

Adult Literacy and Lifeskills Survey (ALL) 2003: U.S. Nonresponse Bias Analysis

September 2009

NCES 2009-063
U.S. DEPARTMENT OF EDUCATION



Page intentionally left blank.

Adult Literacy and Lifeskills Survey (ALL) 2003: U.S. Nonresponse Bias Analysis

September 2009

Wendy Van de Kerckhove
Tom Krenzke
Leyla Mohadjer
Westat

Edited by:
Lydia Malley
Education Statistics Services Institute – Child Trends

Eugene Owen
Project Officer
National Center for Education Statistics

U.S. Department of Education

Arne Duncan
Secretary

Institute of Education Sciences

John Q. Easton
Acting Director

National Center for Education Statistics

Stuart 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, and 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

September 2009

The NCES World Wide Web Home Page address is <http://nces.ed.gov>.

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

Suggested Citation

Van de Kerckhove, W., Krenzke, T., and Mohadjer, L. (2009). *Adult Literacy and Lifeskills Survey (ALL) 2003: U.S. Nonresponse Bias Analysis* (NCES 2009-063). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

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

Eugene Owen
(202) 502-7422
Eugene.Owen@ed.gov

Executive Summary

Most survey estimates are subject to some unknown level of nonresponse bias resulting from failure to obtain responses from all sampled units. There are two types of nonresponse: unit nonresponse refers to the failure of an eligible sample unit to participate in the survey, whereas item nonresponse occurs when a response to a specific survey question is missing for a respondent. For the Adult Literacy and Lifeskills Survey (ALL), survey estimates of literacy-related skills are subject to unit nonresponse from the screener, background questionnaire (BQ), and assessment, and also subject to item nonresponse to BQ items.

The statistical standards set forth by the National Center for Education Statistics require a nonresponse bias analysis to be performed for any data collection stage where the response rate was below 85 percent. At the unit level, the potential bias was analyzed at the BQ level, which had a weighted response rate of 80 percent. At the item level, there were eight BQ items with less than an 85 percent response rate. The unit analyses were composed of a bivariate analysis, multivariate analysis, and an analysis of the weighting effects on nonresponse bias. The item analyses included a similar study into the effect of nonresponse.

For the bivariate unit-level analysis, using the sample base weights, estimated percentages among the respondents were compared to that of the total eligible sample in order to identify any potential bias due to nonresponse. The analysis was limited to using variables that are known for both respondents and nonrespondents. While some statistically significant differences exist,¹ the potential for bias is very small since the absolute difference between estimated percentages was less than 3 percentage points among all domains² considered. The multivariate analysis probed further into identifying where some potential for nonresponse bias may exist. The analysis showed that the lowest response rates were among metropolitan statistical areas in the Northeast. However, these areas and other pockets with low response rates were the focus of the nonresponse weighting adjustments. The analysis showed that the nonresponse weighting adjustments were highly effective in reducing the bias. The general conclusion of this analysis is that the potential amount of nonresponse bias due to unit nonresponse in ALL is likely to be negligible.

¹ A chi-square test and *F* test with a significance level of $\alpha = 0.05$ were used to support the analysis. Details are stated in chapter 2.

² Domains are defined as categories of the analysis variables.

Eight BQ items had response rates of less than 85 percent; seven were related to income, and the other item provided the type of certificate, degree, or diploma, for their education attained. The item nonresponse analysis revealed some low response rates in certain domains and some indication that a potential for bias may exist.

