NCES 2009-476

U. S. DEPARTMENT OF EDUCATION

Technical Report and Data File User's Manual

For the 2003 National Assessment of Adult Literacy

ies 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

CHAPTER 1: THE 2003 NATIONAL ASSESSMENT OF ADULT LITERACY: AN OVERVIEW 1-1
1.1 INTRODUCTION 1-1
1.2 DEFINING LITERACY 1-2
1.2.1 Prose, Document, and Quantitative Literacy 1-2
1.2.2 Establishing Literacy Levels 1-3
1.3 COMPONENTS OF THE NATIONAL ASSESSMENT OF ADULT LITERACY 1-5
1.3.1 Background Questionnaire 1-5
1.3.2 Main Household Assessment 1-5
1.3.3 Prison Survey 1-6
1.3.4 Fluency Addition to the NAAL (FAN) 1-6
1.3.5 Adult Literacy Supplemental Assessment (ALSA) 1-7
1.3.6 Health Literacy Component 1-8
1.4 CONDUCTING THE SURVEY 1-9
1.4.1 Field Tests 1-9
1.4.2 Data Collection 1-9
1.5 SAMPLE DESIGN 1-10
1.5.1 Household Sample 1-11
1.5.2 Prison Sample 1-11
1.6 REDUCING THE RISK OF DATA DISCLOSURE 1-11
1.7 RESPONSE RATES AND NONRESPONSE BIAS ANALYSIS 1-12
1.8 WEIGHTING AND VARIANCE ESTIMATION 1-12
1.9 SCORING 1-12
1.10 ITEM ANALYSIS, SCALING, AND ESTIMATES OF SUBPOPULATION PROFICIENCIES 1-13
1.11 THE LITERACY OF ADULTS WITHOUT COGNITIVE DATA 1-13
1.12 VARIABLE CONSTRUCTION AND FILE DEVELOPMENT 1-13

CHAPTER 2: DEVELOPMENT OF THE SURVEY INSTRUMENTS 2-1
2.1 CONCEPTUAL FRAMEWORK OF THE LITERACY ASSESSMENT 2-1
2.2 HOUSEHOLD BACKGROUND QUESTIONNAIRE 2-3
2.2.1 Development Procedures 2-4
2.2.2 Cognitive Laboratory Analyses 2-6
2.2.3 Key Constructs and Variables 2-7
2.2.3.1 General and Language Background 2-8
2.2.3.2 Educational Background and Experiences 2-8
2.2.3.3 Political and Social Participation 2-9
2.2.3.4 Labor Force Participation 2-10
2.2.3.5 Literacy Practices 2-10
2.2.3.6 Job Training and Skills 2-11
2.2.3.7 Demographic Information 2-11
2.2.3.8 Family Literacy 2-11
2.2.3.9 Household Income and Welfare Participation 2-12
2.2.3.10 Health 2-12
2.2.3.11 Additional Demographics 2-13
2.2.4 Spanish Version 2-13

CONTENTS—Continued

2.3 PRISON BACKGROUND QUESTIONNAIRE 2-13
2.3.1 Development Procedures 2-14
2.3.2 Cognitive Laboratory Analyses 2-14
2.3.3 Key Constructs and Variables Different From Household Questionnaire 2-15
2.3.3.1 General and Language Background 2-16
2.3.3.2 Educational Background and Experiences 2-16
2.3.3.3 Prison Experiences 2-16
2.3.3.4 Prison Work Assignments and Labor Force Participation 2-16
2.3.3.5 Political and Social Participation 2-17
2.3.3.6 Literacy Practices 2-17
2.3.3.7 Demographic Information 2-17
2.3.3.8 Household Income and Welfare Participation 2-17
2.3.3.9 Health 2-18
2.3.3.10 Additional Demographics 2-18
2.4 2003 NATIONAL ASSESSMENT OF ADULT LITERACY COGNITIVE ITEMS 2-18
2.4.1 The Prose, Document, and Quantitative Scales 2-19
2.4.2 Development of Items and Scoring Rubrics 2-20
2.4.2.1 Development of Scoring Rubrics 2-22
2.4.3 Health Literacy 2-26
2.4.3.1 Definition of Health Literacy and Purpose of the Health Literacy Assessment 2-27
2.4.3.2 Health Item Development 2-28
2.4.3.3 Types of Health Literacy Tasks 2-28
2.4.4 Cognitive Laboratory Analyses 2-30
2.5 FLUENCY ADDITION TO THE 2003 NATIONAL ASSESSMENT OF ADULT LITERACY 2-33
2.5.1 Purpose 2-33
2.5.2 Development Procedures 2-33
2.5.2.1 Word Lists 2-34
2.5.2.2 Pseudoword Lists 2-35
2.5.2.3 Passages 2-37
2.5.3 Cognitive Laboratory Procedures 2-38
2.5.4 Automatic Scoring of FAN Lists and Passages 2-39
2.6 ADULT LITERACY SUPPLEMENTAL ASSESSMENT 2-40
2.6.1 Purpose 2-42
2.6.2 Development Procedures 2-42
2.6.3 Cognitive Laboratory Procedures 2-43
2.7 BIAS AND SENSITIVITY REVIEW OF NAAL ITEMS 2-43
2.7.1 Expert Panels 2-44
2.7.2 Identifying Biased and Sensitive Items 2-45
2.8 FIELD-TEST BOOKLET DESIGN 2-45
2.9 MAIN ASSESSMENT BOOKLET DESIGN 2-46
2.9.1 2003 National Assessment of Adult Literacy Cognitive Items 2-46
2.9.2 2003 National Assessment of Adult Literacy Fluency Items 2-50

