do you get from family members or friends with…

**HAND CARD 3**

Would you say…

| | A lot | Some | A little | or | None |
|---|---|---|---|---|---|
| A. Filling out forms? | 1 | 2 | 3 | | 4 |
| B. Reading or explaining newspaper articles or other written information? | 1 | 2 | 3 | | 4 |
| C. Writing notes, letters (and email?) | 1 | 2 | 3 | | 4 |
| D. Using basic arithmetic, that is, adding, subtracting, multiplying, or dividing, such as filling out order forms or balancing a checkbook? | 1 | 2 | 3 | | 4 |

**If answered 1 to D5, go to F1-C; otherwise continue.**

## Section F. Job Training and Skills

F-1. During the past year, did you participate in any training or education, including courses, workshops, formal on-the-job training or apprenticeships to:

| | YES | NO |
|---|---|---|
| A. [Employed within past year only.] Help you do your job better? ............. | 1 | 2 |
| B. [Employed within past year only.] Help you get a promotion or a new job? ............. | 1 | 2 |
| C. [Not employed for entire past year only.] Help you get a job? ............. | 1 | 2 |

**If answered 2 to all parts of F-1, go to box before F-7; otherwise continue.**

F-2. Did this training or education include instruction intended to:

| | YES | NO |
|---|---|---|
| A. Improve your English reading skills? ............. | 1 | 2 |
| B. Improve your English writing skills? ............. | 1 | 2 |
| C. Improve your arithmetic or mathematics skills? ............. | 1 | 2 |
| D. Improve your computer skills? ............. | 1 | 2 |
| E. Help you communicate or work better with co-workers? ............. | 1 | 2 |

**If answered 1 to D-5, go to box before F-7; otherwise continue.**

F-3. Did your employer require you to participate in this training or education?

A. YES ............. 1
B. NO ............. 2

F-4. Did your employer pay at least part of the cost of this training or education?

A. YES ............. 1
B. NO ............. 2

F-5. Did your employer pay for any of your time when you participated in this training or education?

A. YES ............. 1
B. NO ............. 2

F-6. Was any of this training or education provided through a union or trade association agreement?

A. YES ............. 1
B. NO ............. 2

**If age 66 or older (based on A-3) and answered 9 to D-1, go to G-1; otherwise continue.**

F-7. How much do you think your reading skills limit your job opportunities—for example, to get a promotion or a (different) job you would like to have? Would you say a lot, some, a little, or not at all?

| | |
|---|---|
| A. A LOT ........................................ | 1 |
| B. SOME ........................................ | 2 |
| C. A LITTLE ........................................ | 3 |
| D. NOT AT ALL ........................................ | 4 |

F-8. How much do you think your writing skills limit your job opportunities—for example, to get a promotion or a (different) job you would like to have? Would you say a lot, some, a little, or not at all?

| | |
|---|---|
| A. A LOT ........................................ | 1 |
| B. SOME ........................................ | 2 |
| C. A LITTLE ........................................ | 3 |
| D. NOT AT ALL ........................................ | 4 |

F-9. How much do you think your math skills limit your job opportunities—for example, to get a promotion or a (different) job you would like to have? Would you say a lot, some, a little, or not at all?

| | |
|---|---|
| A. A LOT ........................................ | 1 |
| B. SOME ........................................ | 2 |
| C. A LITTLE ........................................ | 3 |
| D. NOT AT ALL ........................................ | 4 |

F-10. How much do you think your computer skills limit your job opportunities—for example, to get a promotion or a (different) job you would like to have? Would you say a lot, some, a little, or not at all?

| | |
|---|---|
| A. A LOT ........................................ | 1 |
| B. SOME ........................................ | 2 |
| C. A LITTLE ........................................ | 3 |
| D. NOT AT ALL ........................................ | 4 |

## Section G. Demographic Information

Now I'm going to ask you some questions about your family.

G-1. In what country was your mother (stepmother or female guardian) born?

A. Look up table ..........

G-2. What was the highest level of education your mother (stepmother or female guardian) completed? (If went to school outside U.S., probe for equivalent)

| | |
|---|---|
| A. LESS THAN HIGH SCHOOL (0-8 YEARS ) (SPECIFY GRADE) | 1 |
| B. SOME HIGH SCHOOL (9-12 YEARS BUT DID NOT GRADUATE) (SPECIFY GRADE) | 2 |
| C. GED OR HIGH SCHOOL EQUIVALENCY | 3 |
| D. HIGH SCHOOL GRADUATE (12 YEARS; ACCELERATED OR EARLY GRADUATE PROGRAM) | 4 |
| E. ATTENDED A VOCATIONAL, TRADE, OR BUSINESS SCHOOL AFTER HIGH SCHOOL | 5 |
| F. COLLEGE: LESS THAN TWO YEARS | 6 |
| G. COLLEGE: ASSOCIATE'S DEGREE (A.A.) | 7 |
| H. COLLEGE: TWO YEARS OR MORE, NO DEGREE | 8 |
| I. COLLEGE GRADUATE (B.S. OR B.A.) | 9 |
| J. POSTGRADUATE/NO DEGREE | 10 |
| K. POSTGRADUATE/DEGREE (M.S., M.A., Ph.D., M.D., ETC.) | 11 |

G-3. In what country was your father (stepfather or male guardian) born?

A. Look up table ..........

G-4. What was the highest level of education your father (stepfather or male guardian) completed? (If went to school outside U.S., probe for equivalent.)

| | |
|---|---|
| A. LESS THAN HIGH SCHOOL (0-8 YEARS ) (SPECIFY GRADE) | 1 |
| B. SOME HIGH SCHOOL (9-12 YEARS BUT DID NOT GRADUATE) (SPECIFY GRADE) | 2 |
| C. GED OR HIGH SCHOOL EQUIVALENCY | 3 |
| D. HIGH SCHOOL GRADUATE (12 YEARS; ACCELERATED OR EARLY GRADUATE PROGRAM) | 4 |
| E. ATTENDED A VOCATIONAL, TRADE, OR BUSINESS SCHOOL AFTER HIGH SCHOOL | 5 |
| F. COLLEGE: LESS THAN TWO YEARS | 6 |
| G. COLLEGE: ASSOCIATE'S DEGREE (A.A.) | 7 |
| H. COLLEGE: TWO YEARS OR MORE, NO DEGREE | 8 |
| I. COLLEGE GRADUATE (B.S. OR B.A.) | 9 |
| J. POSTGRADUATE/NO DEGREE | 10 |
| K. POSTGRADUATE/DEGREE (M.S., M.A., Ph.D., M.D., ETC.) | 11 |

## Section H. Family Literacy

**If respondent is under age 18, add parentheses to H-1.**

H-1. During the past month, how many children (other than you) under 18 lived in this household for 10 or more days?

Number of children ___________________________________

**If answered "0" or "none" go to H-13; otherwise continue.**

H-2. What are their ages?____________________________________________________

**Repeat H-3 for every child under 18.**

H-3. How are you related to the [age of child] year old?

| | |
|---|---|
| A. PARENT/GUARDIAN/STEP-PARENT | 1 |
| B. GRANDPARENT/STEP-GRANDPARENT/GREAT GRANDPARENT | 2 |
| C. SIBLING/STEP-SIBLING/HALF SIBLING | 3 |
| D. OTHER RELATIVE | 4 |
| E. NOT RELATED | 5 |

**If answered 3, 4 or 5 for H-3, go to H-13, otherwise continue.**

H-4. During the past year, have you participated in any parenting groups or classes?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |

**If no children under age 8, go to instructions before H-10.**

Now I'd like to ask you some questions with regard to your child (children/grandchild/grandchildren) who is (are) under 8. [Please answer these questions only with regard to that (these) child (children/grandchild/grandchildren) not your older child (children/grandchild/grandchildren.)]

H-5. Since last [insert the current day of the week], have you read to or with your child (children/grandchild/grandchildren)?

| | | |
|---|---|---|
| A. YES | 1 | (H-6) |
| B. NO | 2 | (H-7) |

H-6. Since last [insert the current day of the week], on how many different days did you read to or with your child (children/grandchild/grandchildren)? Would you say it was...

| | |
|---|---|
| A. Every day, | 1 |
| B. 5 or 6 days, | 2 |
| C. 3 or 4 days, or | 3 |
| D. 1 or 2 days? | 4 |

H-7. **[HAND CARD.]** During the past month, about how often did you try to teach your child (children/grandchild/grandchildren) the letters of the alphabet? Would you say every day, a few times a week, once a week, less than once a week, never, or does (do) your child (children/grandchild/grandchildren) already know the letters of the alphabet?

**HAND CARD 5**

| | |
|---|---|
| A. EVERY DAY | 1 |
| B. A FEW TIMES A WEEK | 2 |
| C. ONCE A WEEK | 3 |
| D. LESS THAN ONCE A WEEK | 4 |
| E. NEVER | 5 |
| F. CHILD (CHILD/GRANDCHILD/GRANDCHILDREN) ALREADY KNOWS THE LETTERS OF THE ALPHABET | 6 |

H-8. **[HAND CARD.]** During the past month, how often did you point out words to your child (children/grandchild/grandchildren) and ask him (her/them) what they say? Would you say every day, a few times a week, once a week, less than once a week, never, or does (do) your child (children/grandchild/grandchildren) already read well?