There were no weighting adjustments at the item level or imputation procedures completed on these items that may have reduced the bias due to item nonresponse. Users of these data items need to be cautious when interpreting results, most notably with the following categories with the greatest potential for bias for each item:

- Item K9 (personal income range \$20,000 or more). The greatest potential for bias is for geographic areas with less than or equal to 10 percent below 150 percent of the poverty threshold.
- Item K12 (total income range). The greatest potential for bias is among ages 26-65 in rural areas.
- Item K7 (personal income range). The greatest potential for bias is among ages 45-65 in geographic areas with less than 2 percent limited English proficiency, especially in areas with greater than 20 percent with less than a high school education.
- Item K14 (total income range \$40,000 or more). The greatest potential for bias is among homeowners, age 26-65 in geographic areas with less than or equal to 10 percent below 150 percent of the poverty threshold.
- Item F4 (type of certificate, degree, or diploma). The greatest potential for bias is among ages 46-65, especially in geographic areas with less than or equal to 10 percent with less than a high school education, or more than 20 percent.
- Item K11 (total household income). The greatest potential for bias is among non-Hispanic Black adults aged 16-25.

Acknowledgments

Many people contributed to making this report possible, and the authors wish to thank all those who have assisted with various aspects of the report.

Page intentionally left blank.

Contents

	<u>Page</u>
EXECUTIVE SUMMARY	iii
ACKNOWLEDGMENTS	v
LIST OF TABLES.....	viii
LIST OF EXHIBITS.....	ix
INTRODUCTION	1
EVALUATING BIAS DUE TO UNIT NONRESPONSE.....	3
Weighted Response Rates	4
Bivariate Analysis	4
Multivariate Analysis	6
Potential for Nonresponse Bias Remaining After Weighting Procedures.....	7
Summary	9
EVALUATING BIAS DUE TO ITEM NONRESPONSE	10
Weighted Response Rates	10
Bivariate Analysis	10
Multivariate Analysis.....	11
Summary	12
CONCLUSIONS	13
REFERENCES	14
APPENDIX A – TABLES AND FIGURES	A-1
APPENDIX B – TECHNICAL NOTES	B-1

List of Tables

<u>Table</u>		<u>Page</u>
1	Weighted response rates, by stage of data collection: 2003	2
2	Variables used in unit nonresponse bias analysis by description, source, and values: 2003	3
A-1	Number of respondents, eligibles, weighted unit response rates for the ALL Background Questionnaire, by key characteristics: 2003	A-1
A-2	Sample distribution of unit respondents vs. eligibles for the ALL Background Questionnaire, by key characteristics: 2003	A-2
A-3	Estimates of unit nonresponse bias for the ALL Background Questionnaire, by key characteristics: 2003	A-3
A-4	Weighted response rates for the ALL Background Questionnaire, by response cell: 2003	A-5
A-5	Weighted response rates for the ALL Background Questionnaire: 2003	A-6
A-6	Weighted response rates for the ALL Background Questionnaire, by key characteristics: 2003	A-7
A-7	Weighted response rates for the ALL Background Questionnaire, by BQ items: 2003	A-10
A-8	Weighted response rates for the ALL Background Questionnaire, by key characteristics: 2003	A-11
A-9	Weighted response rates for the ALL Background Questionnaire, by key characteristics: 2003	A-19

List of Exhibits

<u>Exhibit</u>		<u>Page</u>
A-1	CHAID multivariate analysis of response indicators for the ALL Sample: 2003	A-27
A-2	CHAID multivariate analysis of item nonresponse for item K9 for the ALL Background Questionnaire: 2003.....	A-28
A-3	CHAID multivariate analysis of item nonresponse for item K12 for the ALL Background Questionnaire: 2003.....	A-28
A-4	CHAID multivariate analysis of item nonresponse for item K7 for the ALL Background Questionnaire: 2003	A-29
A-5	CHAID multivariate analysis of item nonresponse for item K14 for the ALL Background Questionnaire: 2003.....	A-29
A-6	CHAID multivariate analysis of item nonresponse for item F4 for the ALL Background Questionnaire: 2003.....	A-30
A-7	CHAID multivariate analysis of item nonresponse for item D43 for the ALL Background Questionnaire: 2003.....	A-31
A-8	CHAID multivariate analysis of item nonresponse for item K11 for the ALL Background Questionnaire: 2003.....	A-32

Page intentionally left blank.