CHAPTER 3: FIELD TEST 3-1
3.1 INTRODUCTION 3-1

CONTENTS—Continued

3.2 SAMPLE DESIGN 3-3
3.2.1 Target Population 3-3
3.2.2 Target Sample Sizes 3-3
3.2.3 Household Sample Design 3-6
3.2.3.1 Primary Sampling Unit and Segment Sample 3-6
3.2.3.2 Sample of Households Within Segments 3-6
3.2.3.3 Sample of Eligible Persons Within Selected Households 3-6
3.2.4 Volunteer Sample Design 3-7
3.3 DATA COLLECTION 3-7
3.3.1 Data Collection Instruments 3-7
3.3.1.1 Screener 3-8
3.3.1.2 Background Questionnaire 3-9
3.3.1.3 Core and Main Assessments 3-9
3.3.1.4 Language of Administration 3-11
3.3.2 General Features of the Computer-Assisted Interviewing Systems 3-12
3.3.2.1 Interviewer Management System 3-12
3.3.2.2 Supervisor Management System 3-12
3.3.3 Interviewer Materials 3-13
3.3.3.1 Advance Materials 3-13
3.3.3.2 Household Folder 3-14
3.3.3.3 Handcards 3-14
3.3.4 Field Staff Recruitment and Training 3-14
3.3.4.1 Field Staff Recruitment 3-14
3.3.4.2 Approach to Training 3-15
3.3.4.3 Supervisor Training 3-16
3.3.4.4 Interviewer Training 3-16
3.3.4.5 General Interviewer Techniques and CAPI Train 3-16
3.3.5 Conduct of the Data Collection Effort 3-17
3.3.5.1 Field Staff and Organization 3-17
3.3.5.2 Management 3-17
3.3.5.3 Data Collection Monitoring With Computer-Generated Reports 3-17
3.3.6 Data Collection Procedures 3-17
3.3.6.1 Contact Protocol 3-18
3.3.6.1.1 Household Sample 3-18
3.3.6.1.2 Volunteer Sample 3-18
3.3.6.2 Scheduling Protocol 3-19
3.3.6.2.1 Household Sample 3-19
3.3.6.2.2 Volunteer Sample 3-19
3.3.6.3 NAAL Hotline 3-19
3.3.7 Quality Control Measures and Feedback to Staff 3-19
3.3.7.1 Quality Control Measures 3-19
3.3.7.2 Feedback to Staff 3-20
3.3.7.2.1 Weekly Newsletter 3-20
3.3.7.2.2 Weekly Supervisor Calls 3-20
3.3.7.2.3 E-mail 3-21
3.3.7.3 CAPI Help Desk 3-21
3.3.8 Cooperation Rates 3-21
3.3.8.1 Cooperation Rates for the Field Test 3-21

CONTENTS—Continued

3.3.8.2 General Cooperation Rate Factors 3-21
3.3.8.3 Field Test Results 3-22
3.3.8.4 Summary of Cooperation and Cooperation Rates and Actual Sample Sizes 3-23
3.3.8.4.1 Probability Sample 3-23
3.3.8.4.2 Volunteer Sample 3-26
3.3.8.4.3 Overall Sample 3-27
3.3.8.5 Demographic Profiles of Respondents 3-28
3.4 DATA PREPARATION AND PROCESSING 3-28
3.4.1 Overview of Data Preparation and Processing Activities for the Screener, the Background Questionnaire, and the Interviewer Guide 3-28
3.4.2 Receipt and Batching of Cases 3-29
3.4.3 Processing of Hard-Copy Assessment Materials 3-29
3.4.3.1 Procedures 3-29
3.4.3.2 Core Assessment Validation 3-29
3.4.4 Delivery of Assessment Booklets for Scoring 3-31
3.4.5 Quality Control of Data in the Study Database 3-31

CHAPTER 4: FIELD-TEST SCORING AND ANALYSIS OF FIELD-TEST DATA 4-1
4.1 DEVELOPMENT OF SCORING PROCEDURES 4-1
4.1.1 Refinement of Scoring Rubrics 4-1
4.1.2 Development of Scorer Training Materials 4-2
4.1.3 Field-Test Scoring Procedures 4-2
4.2 ANALYSIS OF FIELD-TEST DATA 4-3
4.2.1 Background Questionnaire Analyses 4-3
4.2.2 Analysis of Field-Test Cognitive Data 4-4
4.2.2.1 Item Analyses 4-6
4.2.2.2 Flagged Items 4-10
4.2.2.3 Partial Credit 4-11
4.2.2.4 Algorithm Used to Identify Adult Literacy Supplemental Assessment Respondents 4-12
4.2.3 Analysis of Fluency Addition to the NAAL Field-Test Data 4-14

CHAPTER 5: FIELD TEST OF THE FLUENCY ADDITION TO NAAL 5-1
5.1 INTRODUCTION 5-1
5.2 SAMPLE DESIGN 5-2
5.3 DATA COLLECTION 5-2
5.3.1 Data Collection Instruments 5-2
5.3.2 Interviewer Materials 5-3
5.3.3 Field Staff Recruitment and Training 5-4
5.3.3.1 Field Staff Recruitment 5-4
5.3.3.2 Field Staff Training 5-4
5.3.4 Data Collection Effort 5-5
5.3.4.1 Field Staff and Organization 5-5
5.3.4.2 Monitoring of Data Collection Progress 5-5