**HAND CARD 5**

| | |
|---|---|
| A. EVERY DAY | 1 |
| B. A FEW TIMES A WEEK | 2 |
| C. ONCE A WEEK | 3 |
| D. LESS THAN ONCE A WEEK | 4 |
| E. NEVER | 5 |
| F. CHILD (CHILD/GRANDCHILD/GRANDCHILDREN) ALREADY READS WELL | 6 |

H-9. **[HAND CARD.]** During the past month, about how often did you sing songs, recite poems or nursery rhymes, or engage in other activities that included rhyming words with your child (children/grandchild/grandchildren)? Would you say...

**HAND CARD 5**

| | |
|---|---|
| A. Every day, | 1 |
| B. A few times a week, | 2 |
| C. Once a week, | 3 |
| D. Less than once a week, or | 4 |
| E. Never? | 5 |

**[If no child age 5 or older, go to H-13.]** Now I'd like to ask you some questions about your school-age child (children/grandchild/grandchildren).

H-10. **[HAND CARD.]** During a typical school month, how often do you talk to your school-age child (children/grandchild/grandchildren) about things they have studied in school? Would you say...

**HAND CARD 5**

| | | |
|---|---|---|
| A. | Every day, | 1 |
| B. | A few times a week, | 2 |
| C. | Once a week, | 3 |
| D. | Less than once a week, or | 4 |
| E. | Never? | 5 |

H-11. **[HAND CARD.]** During a typical school month, how often do you help or work with your school-age child (children/grandchild/grandchildren) on homework? Would you say...

**HAND CARD 5**

| | | |
|---|---|---|
| A. | Every day, | 1 |
| B. | A few times a week, | 2 |
| C. | Once a week, | 3 |
| D. | Less than once a week, or | 4 |
| E. | Never? | 5 |

H-12. During the past year, have you **(item)**

| | | YES | NO |
|---|---|---|---|
| A. | Volunteered to help out at your child's (one of your children's/ grandchildren/grandchild) school(s), including in the classroom, on a field trip, or at a school event such as a party or school fair? | 1 | 2 |
| B. | Gone to a PTA or other type of parent meeting at your child's (one of your children's/grandchildren/grandchild) school(s)? | 1 | 2 |
| C. | Spoken individually with your child's (one of your children's/ grandchildren/grandchild) teacher(s) to see how he or she was doing in school? | 1 | 2 |
| D. | Sent food, or other items to share in your child's (one of your children's/grandchildren/grandchild) classroom(s)? | 1 | 2 |

H-13. Now I'm going to read you a series of statements. Please tell me if each of the following statements is true or false.

| | | TRUE | FALSE |
|---|---|---|---|
| A. | There are 25 or more books in your home right now. | 1 | 2 |
| B. | There is a variety of magazines and other reading materials in your home. | 1 | 2 |
| C. | [Read only if there are children over age 2 in the household]. The child (children/grandchild/grandchildren) living in your home often see you reading. | 1 | 2 |
| D. | [Read only if there are children over age 2 in the household.] The child (children/grandchild/grandchildren) living in your home have their own books. | 1 | 2 |

H-14. How many computers do you have in your household that can be used for word processing, that is, writing letters or other documents?

____________________ Computers

H-15. How many computers do you have in your household that can access the Internet or World Wide Web?

____________________ Computers

## Section I. Household Income and Welfare Participation

I would like to ask you some questions about your household.

I-1. **[HAND CARD.]** First, which letter on this card describes your current marital status?

**HAND CARD 6**

| | |
|---|---|
| A. NEVER MARRIED | 1 |
| B. MARRIED, LIVING WITH SPOUSE | 2 |
| C. MARRIED, SPOUSE LIVING ELSEWHERE | 3 |
| D. LIVING AS MARRIED | 4 |
| E. SEPARATED OR DIVORCED | 5 |
| F. WIDOWED | 6 |

I-2. Including yourself, how many people in your household are employed or work for pay or wages?

| | |
|---|---|
| A. NONE | 0 |
| B. ONE | 1 |
| C. TWO | 2 |
| D. THREE OR MORE | 3 |

I-3. Did you or anyone in your household receive any of the following during the past 12 months? [Do not read the words in parentheses. They are there for clarification if the respondents ask. For each question to which a respondent answers "Yes," ask, "Is that you, someone else, or both you and someone else in your household?"]

| | Yes, me | Yes, someone else | Yes, someone else and me | No |
|---|---|---|---|---|
| A. Social Security or Railroad Retirement payments | 1 | 2 | 3 | 4 |
| B. Supplemental Security Income (SSI) | 1 | 2 | 3 | 4 |
| C. Other retirement, survivor, or disability payments (other than Social Security or Railroad Retirement) | 1 | 2 | 3 | 4 |
| D. Food stamps | 1 | 2 | 3 | 4 |
| E. WIC supplemental nutrition benefits (Women, infants, and children supplemental nutrition benefits) | 1 | 2 | 3 | 4 |
| F. Rent subsidy, such as Section 8 or public housing | 1 | 2 | 3 | 4 |
| G. Temporary Assistance for Needy Families (TANF), public assistance, or public welfare payments from the state or local welfare office | 1 | 2 | 3 | 4 |
| H. Interest from savings or other bank accounts (other than dividends) | 1 | 2 | 3 | 4 |
| I. Dividend income from stocks or mutual funds or income from rental property, royalty, estates, or trusts | 1 | 2 | 3 | 4 |

**If answered 1 or 3 to I-3G, go to I-4. If answered 1 or 3 to I-3B, go to I-8B; otherwise, go to I-8.**

I-4. In the past 12 months, was there a time when you did not receive welfare payments?

| | | |
|---|---|---|
| A. YES | 1 | (I-5) |
| B. NO | 2 | (I-6) |

I-5. In the past 12 months, how long were you off welfare?

________________ WEEKS

________________ MONTHS

I-6. About how long, in total, have you received welfare payments in your lifetime?

| | |
|---|---|
| A. Less than 6 months, | 1 |
| B. 6 months to one year, | 2 |
| C. More than 1 year but less than 2 years, | 3 |
| D. 2 to 3 years, or | 4 |
| E. More than 3 years? | 5 |

I-7. During the past year, did you take any classes sponsored by a program to help you get a job and get off welfare?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |

**If answered 4 to I-3B, go to I-8A. If answered 4 to I-3D, go to I-8B. If answered 4 to I-3E, go to 1-8C. If answered 4 to I-3G, go to I-8D.**

I-8. Have you ever received...

| | YES | NO |
|---|---|---|
| A. Supplemental Security Income (SSI)? | 1 | 2 |
| B. Food stamps? | 1 | 2 |
| C. WIC supplemental nutrition benefits? | 1 | 2 |
| D. Temporary Assistance to Needy Families (TANF), Aid to Families with Dependent Children (AFDC), public assistance or public welfare payments? | 1 | 2 |

**If answered 2 to I-8D go to J-1; otherwise continue.**

I-9. How long has it been since you last received welfare payments?

| | |
|---|---|
| A. More than 1 year but less than 2 years, | 1 |
| B. 2 to 3 years, or | 2 |
| C. More than 3 years? | 3 |

I-10. About how long, in total, have you received welfare payments in your lifetime?

| | |
|---|---|
| A. Less than 6 months, | 1 |
| B. 6 months to one year, | 2 |
| C. More than 1 year but less than 2 years, | 3 |
| D. 2 to 3 years, or | 4 |
| E. More than 3 years? | 5 |

I-11. Why did you stop getting welfare payments? Was it because you...

| | YES | NO |
|---|---|---|
| A. Reached the time limit set by welfare? | 1 | 2 |
| B. Were discontinued for non-compliance? | 1 | 2 |
| C. Got a job? | 1 | 2 |
| D. Got a raise and earned too much money? | 1 | 2 |
| E. Got married? | 1 | 2 |
| F. Got child support? | 1 | 2 |
| G. Received too much income from a source other than a job or child support? | 1 | 2 |
| H. Moved? | 1 | 2 |

**If answered 2 for I-11A through I-11H, go to I-12; otherwise, go to J-1.**

I-12. Was there some other reason you stopped receiving welfare? ___________________________________

## Section J. Health Questions

J-1. In general, how would you rate your overall health? Would you say it is...

| | | |
|---|---|---|
| A. | Excellent, | 1 |
| B. | Very Good, | 2 |
| C. | Good, | 3 |
| D. | Fair, or | 4 |
| E. | Poor? | 5 |

J-2. Do you have any difficulty seeing the words and letters in ordinary newspaper print even when wearing glasses or contact lenses, if you usually wear them?

| | | |
|---|---|---|
| A. | YES | 1 |
| B. | NO | 2 |

J-3. Do you have any difficulty hearing what is said in a normal conversation with another person even when using a hearing aid, if you usually wear one?

| | | |
|---|---|---|
| A. | YES | 1 |
| B. | NO | 2 |

J-4. Have you ever been diagnosed or identified as having a learning disability?

| | | |
|---|---|---|
| A. | YES | 1 |
| B. | NO | 2 |

J-5. Do you have any other health problem, impairment, or disability now that keeps you from participating fully in work, school, housework, or other activities?

| | | |
|---|---|---|
| A. | YES | 1 |
| B. | NO | 2 |

J-6. Do you have any kind of medical insurance or are you enrolled in any kind of program that helps to pay for your health care?