1. Introduction

The 2003 Adult Literacy and Lifeskills Survey (ALL) was designed to assess the current status of literacy skills in America and to measure how these skill levels may have changed during the past decade in comparison with the 1994 International Adult Literacy Survey (IALS) results. ALL was designed and carried out according to internationally defined procedures, and instruments were developed to allow international-level comparisons across the participating countries. The development and management of the study were co-ordinated by Statistics Canada and the Educational Testing Service (ETS) in collaboration with the National Center for Education Statistics (NCES) of the United States Department of Education, the Organization for Economic Cooperation and Development (OECD), the Regional Office for Latin America and the Caribbean (OREALC) and the Institute for Statistics (UIS) of the United Nations Educational, Scientific and Cultural Organization (UNESCO). The sampling frame and target population are discussed in the ALL Sample Design and Data Collection section in appendix B.

The ALL estimates of literacy-related skills in the United States are subject to potential bias due to nonresponse at various levels of data collection. This report provides the results of a systematic analysis of bias due to the nonresponding persons in the ALL Sample. This report is in accordance with *National Center for Education Statistics (NCES): Standard 4-4*.

NCES Standard 4-4-1 (NCES, 2003) states that “any survey stage of data collection with a unit or item response rate less than 85 percent must be evaluated for the potential magnitude of nonresponse bias before the data or any analysis using the data may be released.” ALL had three stages of data collection where unit nonresponse occurred at the following:

- The screener;
- The background questionnaire (BQ); and
- The assessment.

Participation in the BQ was dependent upon the completion of the screener. Likewise, participation in the assessment was dependent upon completion of the BQ. Table 1 shows the final response rates for each data collection stage, including the overall response rate of the survey. Table 1 shows the BQ stage as being below the 85 percent response rate requirement. Therefore, the unit-level analysis is focused on the BQ stage. The key variables for the analysis are defined in chapter 2.

Sampling weights were used in the response rates computations. Weighted response rates estimate the coverage of the target adult population from the sample.

Table 1. Weighted response rates, by stage of data collection: 2003

Data collection stage	Weighted response rate (%)
Screener	86.7
Background questionnaire	80.0
Assessment	98.1
Overall	68.0

NOTE: The overall rate is the product of the rates from the three stages of data collection.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Adult Literacy and Lifeskills Survey (ALL), 2003.

Overview of the ALL Nonresponse Bias Report

Chapter 2 addresses the unit nonresponse in the ALL BQ. Data from the 2000 Census and the ALL Screener were used to describe the sample of respondents and nonrespondents and detect any differences between them. Chi-square tests and *t* tests with a significance level of $\alpha = 0.05$ were used. Chapter 2 also describes the attempts to reduce the impact of nonresponse through weighting adjustments.

Chapter 3 evaluates nonresponse bias with respect to BQ items with less than an 85 percent response rate. Using the base weights and variables known for respondents and nonrespondents, the bias due to item nonresponse is evaluated.

Appendix A includes all the tables and exhibits referred in chapters 2 and 3.

Appendix B continues with an overview of total survey error and the ALL sample design and procedures.

2. Evaluating Bias Due to Unit Nonresponse

A comparison of respondents and nonrespondents, using variables known for both groups, provides some indication of the potential for nonresponse bias in resulting survey estimates. Variables known for both respondents and nonrespondents to the BQ are displayed in table 2 and came from three sources:

- Census 2000 Public Law (PL) 94 county-level data;
- Census 2000 SF3A Block Group (BG) level data; and
- The ALL screener.

Table 2. Variables used in unit nonresponse bias analysis, by description, source, and values: 2003