CONTENTS—Continued

5.3.4.3 Contact Protocol 5-5
5.3.4.4 Data Collection Procedures 5-6
5.3.5 Quality Control Measures and Feedback to Staff 5-6
5.3.6 Summary of Data Collection Results 5-7
5.4 DATA PREPARATION AND PROCESSING 5-8
5.4.1 Receipt, Batching, and Editing of Cases 5-8
5.4.2 Processing and Delivery of Hard Copy Assessment Materials 5-9
5.4.3 Processing and Delivery of Oral Module Data 5-9
5.5 REFINEMENTS FOR MAIN HOUSEHOLD STUDY 5-9

CHAPTER 6: FIELD TEST OF ADULT LITERACY SUPPLEMENTAL ASSESSMENT 6-1
6.1 PRETEST PHASE 1 6-1
6.2 PRETEST PHASE 2 6-2
6.3 PRETEST PHASE 3 6-3
6.4 RESULTS OF THE FIELD TEST 6-4

CHAPTER 7: SAMPLE DESIGN 7-1
7.1 INTRODUCTION 7-1
7.2 HOUSEHOLD SAMPLES 7-2
7.2.1 Key Design Features and Sizes of the Household Samples 7-2
7.2.2 Sampling Frames for the Household Study 7-6
7.2.2.1 Sampling Frame for Primary Sampling Units 7-7
7.2.2.2 Sampling Frame for Segments 7-8
7.2.2.3 Sampling Frame for Housing Units 7-9
7.2.2.4 Sampling List of Persons 7-9
7.2.3 Sample Selection for the Household Study 7-10
7.2.3.1 Selection of Primary Sampling Units 7-10
7.2.3.1.1 Stratification of Primary Sampling Units 7-10
7.2.3.1.2 PSU Selection 7-13
7.2.3.2 Selection of Segments 7-18
7.2.3.3 Selection of Housing Units 7-18
7.2.3.4 Selection of Persons 7-20
7.2.3.5 Procedures for Selecting Missed Structures and Hidden Housing Units 7-20
7.2.3.6 Overall Probabilities of Selection 7-21
7.3 PRISON SAMPLE 7-23
7.3.1 Sampling Frames for the Prison Study 7-23
7.3.1.1 Overview of Prison Study Sampling Frames 7-23
7.3.1.2 Data Sources Used to Create Sampling Frames for the Prison Study 7-24
7.3.2 Selection of Prison Sample 7-25
7.3.3 Selection of Inmates Within Facilities 7-27
7.4 ASSIGNMENT OF BOOKLETS TO RESPONDENTS 7-32

CONTENTS—Continued

CHAPTER 8: MAIN HOUSEHOLD STUDY DATA COLLECTION AND QUALITY CONTROL 8-1
8.1 INTRODUCTION 8-1
8.2 LISTING 8-2
8.2.1 Staff Organization for Listing 8-2
8.2.2 Listing Materials 8-2
8.2.3 Training Listers 8-3
8.2.4 Listing Operation 8-4
8.2.5 Quality Control Procedures 8-5
8.2.5.1 Quality Control of Listing Sheets 8-6
8.2.5.2 Quality Control of the Listing Operation 8-6
8.2.5.2.1 Hidden Housing Unit Procedure 8-6
8.2.5.2.2 Missed Structure Procedure 8-7
8.3 DATA COLLECTION INSTRUMENTS AND INTERVIEWER MATERIALS 8-7
8.3.1 Assessment Materials 8-8
8.3.2 CAPI Data Collection Instruments and System Features 8-8
8.3.2.1 Screener 8-8
8.3.2.2 Background Questionnaire 8-9
8.3.2.3 Interviewer Guide 8-10
8.3.2.4 Interviewer Management System 8-10
8.3.2.5 Supervisor Management System 8-13
8.3.3 Interviewer Materials 8-13
8.3.3.1 Advance Materials 8-14
8.3.3.2 Household Folder 8-15
8.3.3.3 Handcards 8-16
8.3.3.4 Noninterview Report Forms 8-16
8.3.3.5 Interviewer Manual 8-17
8.3.3.6 Oral Module Booklet 8-17
8.4 FIELD STAFF TRAINING 8-17
8.4.1 Approach to Training 8-18
8.4.2 Train-the-Trainer Trainings 8-18
8.4.3 Supervisor Training 8-19
8.4.4 General Interviewer Techniques 8-19
8.4.5 Project-Specific Training 8-20
8.4.5.1 Interviewer Training Materials 8-20
8.4.5.2 Interviewer Training Techniques 8-20
8.4.5.3 In-Person Training Program for Interviewers 8-22
8.5 CONDUCT OF THE FIELD WORK 8-24
8.5.1 Field Organization 8-25
8.5.1.1 Recruiting Field Staff 8-25
8.5.1.2 Field Staff Attrition and Replacement 8-26
8.5.2 Field Management 8-28
8.5.2.1 Reporting Structure 8-28
8.5.2.2 Reporting Procedures and Tools 8-29
8.6 DATA COLLECTION OPERATIONS 8-30
8.6.1 General Approach 8-30
8.6.2 Schedule and Production 8-31
8.6.3 Data Collection Procedures 8-33