| | | | |
|---|---|---|---|
| A. | YES | 1 | (J-7) |
| B. | NO | 2 | (J-8) |

**If answered 1 for B-10, display "school" in J-7A.**

J-7. Is your program...

| | YES | NO |
|---|---|---|
| A. Health insurance through your work (school) or a family member's work? | 1 | 2 |
| B. Medicare (Medicare is the health insurance for people 65 or older or people with disabilities)? | 1 | 2 |
| C. Health insurance you or someone else in your family purchased directly from an insurance company or other organization that is not related to past or current employment? | 1 | 2 |
| D. Health insurance provided as part of military service? | 1 | 2 |
| E. Medicaid or [if applicable, fill in state name]? | 1 | 2 |
| F. Other? (Specify) | 1 | 2 |

State names for Medicaid:

| | |
|---|---|
| Alaska | Medical Assistance Program |
| Arizona | AHCCCS, Acute Care Program or Long Term Care System (ALTCS) |
| California | Medi-Cal |
| Connecticut | Connecticut Access (CONNECT CARD) |
| D.C. | Medical Assistance |
| Florida | MediPass |
| Georgia | Better Health Care Program or Medical Assistance |
| Hawaii | Med-QUEST, Maluhia or Medical Assistance |
| Idaho | Healthy Connections or Medical Assistance |
| Illinois | MediPlan |
| Indiana | Hoosier Healthwise |
| Iowa | MediPAS (Medical Assistance) |
| Kansas | PrimeCare, Community Care Kansas (CCK) or HealthConnect |
| Kentucky | Kentucky Patient Access and Care System (KenPAC) or Medical Assistance |
| Louisiana | CommunityCARE Program |
| Maine | PrimeCare |
| Maryland | Maryland Access to Care (MAC) or Medical Assistance |
| Massachusetts | MassHealth |
| Minnesota | Prepaid Medical Assistance Program (PMAP) or Health Care Programs |
| Mississippi | HealthMACS |
| Missouri | MC Plus |
| Montana | Passport to Health |
| Nebraska | Primary Care Plus (+) or Health Connection |
| Nevada | MAPnet |
| New Jersey | New Jersey Care 2000 |
| New Mexico | Primary Care Network |
| New York | MAX |
| North Carolina | Carolina Access |
| North Dakota | North Dakota Access to Care (No DAC) |
| Ohio | Accessing Better Care (ABC) Program |
| Oklahoma | SoonerCare |
| Oregon | Oregon Health Plan (OHP), Kaiser-S/HMO or Medical Assistance |
| Pennsylvania | HealthPASS, Family Care Network (FCN), Lancaster Community Health Plan, Blue Card or Green Card or ACCESS |
| Rhode Island | Rite Care or Medical Assistance |
| South Carolina | South Carolina Health Access Plan (SCHAP) |
| South Dakota | Primary Care Provider Program |
| Tennessee | TennCare |
| Texas | LoneSTAR (State of Texas Access Reform) |
| Vermont | Dr. Dynosaur, Vermont Health Access Program (VHAP) or AIM |
| Virginia | Medallion, Options or Medical Assistance |
| Washington | Health Access Spokane, Kaiser-S/HMO or Healthy Options |
| West Virginia | West Virginia Physician Assured Access System (PAAS) |
| Wisconsin | Medical Assistance Program |

J-8. [Ask only of people with children other than the respondent under age 18 living in the home.] Do the children living in this household have any type of medical insurance or health care coverage?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |
| C. AT LEAST ONE CHILD (BUT NOT ALL THE CHILDREN) HAS MEDICAL INSURANCE | 3 |

J-9. **[HAND CARD.]** Now I'd like to find out how you usually get information about health issues, such as diet, exercise, disease prevention, or a specific disease or health condition. How much information about health issues do you get from…

**HAND CARD 3**

Would you say…

| | A lot | Some | A little | or | None |
|---|---|---|---|---|---|
| A. Newspapers | 1 | 2 | 3 | | 4 |
| B. Magazines | 1 | 2 | 3 | | 4 |
| C. Internet | 1 | 2 | 3 | | 4 |
| D. Radio and television | 1 | 2 | 3 | | 4 |
| E. Books or brochures | 1 | 2 | 3 | | 4 |
| F. Family members, friends, or co-workers | 1 | 2 | 3 | | 4 |
| G. Talking to health care professionals, such as doctors, nurses, therapists, or psychologists | 1 | 2 | 3 | | 4 |

J-10. I would like to ask you about some topics related to maintaining health. In the past year, have you...

| | YES | NO |
|---|---|---|
| A. Gotten a flu shot? | 1 | 2 |
| B. [If female age 40 or older] Had a mammogram? | 1 | 2 |
| C. [If female between 18 and 65] Had a pap smear? | 1 | 2 |
| D. [If age 50 or older] Been screened for colon cancer? | 1 | 2 |
| E. Had your vision checked? | 1 | 2 |
| F. [If male] Been screened for prostate cancer? | 1 | 2 |
| G. [If age 50 or older] Been screened for osteoporosis? | 1 | 2 |
| H. [If age 65 or older] Had the pneumonia shot or pneumonia vaccine? | 1 | 2 |
| I. Visited a dentist? | 1 | 2 |

## Section K. Additional Demographics

K-1. **[HAND CARD.]** Which number on this card corresponds to your approximate total *personal* income for the past 12 months? Please include all your personal income, including income from your job, investments, Social Security or retirement, and welfare.

**HAND CARD 7**

| | | |
|---|---|---|
| A. | LESS THAN $5,000 | 1 |
| B. | $5,000 to $ 7,499 | 2 |
| C. | $7,500 to $9,999 | 3 |
| D. | $10,000 to $12,499 | 4 |
| E. | $12,500 to $14,999 | 5 |
| F. | $15,000 to $19,999 | 6 |
| G. | $20,000 to $29,999 | 7 |
| H. | $30,000 to $39,999 | 8 |
| I. | $40,000 to $49,999 | 9 |
| J. | $50,000 to $59,999 | 10 |
| K. | $60,000 to $74,999 | 11 |
| L. | $75,000 to $99,999 | 12 |
| M. | $100,000 OR MORE | 13 |
| N. | NO PERSONAL INCOME | 95 |

K-2. **[HAND CARD.]** Which letter on this card corresponds to your approximate *total household* income for the past 12 months? Please include all income for people living in your household, including income from jobs, investments, Social Security or retirement, and welfare. (If undergraduate college student living away from family home, please provide household income for your permanent residence.)

**HAND CARD 8**

| | | |
|---|---|---|
| A. | LESS THAN $5,000 | 1 |
| B. | $5,000 to $ 7,499 | 2 |
| C. | $7,500 to $9,999 | 3 |
| D. | $10,000 to $12,499 | 4 |
| E. | $12,500 to $14,999 | 5 |
| F. | $15,000 to $19,999 | 6 |
| G. | $20,000 to $29,999 | 7 |
| H. | $30,000 to $39,999 | 8 |
| I. | $40,000 to $49,999 | 9 |
| J. | $50,000 to $59,999 | 10 |
| K. | $60,000 to $74,999 | 11 |
| L. | $75,000 to $99,999 | 12 |
| M. | $100,000 to $149,999 | 13 |
| N. | $150,000 or more | 14 |
| O. | NO HOUSEHOLD INCOME | 95 |

**Note: Follow-up probes were asked of respondents who refused to answer K-1 and/or K-2. These probes were designed to get a broad range for the respondent's income.**

K-3. Are you Hispanic or Latino?

| | | | |
|---|---|---|---|
| A. | YES | 1 | (K-4) |
| B. | NO | 2 | (K-5) |

K-4. **[HAND CARD.]** Which of the groups on this card describes your Hispanic or Latino origin? Choose one or more.

**HAND CARD**
9

| | | |
|---|---|---|
| A. | MEXICAN, MEXICAN AMERICAN, OR CHICANO | 1 |
| B. | PUERTO RICAN OR PUERTO RICAN AMERICAN | 2 |
| C. | CUBAN OR CUBAN AMERICAN | 3 |
| D. | CENTRAL OR SOUTH AMERICAN | 4 |
| E. | OTHER HISPANIC OR LATINO BACKGROUND | 5 |

K-5. **[HAND CARD.]** Which of the groups on this card best describes you? Choose one or more.

**HAND CARD**
10

| | | |
|---|---|---|
| A. | WHITE | 1 |
| B. | BLACK OR AFRICAN AMERICAN | 2 |
| C. | ASIAN | 3 |
| D. | AMERICAN INDIAN OR ALASKA NATIVE | 4 |
| E. | NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER | 5 |

**Interviewer: Code language in which interview was conducted. 1 = English; 2 = Spanish.**

## *HAND CARD #1*

No difficulty

Some difficulty

Moderate difficulty

Great deal of difficulty

## *HAND CARD #2*

Still in high school

Less than high school (0-8 years)

Some high school (9-12 years but did not graduate)

GED or high school equivalency

High school graduate (12 years; accelerated or early graduate program)

Attended a vocational, trade, or business school after high school

College: less than two years

College: Associate's degree (A.A.)

College: two years or more, no degree

College graduate (B.S. or B.A.)

Postgraduate/No degree

Postgraduate/degree (M.S., M.A., Ph.D., M.D., etc.)