CONTENTS—Continued

8.7 QUALITY CONTROL MEASURES AND FEEDBACK TO STAFF 8-34
8.7.1 Validation 8-35
8.7.2 Computer-Generated Reports 8-36
8.7.3 Observation 8-37
8.7.3.1 Tape-Recorded Interviews 8-38
8.7.3.2 In-Person Observations 8-38
8.7.4 CAPI Help Desk 8-38
8.8 ACHIEVING HIGH RESPONSE RATES 8-39
8.8.1 Interviewers' Ability to Obtain Cooperation 8-39
8.8.2 Callback Procedures 8-40
8.8.3 Efforts to Convert Nonresponse 8-40
8.8.3.1 Refusals 8-40
8.8.3.2 Not at Home 8-41
8.8.3.3 Language Problem 8-41
8.8.3.4 Illness 8-41
8.8.3.5 Vacant Housing Unit 8-41
8.8.4 Strategies to Increase Response Rates 8-41
8.8.4.1 Neighbor Information 8-42
8.8.4.2 Hard-Copy Screener 8-42
8.8.4.3 FedEx Mailings 8-42
8.8.4.4 Interviewer Incentives 8-43
8.8.2 Reasons for Nonresponse 8-43
8.8.3 Demographic Profiles of Respondents 8-47
8.9 DATA PREPARATION AND PROCESSING 8-49
8.9.1 Overview of Data Preparation and Processing Activities 8-50
8.9.2 Receipt, Batching, and Editing of Cases 8-50
8.9.3 Processing of Hard-Copy Assessment Materials 8-50
8.9.4 Processing of Oral Module Data and Zip Disks 8-51
8.9.5 Data Processing for the ALSA 8-52
8.9.6 Delivery of Assessment Booklets to the Scoring Contractor 8-52
8.9.7 Quality Control of Data in the Study Database 8-52

CHAPTER 9: CORRECTIONAL INSTITUTION STUDY DATA COLLECTION AND QUALITY CONTROL 9-1
9.1 INTRODUCTION 9-1
9.2 GAINING COOPERATION 9-2
9.3 DATA COLLECTION MATERIALS AND INSTRUMENTS 9-3
9.4 INTERVIEWER SELECTION AND TRAINING 9-4
9.4.1 Sampling Procedures 9-5
9.4.2 Inmate Participation 9-7
9.4.3 Background Questionnaire 9-7
9.4.4 Administrative Procedures 9-7
9.5 DATA COLLECTION 9-8
9.6 QUALITY CONTROL 9-9

CONTENTS—Continued

CHAPTER 10: REDUCING THE RISK OF DATA DISCLOSURE 10-1
10.1 HOUSEHOLD SAMPLE 10-2
10.1.1 Personal Identifiers and Geographic Identifiers 10-3
10.1.2 Data Coarsening 10-3
10.2 PRISON SAMPLE 10-10
10.2.1 Personal Identifiers and Geographic Information 10-11
10.2.2 Ability to Match with Administrative Data and Other External Data 10-11
10.2.3 Data Coarsening 10-12
10.3 DATA SWAPPING 10-14

CHAPTER 11: RESPONSE RATES AND NONRESPONSE BIAS ANALYSIS 11-1
11.1 HOUSEHOLD LITERACY STUDY 11-2
11.1.1 Analysis Weights 11-2
11.1.2 Unit Response Rates 11-3
11.1.3 Item Nonresponse 11-4
11.1.4 Nonresponse Bias Analysis 11-7
11.1.4.1 Evaluating Bias Owing to Screener Nonresponse 11-7
11.1.4.1.1 Screener Bivariate Analysis 11-9
11.1.4.1.2 Screener Multivariate Analysis 11-12
11.1.4.1.3 Potential for Screener Nonresponse Bias Remaining After Weighting Procedures 11-18
11.1.4.2 Evaluating Bias Owing to Background Questionnaire Nonresponse 11-34
11.1.4.2.1 Background Questionnaire Bivariate Analysis 11-34
11.1.4.2.2 Background Questionnaire Multivariate Analysis 11-36
11.1.4.2.3 Potential for Background Questionnaire Nonresponse Bias Remaining After Weighting Procedures 11-43
11.1.5 Conclusion 11-44
11.2 PRISON LITERACY STUDY 11-60
11.2.1 Unit Nonresponse 11-60
11.2.2 Item Nonresponse Rates 11-63
11.2.3 Nonresponse Bias Analysis 11-64
11.2.3.1 Bivariate Analysis 11-64
11.2.3.2 Item Nonresponse Bias Estimates 11-69
11.2.3.3 Multivariate Analysis 11-74
11.2.4 Conclusion 11-77

CHAPTER 12 WEIGHTING AND VARIANCE ESTIMATION 12-1
12.1 HOUSEHOLD SAMPLES 12-1
12.1.1 Preliminary Steps in Weighting 12-4
12.1.2 Screener Base Weights and Nonresponse Adjustments 12-6
12.1.2.1 Screener Base Weights 12-6
12.1.2.2 Screener Nonresponse Adjustment 12-7
12.1.3 Background Questionnaire Base Weights, Nonresponse Adjustments, and Trimming 12-13

CONTENTS—Continued

12.1.3.1 Background Questionnaire Base Weights 12-13
12.1.3.2 Background Questionnaire Nonresponse Adjustment 12-14
12.1.3.2.1 Nonliteracy-Related Nonresponse 12-14
12.1.3.2.2 Literacy-Related Nonresponse 12-16
12.1.3.3 Background Questionnaire Trimming Adjustment 12-18
12.1.4 Calibration Adjustments Prior to Compositing 12-19
12.1.4.1 Control Totals 12-20
12.1.4.2 Calibration 12-20
12.1.5 Compositing Data from the National and State Household Components 12-23
12.1.5.1 Composite Estimation Procedure 12-23
12.1.5.2 Estimating the Compositing Factors 12-25
12.1.6 Computing Final Weights 12-28
12.1.6.1 Compositing the NAAL and SAAL Samples 12-28
12.1.6.2 Raking Composite Weights 12-28
12.1.6.2.1 Control Totals 12-28
12.1.6.2.2 Final Adjustments 12-29
12.1.6.3 Quality Control 12-35
12.1.7 Replicate Weights 12-36
12.1.7.1 Replicate Weights for the National Sample 12-37
12.1.7.2 Replicate Weights for the State Sample 12-38
12.1.7.3 Final Replication for the National and State Samples 12-38
12.1.7.3.1 Active Replicates 12-39
12.1.7.3.2 Replicated Control Totals 12-39
12.1.7.4 Jackknife and Taylor Series Variance Estimation 12-42
12.1.7.5 Evaluation of Variance Estimates 12-43
12.2 WEIGHTING THE PRISON STUDY SAMPLE 12-46
12.2.1 Computing Prison Base Weights 12-46
12.2.2 Prison Nonresponse Adjustment 12-47
12.2.3 Computing Inmate Base Weights 12-48
12.2.4 Accounting for Nonresponse and Noncoverage 12-49
12.2.5 Final Inmate Weights 12-51
12.2.6 Replicate Weights for Variance Estimation for the Prison Study 12-51
12.2.7 Evaluation of Variance Estimates 12-53

CHAPTER 13: SCORING 13-1
13.1 COGNITIVE ITEMS 13-1
13.1.1 Refinement of Training Materials 13-1
13.1.2 Scorer Training 13-2
13.1.3 Scoring Procedures and Quality Control 13-3
13.1.4 Scoring the Prison Sample 13-8
13.2 FLUENCY ADDITION TO NAAL (FAN) 13-9
13.2.1 Background 13-9
13.2.2 Scoring the FAN Data 13-10
13.2.3 FAN Data Products 13-12
13.2.3.1 Span Summary 13-13
13.2.3.2 Articulation Rate 13-13
13.2.3.3 Number of Short Pauses 13-13

CONTENTS—Continued

13.2.3.4 Number of Long Pauses 13-14
13.2.3.5 Number of Deletions 13-14
13.2.3.6 False Start Count 13-14
13.2.3.7 Secondary Machine Scores 13-14
13.3 ADULT LITERACY SUPPLEMENTAL ASSESSMENT 13-15
13.3.1 Background and Rationale 13-15
13.3.2 Field Scoring of Core Items to Identify Respondents for the Adult Literacy Supplemental Assessment 13-16
13.3.2.1 Interviewer Training 13-16
13.3.2.2 Quality Control 13-16
13.3.3 Recording of Responses 13-17
13.3.3.1 Interviewer Training 13-18
13.3.3.2 Quality Control 13-19
13.3.3.3 Data Entry of Questionnaires 13-19

CHAPTER 14: ITEM ANALYSIS, SCALING, AND ESTIMATES OF SUBPOPULATION PROFICIENCIES 14-1
14.1 PROCEDURES AND QUALITY CONTROL 14-1
14.1.1 Analysis Teams 14-1
14.2 ITEM ANALYSIS 14-2
14.2.1 Partial Credit Items 14-2
14.3 SCALING METHODOLOGY 14-3
14.3.1 The Scaling Model 14-4
14.3.2 Linking to the 1992 NALS 14-6
14.3.3 Item Parameter Estimation 14-6
14.4 DIRECT ESTIMATION OF SUBPOPULATION DISTRIBUTIONS OF PROFICIENCY USING DIRECT ESTIMATION WITH THE AM SOFTWARE 14-8
14.4.1 Background 14-8
14.4.2 Direct Estimation of MML Models 14-10
14.4.2.1 Estimation 14-11
14.4.3 Weighting and Variance Estimation in Complex Samples 14-13
14.5 LINKING THE 2003 SCALE TO THE 1992 SCALE 14-15
14.6 MINIMUM SAMPLE SIZES FOR REPORTING SUBGROUP RESULTS 14-16

CHAPTER 15: THE LITERACY OF ADULTS WITHOUT COGNITIVE DATA 15-1
15.1 INTRODUCTION 15-1
15.1.1 Missing Background Data 15-1
15.1.2 Missing Cognitive Data 15-2
15.2 THE NORMAL TREATMENT OF MISSING COGNITIVE DATA 15-2
15.2.1 Omitted, Not Reached 15-3
15.2.2 Cognitive Data Missing by Design 15-4
15.3 REASONS FOR MISSING DATA 15-4
15.3.1 Disposition Codes and Literacy Skills 15-4
15.3.2 Household Sample Internal Evidence for the Validity of Reasons 15-6
15.4 POPULATION ESTIMATES FOR ADULTS WITH UNMEASURABLE LITERACY SKILLS 15-12

CONTENTS—Continued

15.4.1 Using "Reasons" to Improve Treatment of Missing Cognitive Data 15-12
15.4.2 Maintaining Comparability of Population Coverage Between 1992 and 2003 Assessments 15-16

CHAPTER 16: VARIABLE CONSTRUCTION AND FILE DEVELOPMENT 16-1
16.1 FILE CONSTRUCTION PROCEDURES AND FILE CONTENTS 16-1
16.1.1 Household Study 16-2
16.1.2 Prison Study 16-2
16.1.3 NAAL Item Parameter Files 16-2
16.2 CONSTRUCTION OF DERIVED VARIABLES 16-3
16.3 ANALYZING 2003 NAAL DATA USING AM SOFTWARE 16-3
16.4 USING THE ELECTRONIC CODEBOOK 16-4