## *HAND CARD #3*

A lot

Some

A little

None

## *HAND CARD #4*

Working a full-time job for pay or profit, that is, 35 hours or more?

Working for pay or profit part-time, that is, 1 to 34 hours?

Working two or more part-time jobs for pay, totaling 35 or more hours?

Unemployed, laid off, or looking for work?

With a job but not at work because of temporary illness, vacation, or work stoppage?

With a job but on family leave (maternity or paternity leave)?

In school?

Keeping house?

Retired?

Doing volunteer work?

## *HAND CARD #5*

Every day

A few times a week

Once a week

Less than once a week

Never

## *HAND CARD #6*

A. Never married

B. Married, living with spouse

C. Married, spouse living elsewhere

D. Living as married

E. Separated or divorced

F. Widowed

## *HAND CARD #7*

A. Less than $5,000

B. $5,000 to $ 7,499

C. $7,500 to $9,999

D. $10,000 to $12,499

E. $12,500 to $14,999

F. $15,000 to $19,999

G. $20,000 to $29,999

H. $30,000 to $39,999

I. $40,000 to $49,999

J. $50,000 to $59,999

K. $60,000 to $74,999

L. $75,000 to $99,000

M. $100,000 or more

## *HAND CARD #8*

A. Less than $5,000

B. $5,000 to $ 7,499

C. $7,500 to $9,999

D. $10,000 to $12,499

E. $12,500 to $14,999

F. $15,000 to $19,999

G. $20,000 to $29,999

H. $30,000 to $39,999

I. $40,000 to $49,999

J. $50,000 to $59,999

K. $60,000 to $74,999

L. $75,000 to $99,000

M. $100,000 to $149,999

N. $150,000 or more

## *HAND CARD #9*

Mexican, Mexican American, or Chicano

Puerto Rican or Puerto Rican American

Cuban or Cuban American

Central or South American

Other Hispanic or Latino background

## *HAND CARD #10*

White

Black or African American

Asian

American Indian or Alaska Native

Native Hawaiian or other Pacific Islander

# Appendix G

# NAAL PRISON BACKGROUND QUESTIONNAIRE

## Instructions for Reading NAAL Prison Background Questionnaire

The NAAL background questionnaire is designed to be administered using a Computer Assisted Personal Interviewing (CAPI) system. Interviewers read the questions aloud to respondents from the screens of laptop computers, and they record respondents' answers directly on the computers. The interviewers read aloud response options that are written in lower-case letters, but do not read aloud the response options that are written in upper-case letters.

Administering the background questionnaire using a CAPI system allows for the inclusion of complex skip patterns that target questions only at respondents in a particular subgroup. However, these complex skip patterns make the background questionnaire difficult to follow on paper. Skip patterns are indicated on the paper version of the background questionnaire in two different ways.

For some items where a respondent's choice indicates that he or she should skip to a particular item, the skip pattern is indicated in parenthesis following the response item. For example, if a respondent answers "No" to question A-9, "Have you ever taken a class in the United States to learn English-as-a-second-language, sometimes called an ESL class?," the respondent skips to question A-11, since the question is about how long ago he or she took the class is not relevant for this respondent. If the respondent answers "Yes" to question A-9, the respondent continues to question A-9A.

When skip patterns are based on responses to one or more earlier questions in the survey, the skip patterns are indicated in a box that appears before the item that some respondents may skip. For example, question A-4 is skipped by some respondents based on their answer to A-1. The box inserted before question A-4 describes the skip pattern.

For all questions, interviewers had the option to code responses "Don't Know" or "Refused" when appropriate.

# National Study Of America's Adults

## PRISON QUESTIONNAIRE

| IF R IS NOT SCREENER R: | IF R IS SCREENER R: |
|---|---|
| Hello, I am [NAME] from Westat, a research organization in Rockville, MD. We are helping the United States Department of Education with a very important survey about how adults use printed materials. Thousands of people across the United States have already participated in the study in their homes. The United States Department of Education also wants to find out how adults in prisons use printed materials and what they read. Based on a list of all people residing in this institution, you have been selected at random to take part in a special version of this survey designed to collect information from people incarcerated in state and federal prisons. | You have been selected to participate in the survey. |

I will ask you a short set of questions about your background and education. Then, I will give you a booklet containing exercises based on printed materials, such as newspapers, maps, stories, brochures, forms, and advertisements. Others who have completed these exercises found them interesting and fun. The entire survey will take approximately 90 minutes to complete.

Your participation in this survey is voluntary and very important. All of your answers will be kept strictly confidential and will not be available to anyone here at the facility. All information will be reported for a group as a whole and your responses will not be linked to your name. You do not have to answer any questions you do not want to answer.

IF R REQUESTS MORE INFORMATION ABOUT PURPOSE OF SURVEY:
Today, adults increasingly are expected to use printed information in our society, but there is very little information available on whether or not they are well prepared. This survey will provide information about the reading and writing experiences, activities, and skills of adults in the United States. Information will be used by educators, policymakers, and business leaders to design programs in order to improve the literacy skills of adults.

According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless such collection displays a valid OMB control number. The valid OMB control number for this information collection is 1850-0654 and will expire on **06/30/04**. The time required to complete these forms is estimated to average .45 hours per respondent, including the time to review instructions and complete the survey. **If you have any comments concerning the accuracy of**

**the time estimate or suggestions for improving this form, please write to:** U.S Department of Education, Washington, DC 20202-4651. **If you have any comments or concerns regarding the status of your individual submission of this form, write directly to:** NCES, U.S. Department of Education, 1990 K Street, NW, Washington, DC 20006.

## Section A. General and Language Background

**Interviewer: Code gender of respondent. 1 = Male; 2 = Female**

A-1. In what country were you born?

Look up table ....................................................................................................

**If answered U.S.A. for A-1, go to A-3; otherwise continue.**

A-1A. How old were you when you moved to the United States?

________________
Age

A-2. How many years have you lived in the United States?

| | |
|---|---|
| A. 1 TO 5.................................................................................................... | 1 |
| B. 6 TO 10 .................................................................................................. | 2 |
| C. 11 TO 15 ................................................................................................ | 3 |
| D. 16 TO 20................................................................................................. | 4 |
| E. 21 TO 30................................................................................................. | 5 |
| F. 31 TO 40................................................................................................. | 6 |
| G. 41 TO 50................................................................................................. | 7 |
| H. 51 OR MORE ......................................................................................... | 8 |

A-3. What is your date of birth?

________________ ____________ ______________
Month Day Year

**If answered U.S.A. for A-1, go to A-5; otherwise continue.**

A-4. What was the highest level of education you completed before coming to the United States? (If response does not fit categories, probe for equivalent.)

| | |
|---|---|
| A. DID NOT ATTEND SCHOOL BEFORE COMING TO U.S. ...................... | 1 |
| B. PRIMARY (GRADES K-3)............................................................................ | 2 |
| C. ELEMENTARY (GRADES 4-8) .................................................................. | 3 |
| D. SECONDARY (GRADES 9-12) ................................................................. | 4 |
| E. VOCATIONAL (POSTSECONDARY)...................................................... | 5 |
| F. COLLEGE/UNIVERSITY ........................................................................... | 6 |
| G. OTHER (SPECIFY) .................................................................................. | 910 |

A-5. When you were growing up, what language or languages were usually spoken in your home? (Select all that apply.)

Look up table ....................................................................................................

A-6. What language or languages did you learn to speak before you started school? (Select all that apply.)

Look up table ...........................................................................................

**If English only for both A-5 and A-6, go to A-11; otherwise continue.**

A-7. What language did you first learn to read and write?

Look up table ...........................................................................................

A-8. How old were you when you learned to speak English?

| | |
|---|---|
| A. 1-4 YEARS OLD | 1 |
| B. 5-10 YEARS OLD | 2 |
| C. 11-15 YEARS OLD | 3 |
| D. 16-20 YEARS OLD | 4 |
| E. 21 YEARS OR OLDER | 5 |
| F. DOES NOT SPEAK ENGLISH | 95 |

**If English only in A-6, go to A-11; otherwise continue.**

A-9. Have you ever taken a class in the United States to learn English-as-a-second-language, sometimes called an ESL class?

| | | |
|---|---|---|
| A. YES | 1 | |
| B. NO | 2 | (A-11) |

A-9A. Where did you take the ESL class? Was it during your current or a previous incarceration in jail, prison or another correctional facility, or outside of prison? (Select all that apply.)

| | |
|---|---|
| A. CURRENT ADMISSION | 1 |
| B. PRIOR ADMISSION | 2 |
| C. OUTSIDE PRISON/CORRECTIONAL FACILITY | 3 |

A-10. How long ago did you last take an English-as-a-second-language or ESL class in the United States? Was it…

| | |
|---|---|
| A. Within the last two years, | 1 |
| B. 2 to 5 years ago, | 2 |
| C. More than 5 years ago, or | 3 |
| D. Are you taking an ESL class now? | 4 |

A-11. Which language do you usually speak now?

Look up table ...........................................................................................

A-12. What other language do you often speak now?

| | |
|---|---|
| A. Look up table | |
| B. No other language spoken | 95 |

A-13. Other than English, what language do you speak best?

A. None .......................................................... 1
B. Look up table ..........................................................

**If English only in A-6, A-11, A-12, and A-13, go to A-15; otherwise, repeat item A-14 for each non-English language identified in questions A-6, A-11, A-12, and A-13.**

A-14. With regard to **(non-English language in A-6, A-11, A-12, and A-13)**, how well do you…

Would you say…

| | Very well | Well | Not well | or | Not at all |
|---|---|---|---|---|---|
| A. Understand it when it is spoken to you? .................... | 1 | 2 | 3 | | 4 |
| B. Speak it? .............................. | 1 | 2 | 3 | | 4 |
| C. Read it? .............................. | 1 | 2 | 3 | | 4 |
| D. Write it? .............................. | 1 | 2 | 3 | | 4 |

A-15. With regard to the English language, how well do you…

Would you say…

| | Very well | Well | Not well | or | Not at all |
|---|---|---|---|---|---|
| A. Understand it when it is spoken to you? .................... | 1 | 2 | 3 | | 4 |
| B. Speak it? .............................. | 1 | 2 | 3 | | 4 |
| C. Read it? .............................. | 1 | 2 | 3 | | 4 |
| D. Write it? .............................. | 1 | 2 | 3 | | 4 |

**If English only for A-6, go to B-1; otherwise continue.**

A-16. **[HAND CARD.]** How difficult is it for you to **(item)** in English?

**HAND CARD 1**

Would you say you have...

| | No difficulty | Some difficulty | Moderate difficulty | or a | Great deal of difficulty | NEVER TRIED |
|---|---|---|---|---|---|---|
| A. Understand people having a conversation with you? ..... | 1 | 2 | 3 | | 4 | 5 |
| B. Understand television, movies, or videos? ............... | 1 | 2 | 3 | | 4 | 5 |
| C. Understand a telephone conversation? ...................... | 1 | 2 | 3 | | 4 | 5 |

**If answered 1 to A-15C and A-15D, go to B-1. If answered 1 to A-15C, but A-15D was not answered 1, go to A-18; otherwise continue.**

A-17. **[HAND CARD.]** How difficult is it for you to **(item)** written in English?

Would you say you have...

**HAND CARD 1**

| | No difficulty | Some difficulty | Moderate difficulty | or a | Great deal of difficulty | NEVER TRIED |
|---|---|---|---|---|---|---|
| A. Understand a utility bill, such as telephone or electric? ...... | 1 | 2 | 3 | | 4 | 5 |
| B. Understand the dosage information on over-the-counter medicines? .............. | 1 | 2 | 3 | | 4 | 5 |
| C. Look up information in dictionaries, encyclopedias, phone books, or other reference books?.................. | 1 | 2 | 3 | | 4 | 5 |

A-18. **[HAND CARD.]** How difficult is it for you to fill out forms in English, such as at the doctor's office or at school? Would you say you have...

**HAND CARD 1**

A. No difficulty, ........................................................................................ 1
B. Some difficulty, ..................................................................................... 2
C. Moderate difficulty, or a....................................................................... 3
D. Great deal of difficulty? ....................................................................... 4
E. NEVER TRIED ........................................................................................ 5

## Section B. Educational Background and Experiences

B-1. **[HAND CARD.]** I'd like to ask you about your educational background and experiences. What is the highest level of public or private education you completed prior to your most recent admission to prison? (If R went to school outside United States, probe for equivalent.)

**HAND CARD 2**

| | |
|---|---|
| A. LESS THAN HIGH SCHOOL (0-8 YEARS) (SPECIFY GRADE) | 2 |
| B. SOME HIGH SCHOOL (9-12 YEARS BUT DID NOT GRADUATE) (SPECIFY GRADE) | 3 |
| C. GED OR HIGH SCHOOL EQUIVALENCY | 4 |
| D. HIGH SCHOOL GRADUATE (12 YEARS; ACCELERATED OR EARLY GRADUATE PROGRAM) | 5 |
| E. ATTENDED A VOCATIONAL, TRADE, OR BUSINESS SCHOOL AFTER HIGH SCHOOL | 6 |
| F. COLLEGE: LESS THAN TWO YEARS | 7 |
| G. COLLEGE: ASSOCIATE'S DEGREE (A.A.) | 8 |
| H. COLLEGE: TWO YEARS OR MORE, NO DEGREE | 9 |
| I. COLLEGE GRADUATE (B.S. OR B.A.) | 10 |
| J. POSTGRADUATE/NO DEGREE | 11 |
| K. POSTGRADUATE/DEGREE (M.S., M.A., PH.D., M.D., ETC.) | 12 |

B-2. Since your most recent admission to prison, have you completed any additional education?

| | | |
|---|---|---|
| A. YES | 1 | |
| B. NO | 2 | (B-4) |

B-3. Are you currently enrolled in any academic classes?

| | | |
|---|---|---|
| A. YES | 1 | (B-5) |
| B. NO | 2 | |

B-4. Are you on a waiting list for an academic class or program?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |

**If answered 2 to B-2, go to B-6; otherwise continue.**

B-5. **[HAND CARD.]** Since your most recent admission to prison, what is the highest level of education you completed? (If R went to school outside U.S., probe for equivalent.)

**HAND CARD 3**

| | |
|---|---|
| A. CLASSES UP TO THE 9TH GRADE | 1 |
| B. HIGH SCHOOL CLASSES, BUT DID NOT RECEIVE A GED OR HIGH SCHOOL DEGREE | 2 |
| C. GED OR HIGH SCHOOL EQUIVALENCY | 3 |
| D. COLLEGE CLASSES, BUT DID NOT RECEIVE A DEGREE | 4 |
| E. COLLEGE: ASSOCIATE'S DEGREE (A.A.) | 5 |
| F. COLLEGE GRADUATE (B.S. OR B.A.) | 6 |
| G. POSTGRADUATE CLASSES | 7 |

**If answered 1 to B-2, then lookup from B-5; otherwise lookup from B-1.**

B-6. What year did you **(graduate from high school/receive your GED/attend your last year of school)?**

Year ___________________________________

**If answered 2, 3 or 4 to B-1, or 3 to B-5, go to B-8; otherwise continue.**

B-7. What type of high school diploma did you receive? Was it a regular high school diploma from a school in the United States; a regular high school diploma from a school outside the United States run by the United States government, such as a Department of Defense school; a regular high school diploma from a school outside the United States, not run by the United States government; a GED or high school equivalency degree; a certificate of completion that was different from a regular high school diploma; or something else?

| | |
|---|---|
| A. REGULAR HIGH SCHOOL DIPLOMA FROM A SCHOOL IN THE UNITED STATES | 1 |
| B. REGULAR HIGH SCHOOL DIPLOMA FROM A SCHOOL OUTSIDE THE UNITED STATES RUN BY THE UNITED STATES GOVERNMENT, SUCH AS A DEPARTMENT OF DEFENSE SCHOOL | 2 |
| C. REGULAR HIGH SCHOOL DIPLOMA FROM A SCHOOL OUTSIDE THE UNITED STATES, NOT RUN BY THE UNITED STATES GOVERNMENT | 3 |
| D. GED | 4 |
| E. CERTIFICATE OF COMPLETION | 5 |
| F. DID NOT RECEIVE HIGH SCHOOL DIPLOMA | 6 |
| G. OTHER (SPECIFY) | 910 |

**If answered 10, 11, 12 or DK/RF to B-1, go to box before B-9; otherwise continue.**

B-8. Now I want you to think about your education prior to prison. What was the main reason you stopped your public or private schooling when you did? Please listen to me read all the answer categories, and then tell me which one best describes the reason you stopped your schooling. Was it…

| | | |
|---|---|---|
| A. | Financial problems, | 2 |
| B. | Did not do well in school, | 3 |
| C. | Did not like school or was bored in school, | 4 |
| D. | Expelled from school or asked to leave, | 5 |
| E. | Wanted to work, | 6 |
| F. | Wanted to go into the military, | 7 |
| G. | Personal illness, disability, or pregnancy, | 8 |
| H. | Family reasons such as the illness or death of one of your parents, | 9 |
| I. | School not available or not accessible, | 10 |
| J. | Did not feel safe in school, | 11 |
| K. | Sent to jail, prison, or detention center, or | 13 |
| L. | Other? | 14 |

**If answered 2-6 or DK/RF to B-1, go to B-10; otherwise continue. If answered 7 or 9 to B-1, read "attend your last year of college" for B-9. If answered 11 or 12 for B-1, read "receive your undergraduate degree" for B-9. If answered 8 for B-1, read "receive your associate's degree" for B-9. If answered 10 to B-1, read "graduate from college" for B-9.**

B-9. What year did you **(attend your last year of college/receive your undergraduate degree/receive your associate's degree/graduate from college)?**

Year ___________________________________

B-10. Have you ever taken part in a program other than in regular school in order to improve your *basic skills*, that is, basic reading, writing, and arithmetic skills? The program may have been in prison or it may have been outside of prison.

| | | | |
|---|---|---|---|
| A. | YES | 1 | |
| B. | NO | 2 | (B-14) |

B-11. Where did you take the basic skills class? Was it during your current or a previous incarceration in jail, prison or another correctional facility, or outside of prison? (Select all that apply.)

| | | |
|---|---|---|
| A. | CURRENT ADMISSION | 1 |
| B. | PRIOR ADMISSION | 2 |
| C. | OUTSIDE PRISON/CORRECTIONAL FACILITY | 3 |