CHAPTER 17: THE NAAL HEALTH LITERACY COMPONENT 17-1
17.1 DEFINING HEALTH LITERACY AND DETERMINING THE CONTENT OF THE HEALTH LITERACY COMPONENT 17-1
17.2 IDENTIFYING HEALTH LITERACY TASKS 17-1
17.2.1 Overlap of the Health Literacy Tasks with the Tasks on the Prose, Document and Quantitative Scales 17-2
17.2.2 Framework of the Health Literacy Scale 17-2
17.3 SCALING THE HEALTH ITEMS 17-3
17.3.1 The Scaling Model 17-4
17.3.2 Item Parameter Estimation 17-4
17.4 REPORTING RESULTS ON THE HEALTH LITERACY SCALE 17-5
17.4.1 Establishing Health Literacy Performance Levels 17-5
17.4.2 Relationship of Health Scale to Prose, Document, and Quantitative Scales 17-12

References: R-1
APPENDIX A: DERIVED VARIABLES A-1
APPENDIX B: INSTRUCTIONS ON USING AM SOFTWARE TO ANALYZE THE 2003 NAAL DATA B-1
APPENDIX C: ELECTRONIC CODEBOOK FOR WINDOWS USER'S MANUAL C-1
APPENDIX D: NAAL ITEM DESCRIPTIONS AND PARAMETERS D-1
APPENDIX E: P-VALUES OF NAAL 2003 ITEMS E-1
APPENDIX F: NAAL HOUSEHOLD BACKGROUND QUESTIONNAIRE F-1
APPENDIX G: NAAL PRISON BACKGROUND QUESTIONNAIRE G-1

LIST OF TABLES

Table **Page**

Table 2-1. Number of background questionnaire cognitive laboratory participants, by selected characteristics: 2003 2-7
Table 2-2. Coding of the 2003 prose tasks, by type of text content, type of process/strategy, and Lexile score: 2003 2-24
Table 2-3. Coding of the 2003 document tasks, by type of text content, and type of process/strategy: 2003 2-25
Table 2-4. Coding of the 2003 quantitative tasks, by type of text content, type of operation, and number of operations: 2003 2-26
Table 2-5. NAAL health items, by distribution across the clinical, prevention, and navigation health domains: 2003 2-29
Table 2-6. Number of NAAL assessment cognitive laboratory participants – round 1, by selected characteristics: 2003 2-31
Table 2-7. Number of NAAL assessment cognitive laboratory participants – round 2, by selected characteristics: 2003 2-32
Table 2-8. Oral reading fluency passages included in the fluency addition to the NAAL (FAN) operational assessment, by selected properties: 2003 2-38
Table 2-9. Oral reading fluency cognitive laboratory participants, by selected characteristics: 2003 2-39
Table 2-10. Number of Adult Literacy Supplemental Assessment (ALSA) questions, by type of question: 2003 2-43
Table 2-11. NAAL field-test booklet design: 2003 2-46
Table 2-12. NAAL operational booklet design: 2003 2-48
Table 2-13. Booklet layout for operational Fluency Addition to the NAAL (FAN) administration: 2003 2-49
Table 2-14. Mapping between assessment booklets and Fluency Addition to the NAAL (FAN) passages: 2003 2-50
Table 3-1. Minimum expected sample sizes per booklet group, by age and race/ethnicity: 2001 3-4
Table 3-2. Expected sample yield from the household sample, by age and race/ethnicity: 2001 3-5
Table 3-3. Expected sample yield from the volunteer sample, by age and race/ethnicity: 2001 3-5
Table 3-4. Expected sample yield from the household and volunteer samples, by age and race/ethnicity: 2001 3-5
Table 3-5. Average administration times for item blocks: 2001 3-10
Table 3-6. Actual sample yield from the household sample, by age and race/ethnicity: 2001 3-23
Table 3-7. Expected and actual cooperation rates from the household sample, by survey component: 2001 3-24
Table 3-8. Screener nonresponse in the household sample, by reason for nonresponse: 2001 3-25
Table 3-9. Background questionnaire nonresponse in the household sample, by reason for nonresponse: 2001 3-25
Table 3-10. Assessment nonresponse in the household sample, by reason for nonresponse: 2001 3-26
Table 3-11. Actual sample yield from the volunteer sample, by age and race/ethnicity: 2001 3-27
Table 3-12. Actual sample yield from the household and volunteer samples, by age and race/ethnicity: 2001 3-27
Table 3-13. Actual sample yield from the household and volunteer samples, by race/ethnicity and age: 2001 3-28

LIST OF TABLES—Continued

Table **Page**

Table 4-1. Field-test booklet spiral: 2003 4-5
Table 4-2. Number of NAAL noncore and core field-test items, by percent correct (*p*-value): 2003 4-7
Table 4-3. Classification rules for differential item functioning (DIF) categories: 2003 4-8
Table 4-4. Number of NAAL noncore and core field-test items, by differential item functioning (DIF) categories: 2003 4-9
Table 4-5. Number of NAAL noncore and core field-test items, by biserial correlation: 2003 4-10
Table 4-6. Number of NAAL noncore and core field-test items, by interrater reliability (exact agreement): 2003 4-10
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 4-13
Table 4-8. Fluency Addition to the NAAL (FAN) passages, by field-test results: 2003 4-15
Table 5-1. Percent of interviews completed in the FAN field test, by race/ethnicity, language status, and educational attainment: 2002 5-8
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 7-4
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 7-6
Table 7-3. Distribution of formed primary sampling units, by Metropolitan Statistical Area (MSA) and region status: 2003 7-7
Table 7-4. Distributions of primary sampling unit characteristics, by Metropolitan Statistical Area (MSA) status, mean, median, and percentiles: 2003 7-8
Table 7-5. Percentile distributions of NAAL-SAAL sampled segments: 2003 7-9
Table 7-6. Variables used in NAAL noncertainty primary sampling unit stratification: 2003 7-12
Table 7-7. Variables used in SAAL noncertainty primary sampling unit stratification: 2003 7-13
Table 7-8. NAAL-SAAL selection probabilities: 2003 7-14
Table 7-9. Distribution of sampled PSUs, by certainty status: 2003 7-17
Table 7-10. Number of overlapping PSUs in the NAAL and SAAL samples: 2003 7-17
Table 7-11. Number of housing units added by the missed structure and hidden housing unit procedures: 2003 7-21
Table 7-12. Summary of data sources used to create the prison sampling frame: 2003 7-25
Table 7-13. Distribution of selected prisons, by cooperation status and census region: 2003 7-27
Table 7-14. Background questionnaire sample counts by cooperation status and prison characteristics: 2003 7-30
Table 7-15. Exercise sample counts by cooperation status and prison characteristics: 2003 7-31
Table 7-16. Exercise sample counts by cooperation status and background questionnaire variables in the prison study, by inmate characteristic: 2003 7-32
Table 8-1. Overview of the NAAL study-specific interviewer training session: 2003 8-23
Table 8-2. Characteristics of field workers on NAAL main study, by gender, age, experience, and assignments: 2003 8-27
Table 8-3. Average administration time, by instrument: 2003 8-34
Table 8-4. Sample counts by instrument and reasons for nonresponse: 2003 8-45
Table 8-5. Demographic profile of respondents in the combined household sample, by instrument: 2003 8-48