B-12. How long ago did you last take a class to improve your basic skills. Was it…

| | |
|---|---|
| A. Within the last two years, | 1 |
| B. 2 to 5 years ago, | 2 |
| C. More than 5 years ago, or | 3 |
| D. Are you currently taking a basic skills class? | 4 |

**If answered 1 to A-9 and B-10, go to B-13; otherwise go to B-14.**

B-13. Was the basic skills class part of the English-as-a-second-language or ESL class you took, or was it a separate class?

| | |
|---|---|
| A. PART OF ESL CLASS | 1 |
| B. SEPARATE CLASS | 2 |

B-14. Have you received any type of information technology skill certification sponsored by a hardware or software manufacturer or an industry or professional association?

| | | |
|---|---|---|
| A. YES | 1 | |
| B. NO | 2 | (B-17) |

B-15. Did you have to pass a test to get the certification?

| | | |
|---|---|---|
| A. YES | 1 | |
| B. NO | 2 | (B-17) |

B-16. How did you prepare for the test? Was it...

| | YES | NO |
|---|---|---|
| A. A class offered in prison, jail or other correctional facility? | 1 | 2 |
| B. A class offered by a four-year college or university? | 1 | 2 |
| C. A class offered by a community college? | 1 | 2 |
| D. A class offered by a technical school or private vendor? | 1 | 2 |
| E. A class offered by a high school or vocational secondary school? | 1 | 2 |
| F. A class offered directly by a hardware or software manufacturer, such as Microsoft, Oracle, Novell, or Cisco? | 1 | 2 |
| G. On the job training or apprenticeship? | 1 | 2 |
| H. Independent study? | 1 | 2 |
| I. Other? (Specify) | 1 | 2 |

B-17. Other than information technology, have you ever received any type of job-related skill certification recognized by a licensing board or an industry or professional association?

| | | |
|---|---|---|
| A. YES | 1 | |
| B. NO | 2 | (C-1) |

B-18. Did you have to pass a test to get the certification?

| | | |
|---|---|---|
| A. YES | 1 | |
| B. NO | 2 | (C-1) |

B-19. How did you prepare for the test? Was it...

| | YES | NO |
|---|---|---|
| A. A class offered in prison, jail or other correctional facility? | 1 | 2 |
| B. A class offered by a four-year college or university? | 1 | 2 |
| C. A class offered by a community college? | 1 | 2 |
| D. A class offered by a technical school or private vendor? | 1 | 2 |
| E. A class offered by a high school or vocational secondary school? | 1 | 2 |
| F. On the job training or apprenticeship? | 1 | 2 |
| G. Independent study? | 1 | 2 |
| H. Other? (Specify) | 1 | 2 |

## Section C. Prison Experiences

C-1. Now I'd like to ask you a few questions about your prison experiences. Since your most recent admission to prison, have you participated in…

| | YES | NO |
|---|---|---|
| A. Employment counseling (including how to find a job or interviewing skills)? | 1 | 2 |
| B. Classes in parenting or child rearing skills? | 1 | 2 |
| C. Classes in life skills and community adjustment including anger management, conflict resolution, personal finance, or other life skills? | 1 | 2 |
| D. Drug or alcohol groups (for example, Alcoholics Anonymous, Al-Anon, Narcotics Anonymous, or other drug or alcohol-related groups)? | 1 | 2 |
| E. Inmate assistance groups, for example, inmate liaison, advisory, or workers' councils? | 1 | 2 |
| F. A religious study group? | 1 | 2 |
| G. An ethnic/racial organization (for example, NAACP, African-American or Black Culture group, Hispanic Committee, Aztlan, or Lakota? | 1 | 2 |
| H. Any other programs or groups? (Specify) | 1 | 2 |

C-2. Since your most recent admission to prison, have you been a student in a vocational training program, excluding prison work assignments?

| | | |
|---|---|---|
| A. YES | 1 | |
| B. NO | 2 | (C-6) |

C-3. Are you currently a student in a vocational training program?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |

C-4. Since your most recent admission to prison, how long have you spent altogether in vocational training?

| | |
|---|---|
| A. LESS THAN ONE MONTH | 1 |
| B. NUMBER OF MONTHS (SPECIFY) | 2 |
| C. NUMBER OF YEARS (SPECIFY) | 3 |

C-5. Did this training or education include instruction intended to:

| | YES | NO |
|---|---|---|
| A. Improve your English reading skills? | 1 | 2 |
| B. Improve your English writing skills? | 1 | 2 |
| C. Improve your arithmetic or mathematics skills? | 1 | 2 |
| D. Improve your computer skills? | 1 | 2 |
| E. Help you communicate or work better with other people? | 1 | 2 |

**If answered 1 to C-3, go to C-7; otherwise continue.**

C-6. Are you on a waiting list for any vocational training programs?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |

C-7. Since last **(current day),** how many hours did you spend in any type of class as a student?

HOURS: ______

C-8. For what offenses are you *now* in prison? (PROBE: Any others?) (Record each offense on a separate line.)

OFFENSE #1: ______________________________
OFFENSE #2: ______________________________
OFFENSE #3: ______________________________
OFFENSE #4: ______________________________
OFFENSE #5: ______________________________

**If only one offense reported in C-8, go to C-10; otherwise continue.**

C-9. For which of these offenses did you receive the longest sentence?

| | |
|---|---|
| A. Look up responses | |
| B. Hasn't been sentenced | 95 |

C-10. In what month and year were you admitted to prison most recently?

______________________________
MONTH / YEAR

C-11. Before your most recent admission to prison, did you ever serve time in prison, jail, or some other correctional facility as a juvenile or adult?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |

C-12. Have you ever been placed on probation, either as a juvenile or as an adult?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |

**If answered 95 to C-9, go to D-1; otherwise continue.**

C-13. Do you have a definite date on which you expect to be released from prison?

| | | |
|---|---|---|
| A. YES | 1 | |
| B. NO | 2 | (C-15) |

C-14. In what month and year will you be released?

______________________________ (D-1)
MONTH / YEAR

C-15. Do you expect to ever be released from prison?

| | | |
|---|---|---|
| A. YES ........................................ | 1 | |
| B. NO ........................................ | 2 | (D-1) |

C-16. In what month and year is your earliest possible release date?

________________________________________

MONTH / YEAR

## Section D. Prison Work Assignments and Labor Force Participation

D-1. The next question is about jobs you've had since your most recent admission to prison on **(date from C-10).** Do you currently have any work assignments?

| | | |
|---|---|---|
| A. YES ........................................ | 1 | |
| B. NO ........................................ | 2 | (D-5) |

D-2. **[HAND CARD.]** How often do you read as part of your current job(s) in prison? Would you say…

**HAND CARD 6**

| | |
|---|---|
| A. Every day, ........................................ | 1 |
| B. A few times a week, ........................................ | 2 |
| C. Once a week, ........................................ | 3 |
| D. Less than once a week, or ........................................ | 4 |
| E. Never? ........................................ | 5 |

D-3. **[HAND CARD.]** How often do you write as part of your current job(s) in prison? Would you say…

**HAND CARD 6**

| | |
|---|---|
| A. Every day, ........................................ | 1 |
| B. A few times a week, ........................................ | 2 |
| C. Once a week, ........................................ | 3 |
| D. Less than once a week, or ........................................ | 4 |
| E. Never? ........................................ | 5 |

D-4. In the last week, how many hours did you work at your job(s)?

HOURS: __________

D-5. In the year before your incarceration on **(date from C-10),** did you receive income from any of the following? (Code all that apply.)

| | YES | NO |
|---|---|---|
| A. Pay from jobs or wages? ........................................ | 1 | 2 |
| B. Educational scholarship or grant? ........................................ | 1 | 2 |
| C. Family or friends? ........................................ | 1 | 2 |
| D. Unemployment insurance compensation and/or workman's compensation? ........................................ | 1 | 2 |
| E. Social Security or other pensions (e.g., veterans, Supplemental Security Income, disability)? ........................................ | 1 | 2 |
| F. Welfare, charity? ........................................ | 1 | 2 |
| G. Illegal sources? ........................................ | 1 | 2 |
| H. Anything else? (Specify) ........................................ | 1 | 2 |

**If date of admission from C-10 is earlier than January 2001, go to box before D-13; otherwise continue.**

D-6. In the past three years, that is, since January 2001, did you work for pay or profit, either full-time or part-time, while not serving time in prison?

| | | |
|---|---|---|
| A. YES ............ | 1 | |
| B. NO ............ | 2 | (Box before D-13) |

D-7. Between January 2001 and **(date from C-10),** how many months did you work?

| | | |
|---|---|---|
| A. LESS THAN 1 MONTH............ | 1 | (Box before D-13) |
| B. NUMBER OF MONTHS (SPECIFY)............ | 2 | |

D-8. Now I would like to ask you some questions about your most recent full-time or part-time job prior to your most recent admission to prison. For what kind of business or industry did you work? (For example, television and radio manufacturing, retail shoe store, state labor department, farm, etc.) (If R worked two or more jobs, probe: Tell me about the job you worked the most hours or the job you considered your primary employment.)

BUSINESS OR INDUSTRY: ______________________________

______________________________

______________________________

______________________________

D-9. What was your occupation, that is, what was your job called? (For example, electrical engineer, stock clerk, typist, farmer, etc.)