LIST OF TABLES—Continued

Table **Page**

Table 10-1. Source, description, and categories of age-related recoded variables for household sample: 2003 10-5
Table 10-2. Source, description, and categories of race/ethnicity recoded variables for household sample: 2003 10-5
Table 10-3. Source, description, and categories of language recoded variables for household sample: 2003 10-6
Table 10-4. Source, description, and categories of education recoded variables for household sample: 2003 10-7
Table 10-5. Source, description, and categories of income-related recoded variables for household sample: 2003 10-9
Table 10-6. Source, description, and categories of recoded variables for prison sample public-use microdata file: 2003 10-12
Table 11-1. Household study unweighted (UW) and weighted (W) unit response rates, by analysis variable, in percent: 2003 11-5
Table 11-2. Household study variables used in screener nonresponse bias analysis, by source and values: 2003 11-8
Table 11-3. Household study sample distribution of screener respondents versus eligible dwelling units, by analysis domain: 2003 11-10
Table 11-4. Household study estimates of screener nonresponse bias, by analysis domain: 2003 11-11
Table 11-5. Household study multivariate CHAID analysis of screener response indicators, by response cell: 2003 11-15
Table 11-6. Household study multivariate logistic regression analysis of screener response indicators, by predictor: 2003 11-17
Table 11-7. Household study screener weighting effects for the national NAAL household sample, by subgroup: 2003 11-20
Table 11-8. Household study screener weighting effects for Kentucky, by subgroup: 2003 11-22
Table 11-9. Household study screener weighting effects for Maryland, by subgroup: 2003 11-24
Table 11-10. Household study screener weighting effects for Massachusetts, by subgroup: 2003 11-26
Table 11-11. Household study screener weighting effects for Missouri, by subgroup: 2003 11-28
Table 11-12. Household study screener weighting effects for New York, by subgroup: 2003 11-30
Table 11-13. Household study screener weighting effects for Oklahoma, by subgroup: 2003 11-32
Table 11-14. Household study variables used in background questionnaire nonresponse bias analysis, by source and values: 2003 11-35
Table 11-15. Household study sample distribution of background questionnaire respondents versus eligible sample persons, by analysis domain: 2003 11-37
Table 11-16. Household study estimates of background questionnaire nonresponse bias, by analysis domain: 2003 11-38
Table 11-17. Household study multivariate CHAID analysis of background questionnaire response indicators, by response cell: 2003 11-41
Table 11-18. Household study multivariate logistic regression analysis of background questionnaire response indicators, by predictor: 2003 11-42
Table 11-19. Household study background questionnaire weighting effects for the national NAAL household sample, by subgroup: 2003 11-45
Table 11-20. Household study background questionnaire weighting effects for Kentucky, by subgroup: 2003 11-48
Table 11-21. Household study background questionnaire weighting effects for Maryland, by subgroup: 2003 11-50

LIST OF TABLES—Continued

Table **Page**

Table 11-22. Household study background questionnaire weighting effects for Massachusetts, by subgroup: 2003 11-52
Table 11-23. Household study background questionnaire weighting effects for Missouri, by subgroup: 2003 11-54
Table 11-24. Household study background questionnaire weighting effects for New York, by subgroup: 2003 11-56
Table 11-25. Household study background questionnaire weighting effects for Oklahoma, by subgroup: 2003 11-58
Table 11-26. Prison study variables used to calculate response rates, by source and values: 2003 11-61
Table 11-27. NAAL prison study weighted unit response rates, by domain, in percent: 2003 11-62
Table 11-28. Prison study item response rates below 85 percent, by variable: 2003 11-63
Table 11-29. Prison study sample distribution for item respondents versus eligibles, by survey domain: 2003 11-65
Table 11-30. Prison study estimate of item nonresponse bias, by survey domain: 2003 11-70
Table 12-1. Screener base weight distribution for the household samples, by sample: 2003 12-7
Table 12-2. Variables used in forming weighting classes for the screener nonresponse adjustment, by sample: 2003 12-10
Table 12-3. Distribution of the background questionnaire base weights for household samples, by sample: 2003 12-14
Table 12-4. Variables forming background questionnaire nonresponse adjustment weighting classes, by sample: 2003 12-15
Table 12-5. Distribution of trimmed weights and trimming factors, by minority status and race: 2003 12-19
Table 12-6. Variables involved in the calibration process prior to compositing, by sample: 2003 12-21
Table 12-7. Calculations of $Ratio_{\delta}$(var) and $\overline{\beta}_{\delta}$, by certainty in national sample, race/ethnicity, and state: 2003 12-27
Table 12-8. Variables involved in the raking process after compositing, by sample: 2003 12-30
Table 12-9. Raking factors for the first round of raking after compositing, by state and domain: 2003 12-31
Table 12-10. Raking factors for the second round of raking after compositing, by state and domain: 2003 12-33
Table 12-11. Distribution of final background questionnaire weights, by domain: 2003 12-35
Table 12-12. Active replicates for subdomains of the NAAL analysis file, by replicate, region, and state: 2003 12-40
Table 12-13. GVF parameter, by subgroup: 2003 12-42
Table 12-14. Standard errors for selected background questionnaire items, by background questionnaire item: 2003 12-44
Table 12-15. Distribution of prison study prison base weights: 2003 12-47
Table 12-16. Prison study prison nonresponse adjustment factors, by adjustment class: 2003 12-48
Table 12-17. Distribution of inmate base weights: 2003 12-49
Table 12-18. Raking factors by raking dimension for the prison study, by domain: 2003 12-50
Table 12-19. Distribution of prison study final inmate weights, by raking dimension: 2003 12-52
Table 12-20. Design effects for selected prison study background questionnaire and raking variables, by item: 2003 12-54
Table 12-21. Prison study active replicates, by selected subdomains: 2003 12-56

LIST OF TABLES—Continued

Table **Page**

Table 13-1. Inter-rater reliability statistics for NAAL household and prison items, by block: 2003 ... 13-4
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 ... 13-12
Table 13-3. Summary of discrepancies in core item scoring between interviewers and home office staff, by core item: 2004 ... 13-18
Table 14-1. NAAL items, by item treatment during scaling: 2003 ... 14-8
Table 14-2. Transformation constants (standard deviations and means) using new item parameters, by literacy scale: 1992 and 2003 ... 14-15
Table 15-1. Percent distribution of final weighted NAAL sample, by presence of cognitive data, reason for missing data, and sample type: 2003 ... 15-5
Table 15-2. Distribution of final weighted NAAL sample of 19,714, by presence of cognitive data and detailed reason for missing data: 2003 ... 15-6
Table 15-3. Percentage of household study population by presence of cognitive data, reason for missing data, and population group: 2003 ... 15-7
Table 15-4. Standard errors for table 15-3 ... 15-9
Table 15-5. Variables used as predictors in the logistic regression models for imputation in the NAAL household study, by value label: 2003 ... 15-14
Table 15-6. Variables used as predictors in the logistic regression models for imputation in the NAAL prison study, by value label: 2003 ... 15-15
Table 15-7. Variables used as predictors in the logistic regression models for imputation in the NALS household study, by value label: 1992 ... 15-17
Table 15-8. Variables used as predictors in the logistic regression models for imputation in the NALS prison study, by value label: 1992 ... 15-18
Table 17-1. Overview of the NAAL literacy levels: 2003 ... 17-7
Table 17-2. Item map for the NAAL health literacy items: 2003 ... 17-9
Table 17-3. Correlations among NAAL literacy scales, by literacy scale: 2003 ... 17-12
Table D-1. Prose literacy item description and parameters for the NAAL, by ID: 2003 ... D-2
Table D-2. Document literacy item description and parameters for the NAAL, by ID: 2003 ... D-5
Table D-3. Quantitative literacy item description and parameters for the NAAL, by ID: 2003 ... D-8
Table D-4. Health literacy item description and parameters for the NAAL, by ID: 2003 ... D-10
Table E-1. P-values of NAAL items: 2003 ... E-2

LIST OF FIGURES

Figure **Page**

Figure 3-1. Distribution of discrepancies in core assessment scoring results, by core item number: 2001 3-30
Figure 8-1. Percentage of closed-out cases, by week of field period and phases of data collection: 2003 8-33
Figure 11-1. Household study multivariate CHAID analysis of screener response indicators: 2003 .. 11-13
Figure 11-2. Household study multivariate CHAID analysis of background questionnaire response indicators: 2003 11-40
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 11-75
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 11-76
Figure 12-1. Distribution of the unknown eligibility adjustment factors for the household samples, by sample: 2003 12-11
Figure 12-2. Distribution of the screener nonresponse adjustment factors for the household samples, by sample: 2003 12-12
Figure 12-3. Distribution of the background questionnaire nonresponse adjustment factors for the household samples, by sample: 2003 12-17
Figure 14-1. Example of item response curve (item N100201) from the NAAL: 2003 14-7
Figure 17-1. Difficulty of selected NAAL health literacy tasks: 2003 17-10
Figure B-1. AM program starting window: 2003 ..... B-3
Figure B-2. Example of calculating mean prose literacy scores for men and women: 2003 ..... B-6
Figure B-3. NAAL cutpoints, by literacy scale and literacy level: 2003 ..... B-7
Figure B-4. NAAL cut scores, by literacy scale: 2003 ..... B-8
Figure B-5. Column headings in AM output file, by NAAL prose scale cut scores: 2003 ..... B-8
Figure B-6. Column headings in AM output file, by NAAL literacy levels: 2003 ..... B-9
Figure B-7. Example of calculating the percentage of men and women within document literacy levels: 2003 ..... B-9
Figure B-8. AM *t* test dialog box: 2003 ..... B-11

LIST OF EXHIBITS

Exhibit **Page**

Exhibit 3-1. Translation guide, NAAL data collection: 2001 field test 3-11

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 PROFICIENCES

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.