OCCUPATION: ______________________________

______________________________

______________________________

______________________________

D-10. What were the most important activities or duties at this job? (For example, typing, keeping account books, filing, selling cars, operating a printing press, finishing concrete, etc.)

ACTIVITIES OR DUTIES: ______________________________

______________________________

______________________________

______________________________

______________________________

______________________________

D-11. Was this full-time, part-time, or occasional work?

| | |
|---|---|
| A. FULL-TIME ............ | 1 |
| B. PART-TIME ............ | 2 |
| C. OCCASIONAL ............ | 3 |

D-12. Were you employed by government, by a **PRIVATE** employer, or were you self-employed or working in a family business?

| | |
|---|---|
| A. GOVERNMENT ............ | 1 |
| B. PRIVATE EMPLOYER ............ | 2 |
| C. SELF-EMPLOYED ............ | 3 |
| D. WORKING IN A FAMILY BUSINESS ............ | 4 |

**If age 66 or older (based on A-3), skip to E-1; otherwise continue.**

D-13. How much do you think your reading skills limit your job opportunities – for example, to get a promotion or a (different) job you would like to have? Would you say a lot, some, a little, or not at all?

| | |
|---|---|
| A. A LOT | 1 |
| B. SOME | 2 |
| C. A LITTLE | 3 |
| D. NOT AT ALL | 4 |

D-14. How much do you think your writing skills limit your job opportunities – for example, to get a promotion or a (different) job you would like to have? Would you say a lot, some, a little, or not at all?

| | |
|---|---|
| A. A LOT | 1 |
| B. SOME | 2 |
| C. A LITTLE | 3 |
| D. NOT AT ALL | 4 |

D-15. How much do you think your math skills limit your job opportunities – for example, to get a promotion or a (different) job you would like to have? Would you say a lot, some, a little, or not at all?

| | |
|---|---|
| A. A LOT | 1 |
| B. SOME | 2 |
| C. A LITTLE | 3 |
| D. NOT AT ALL | 4 |

D-16. How much do you think your computer skills limit your job opportunities – for example, to get a promotion or a (different) job you would like to have? Would you say a lot, some, a little, or not at all?

| | |
|---|---|
| A. A LOT | 1 |
| B. SOME | 2 |
| C. A LITTLE | 3 |
| D. NOT AT ALL | 4 |

## Section E: Political and Social Participation

E-1. **[HAND CARD.]** I'd like to find out how you usually get information about current events, public affairs, and the government. How much information about current events, public affairs, and the government do you get from …

**HAND CARD 4**

Would you say…

| | A lot | Some | A little | or | None |
|---|---|---|---|---|---|
| A. Newspapers? | 1 | 2 | 3 | | 4 |
| B. Magazines? | 1 | 2 | 3 | | 4 |
| C. Radio and television? | 1 | 2 | 3 | | 4 |
| D. Books or brochures? | 1 | 2 | 3 | | 4 |
| E. Family members, friends, other inmates or staff? | 1 | 2 | 3 | | 4 |

**If English only for A-6, go to E-3; otherwise continue.**

E-2. How much of the information you get about current events, public affairs, and the government is presented in **(non-English language in A-6, A-11, A-12 or A-13)?** Would you say…

| | | |
|---|---|---|
| A. | All, | 1 |
| B. | Most, | 2 |
| C. | Some, or | 3 |
| D. | None? | 4 |

E-3. How many hours do you usually watch television, videotapes, or DVDs each day?

| | | |
|---|---|---|
| A. | NONE | 0 |
| B. | 1 HOUR OR LESS | 1 |
| C. | 2 HOURS | 2 |
| D. | 3 HOURS | 3 |
| E. | 4 HOURS | 4 |
| F. | 5 HOURS | 5 |
| G. | 6 OR MORE HOURS | 6 |

E-4. How often do you use the services of a library for any reason? Would you say…

| | | | |
|---|---|---|---|
| A. | Daily, | 1 | (E-8) |
| B. | Weekly, | 2 | (E-7) |
| C. | Monthly, | 3 | (E-5) |
| D. | Once or twice a year, or | 4 | (E-5) |
| E. | Never? | 5 | (E-6) |

E-5. During the past month, have you used or ever wanted to use the library?

| | | | |
|---|---|---|---|
| A. | YES | 1 | (E-7) |
| B. | NO | 2 | (Box before E-9) |

E-6. During the past month, have you ever wanted to use the library?

| | | |
|---|---|---|
| A. YES | 1 | (E-7) |
| B. NO | 2 | (Box before E-9) |

E-7. **[HAND CARD.]** How long does it usually take you to get access to the prison library when you want to use it? Would you say...

**HAND CARD 10**

| | |
|---|---|
| A. Less than two days, | 1 |
| B. Two to six days, | 2 |
| C. Seven to 10 days, or | 3 |
| D. More than 10 days? | 4 |

**If answered 5 for E-4, go to box before E-9; otherwise continue.**

E-8. During the past month, did you borrow any materials from a library?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |

**If answered (born in U.S.A.) or (U.S. territory) to A-1, go to E-10; otherwise continue.**

E-9. Are you a citizen of the United States?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |

E-10. Did you ever serve on active duty in the U.S. Armed Forces?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |

**If answered 2 to E-9, go to F-1; otherwise continue.**

**If in prison for current offense in November 2000 (based on C-10), go to F-1; otherwise continue.**

E-11. In 2000, Al Gore ran on the Democratic ticket against George W. Bush for the Republicans. Do you remember for sure whether or not you voted in that election?

| | | |
|---|---|---|
| A. YES, REMEMBER FOR SURE | 1 | |
| B. NO, DON'T REMEMBER FOR SURE | 2 | (F-1) |
| C. YES, VOTED | 3 | (F-1) |
| D. NO, DIDN'T VOTE | 4 | (F-1) |

E-12. Did you vote in that election?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |

## Section F: Literacy Practices

F-1. **[HAND CARD.]** Now I'd like to talk to you about what you read in English. How often do you read **(item)** in English?

**HAND CARD 6**

Would you say...

| | Every day | A few times a week | Once a week | Less than once a week | or | Never |
|---|---|---|---|---|---|---|
| A. Newspapers or magazines ... | 1 | 2 | 3 | 4 | | 5 |
| B. Books .................................. | 1 | 2 | 3 | 4 | | 5 |
| C. Letters and notes.................. | 1 | 2 | 3 | 4 | | 5 |

**If English only in A-6, go to F-3; otherwise continue.**

F-2. **[HAND CARD.]** How often do you read **(item)** in **(non-English language in A-6, A-11, A-12 or A-13)?**

**HAND CARD 6**

Would you say...

| | Every day | A few times a week | Once a week | Less than once a week | or | Never |
|---|---|---|---|---|---|---|
| A. Newspapers or magazines ... | 1 | 2 | 3 | 4 | | 5 |
| B. Books .................................. | 1 | 2 | 3 | 4 | | 5 |
| C. Letters and notes.................. | 1 | 2 | 3 | 4 | | 5 |

F-3. Do you ever use a computer?

| | | |
|---|---|---|
| A. YES ............................................................................................ | 1 | |
| B. NO .............................................................................................. | 2 | (F-5) |

F-4. **[HAND CARD.]** Now I'd like to ask you about how you use the computer. How often do you **(item)?**

**HAND CARD 6**

Would you say...

| | Every day | A few times a week | Once a week | Less than once a week | or | Never |
|---|---|---|---|---|---|---|
| A. Write using a word processing program?............................ | 1 | 2 | 3 | 4 | | 5 |
| B. Use a spreadsheet program?............................ | 1 | 2 | 3 | 4 | | 5 |
| C. Look up information on a CD-ROM?................................. | 1 | 2 | 3 | 4 | | 5 |

F-5. **[HAND CARD.]** How much help do you get from family members, friends, other inmates, or staff with…

**HAND CARD 4**

Would you say…

| | A lot | Some | A little | or | None |
|---|---|---|---|---|---|
| A. Filling out forms? .................. | 1 | 2 | 3 | | 4 |
| B. Reading or explaining newspaper articles or other written information? ............. | 1 | 2 | 3 | | 4 |
| C. Writing notes and letters? ..... | 1 | 2 | 3 | | 4 |
| D. Using basic arithmetic, that is, adding, subtracting, multiplying, or dividing, such as filling out order forms or balancing a checkbook? ....... | 1 | 2 | 3 | | 4 |

## Section G: Demographic Information

Now I'm going to ask you some questions about your family.

G-1. In what country was your mother (stepmother or female guardian) born?

Look up table ..........................................................................................

G-2. What was the highest level of education your mother (stepmother or female guardian) completed? (If went to school outside U.S., probe for equivalent.)

| | |
|---|---|
| A. LESS THAN HIGH SCHOOL (0-8 YEARS) (SPECIFY GRADE) ............ | 1 |
| B. SOME HIGH SCHOOL (9-12 YEARS BUT DID NOT GRADUATE) (SPECIFY GRADE) ........................................................................ | 2 |
| C. GED OR HIGH SCHOOL EQUIVALENCY ............................................ | 3 |
| D. HIGH SCHOOL GRADUATE (12 YEARS; ACCELERATED OR EARLY GRADUATE PROGRAM) ....................................................................... | 4 |
| E. ATTENDED VOCATIONAL, TRADE, OR BUSINESS SCHOOL AFTER HIGH SCHOOL ..................................................................................... | 5 |
| F. COLLEGE: LESS THAN TWO YEARS .................................................. | 6 |
| G. COLLEGE: ASSOCIATE'S DEGREE (A.A.)........................................... | 7 |
| H. COLLEGE: TWO YEARS OR MORE, NO DEGREE.............................. | 8 |
| I. COLLEGE GRADUATE (B.S. OR B.A.) ................................................ | 9 |
| J. POSTGRADUATE/NO DEGREE .......................................................... | 10 |
| K. POSTGRADUATE/DEGREE (M.S., M.A., Ph.D., M.D., ETC.) ................ | 11 |

G-3. In what country was your father (stepfather or male guardian) born?

Look up table ..........................................................................................

G-4. What was the highest level of education your father (stepfather or male guardian) completed? (If went to school outside United States, probe for equivalent.)

| | |
|---|---|
| A. LESS THAN HIGH SCHOOL (0-8 YEARS) (SPECIFY GRADE) ............ | 1 |
| B. SOME HIGH SCHOOL (9-12 YEARS BUT DID NOT GRADUATE) (SPECIFY GRADE) ........................................................................ | 2 |
| C. GED OR HIGH SCHOOL EQUIVALENCY ............................................ | 3 |
| D. HIGH SCHOOL GRADUATE (12 YEARS; ACCELERATED OR EARLY GRADUATE PROGRAM) ....................................................................... | 4 |
| E. ATTENDED VOCATIONAL, TRADE, OR BUSINESS SCHOOL AFTER HIGH SCHOOL ..................................................................................... | 5 |
| F. COLLEGE: LESS THAN TWO YEARS .................................................. | 6 |
| G. COLLEGE: ASSOCIATE'S DEGREE (A.A.)........................................... | 7 |
| H. COLLEGE: TWO YEARS OR MORE, NO DEGREE.............................. | 8 |
| I. COLLEGE GRADUATE (B.S. OR B.A.) ................................................ | 9 |
| J. POSTGRADUATE/NO DEGREE .......................................................... | 10 |
| K. POSTGRADUATE/DEGREE (M.S., M.A., Ph.D., M.D., ETC.) ................ | 11 |

## Section H: Household Income and Welfare Participation

H-1. **[HAND CARD.]** Which number on this card describes your current marital status?

**HAND CARD 7**

| | |
|---|---|
| A. NEVER MARRIED | 1 |
| B. MARRIED | 2 |
| C. LIVING AS MARRIED | 3 |
| D. SEPARATED OR DIVORCED | 4 |
| E. WIDOWED | 5 |

H-2. Have you ever received…

| | YES | NO |
|---|---|---|
| A. Supplemental Security Income (SSI)? | 1 | 2 |
| B. Food stamps? | 1 | 2 |
| C. [If female] WIC supplemental nutrition benefits? | 1 | 2 |
| D. Temporary Assistance for Needy Families (TANF), Aid to Families with Dependent Children (AFDC), public assistance or public welfare payments? | 1 | 2 |
| E. Rent subsidy, such as Section 8 or public housing? | 1 | 2 |

**If answered 2, DK or RF to H-2D, go to I-1; otherwise continue.**

H-3. How long has it been since you last received welfare payments? (Read responses and stop when R indicates that you are at correct response.)

| | |
|---|---|
| A. 1 year or less, | 1 |
| B. More than 1 year but less than 2 years, | 2 |
| C. 2 to 3 years, or | 3 |
| D. More than 3 years? | 4 |

H-4. About how long, in total, have you received welfare payments in your lifetime? (Read responses and stop when R indicates that you are at correct response.)

| | |
|---|---|
| A. Less than 6 months, | 1 |
| B. 6 months to one year, | 2 |
| C. More than 1 year but less than 2 years, | 3 |
| D. 2 to 3 years, or | 4 |
| E. More than 3 years? | 5 |

H-5. Why did you stop getting welfare payments? Was it because you…

| | YES | NO |
|---|---|---|
| A. Were incarcerated? | 1 | 2 |
| B. Reached the time limit set by welfare? | 1 | 2 |
| C. Were discontinued for non-compliance? | 1 | 2 |
| D. Got a job? | 1 | 2 |
| E. Got a raise and earned too much money? | 1 | 2 |
| F. Got married? | 1 | 2 |
| G. Got child support? | 1 | 2 |
| H. Received too much income from a source other than a job or child support? | 1 | 2 |
| I. Moved? | 1 | 2 |

**If answered 2 for H-5A through H-5I, go to H-6; otherwise go to I-1.**

H-6. Was there some other reason you stopped receiving welfare? ___________________________

## Section I: Health Questions

I-1. In general, how would you rate your overall health? Would you say it is…

| | |
|---|---|
| A. Excellent, | 1 |
| B. Very good, | 2 |
| C. Good, | 3 |
| D. Fair, or | 4 |
| E. Poor? | 5 |

I-2. Do you have any difficulty seeing the words and letters in ordinary newspaper print even when wearing glasses or contact lenses, if you usually wear them?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |

I-3. Do you have any difficulty hearing what is said in a normal conversation with another person even when using a hearing aid, if you usually wear one?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |

I-4. Have you ever been diagnosed or identified as having a learning disability?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |

I-5. Do you have any other health problem, impairment, or disability now that keeps you from participating fully in work, school, or other activities?

| | |
|---|---|
| A. YES | 1 |
| B. NO | 2 |

I-6. **[HAND CARD.]** Now I'd like to find out how you usually get information about health issues, such as diet, exercise, disease prevention, or a specific disease or health condition. How much information about health issues do you get from…

**HAND CARD 4**

Would you say…

| | A lot | Some | A little | or | None |
|---|---|---|---|---|---|
| A. Newspapers? | 1 | 2 | 3 | | 4 |
| B. Magazines? | 1 | 2 | 3 | | 4 |
| C. Radio and television? | 1 | 2 | 3 | | 4 |
| D. Books or brochures? | 1 | 2 | 3 | | 4 |
| E. Talking to health care professionals, such as doctors, nurses, therapists, or psychologists? | 1 | 2 | 3 | | 4 |
| F. Family members, friends, other inmates or staff? | 1 | 2 | 3 | | 4 |

## Section J: Additional Demographics

J-1. Are you Hispanic or Latino?

| | | |
|---|---|---|
| A. YES ........................................ | 1 | |
| B. NO ........................................ | 2 | (J-3) |

J-2. **[HAND CARD.]** Which of the groups on this card describes your Hispanic or Latino origin? Choose one or more.

**HAND CARD 8**

| | |
|---|---|
| A. MEXICAN, MEXICAN AMERICAN, OR CHICANO .................. | 1 |
| B. PUERTO RICAN OR PUERTO RICAN AMERICAN .................. | 2 |
| C. CUBAN OR CUBAN AMERICAN .................. | 3 |
| D. CENTRAL OR SOUTH AMERICAN.................. | 4 |
| E. OTHER HISPANIC OR LATINO BACKGROUND .................. | 5 |

J-3. **[HAND CARD.]** Which of the groups on this card best describes you? Choose one or more.

**HAND CARD 9**

| | |
|---|---|
| A. WHITE .................. | 1 |
| B. BLACK OR AFRICAN AMERICAN.................. | 2 |
| C. ASIAN.................. | 3 |
| D. AMERICAN INDIAN OR ALASKA NATIVE .................. | 4 |
| E. NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER .................. | 5 |

**Interview: Code language in which interview was conducted. 1 = English; 2 = Spanish.**

## *HAND CARD #1*

No difficulty

Some difficulty

Moderate difficulty

Great deal of difficulty

## *HAND CARD #2*

Less than high school (0-8 years)

Some high school (9-12 years but did not graduate)

GED or high school equivalency

High school graduate (12 years; accelerated or early graduate program)

Attended a vocational, trade, or business school after high school

College: less than 2 years

College: Associate's degree (A.A.)

College: two years or more, no degree

College graduate (B.S. or B.A.)

Postgraduate/no degree

## *HAND CARD #3*

Classes up to the 9th grade

High School classes, but did not receive a GED or high school

GED or high school equivalency

College classes, but did not receive a degree

College: Associate's Degree (A.A.)

College graduate (B.S. or B.A.)

Postgraduate classes

## *HAND CARD #4*

A lot

Some

A little

None

## *HAND CARD #5*

Social Security or other pensions (e.g., veterans, Supplemental Security Income, disability)

Welfare, charity

Unemployment insurance compensation / workman's compensation

Family or friends

Illegal sources

Educational scholarship or grant

## *HAND CARD #6*

Every day

A few times a week

Once a week

Less than once a week

Never

## *HAND CARD #7*

Never married

Married

Living as married

Separated or divorced

Widowed

## *HAND CARD #8*

Mexican, Mexican American or Chicano

Puerto Rican or Puerto Rican American

Cuban or Cuban American

Central or South American

Other Hispanic or Latino background

## *HAND CARD #9*

White

Black or African American

Asian

American Indian or Alaska Native

Native Hawaiian or other Pacific Islander

## *HAND CARD #10*

Less Than Two Days

Two to Six Days

Seven to 10 Days

More Than 10 Days