Variable description	Source ¹	Values
Census region	PL-94	1: Northeast, 2: Midwest, 3: South, 4: West
Metropolitan statistical area (MSA) status	PL-94	1: MSA, 2: Non-MSA
Locale indicator ²	SF3A	1: Urban, 2: Suburban, 3: Rural
Majority own/rent	SF3A	1: Majority of BG rent, 2: Majority of BG own
Average household size	SF3A	1: less than or equal to 2.2, 2: greater than 2.2 and less than or equal to 2.8, 3: greater than 2.8
Percentage of persons 25 years or older with less than a high school diploma	SF3A	1: less than or equal to 10%, 2: greater than 10% and less than or equal to 20%, 3: greater than 20%
Percentage of persons 5-64 years old limited English proficiency ³	SF3A	1: 0%, 2: greater than 0% and less than or equal to 2%, 3: greater than 2%
Percentage below 150% of poverty ⁴	SF3A	1: less than or equal to 10%, 2: greater than 10% and less than or equal to 30%, 3: greater than 30%
Age	Screener	1: 16-25, 2: 26-35, 3: 36-45, 4: 46-65
Race/ethnicity ⁵	Screener	1: Hispanic, 2: Non-Hispanic Black only, 3: Other
Sex	Screener	1: Male, 2: Female

¹The SF3A and PL94 variables from the Census 2000 data files provide relevant statistics for the block group or the county where the sampled person resided at the time of sampling.

²This indicator was set to 'urban' if the largest percentage in the segment was inside urbanized areas; 'suburban' if the largest percentage was inside urban clusters; 'Rural' if the largest percentage was the rural population. The terms 'Urbanized areas', 'urban clusters', and 'rural' are Census Bureau defined terms.

³The exact language for this variable from the SF3A is "Percentage of persons 5-64 years old speaking English not well or not at all"

⁴Following the Office of Management and Budget's (OMB) Statistical Policy Directive 14, the Census Bureau uses a set of money income thresholds that vary by family size and composition to determine who is in poverty. The variable used in this report is supplied from the SF3 file, and it represents those with income below 150% of the poverty threshold (U.S. Census Bureau, 2006).

⁵All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as Other include non-Hispanics of all other races, including White, Asian, American Indian/Alaska Native, or multiracial.

Note: BG stands for block group. PL-94 stands for Public Law-94. SF3A stands for Summary File 3A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Adult Literacy and Life Skills Survey (ALL), 2003.

Survey base weights were used to account for the unequal within-household selection probabilities of sample persons, and replicate weights were used to adequately reflect the effect of multistage cluster sampling on variance estimates. Weights were used in all stages of the nonresponse bias analysis.

The Weighted Response Rates section presents response rates for selected analysis variables, revealing areas where the most potential for bias exists, prior to weight adjustments. The Bivariate Analysis section contains chi-square tests that may detect a significant relationship between a response indicator and the analysis variable of interest. Next is a multivariate analysis of the relationship between the response indicator and the analysis variables, followed by a section that addresses the effect of weighting procedures on reducing nonresponse bias.

Weighted Response Rates

Table A-1 contains unit response rates for the BQ. The overall BQ weighted response rate was 80 percent. Response rates were calculated according to Statistics Canada's guidelines for each category of the analysis variables in table 2. The numerator of the response rate consisted of all completed cases. The denominator included all cases except those with hearing impaired, physical disability, or other disability status that precluded participation. There were 12 cases that fell into these categories. The 12 cases were considered nonrespondents for the bivariate and multivariate analysis described below. This is consistent with other NCES projects, including the National Adult Literacy Survey (Kirsch et al 2001), where those with disabilities were treated like other nonrespondents, and provides a slightly more conservative approach to the analysis.

Bivariate Analysis

Table A-2 shows the domain percentages of the variables described in table 2 for respondents and eligibles. A comparison between domain estimates for respondents and eligibles shows little absolute difference. The differences are all within 3 percentage points. Therefore, it appears that there is minimal potential for nonresponse bias. To formally test for independence between the response indicator and analysis variable, a Rao-Scott (RS3) chi-square test (Rao and Thomas 2003) was computed using WesVar (Westat 2002). The following variables were not found to be significantly related to

response status at the 5 percent α level: locale indicator, majority own/rent, average household size, percentage of persons 25 years or older with less than a high school diploma, and percentage of persons 5 to 64 years old with limited English proficiency.

The chi-square test, however, indicates a significant relationship of census region to response status, with a lower percentage in the Northeast among respondents than eligibles. In addition, metropolitan statistical area (MSA) status appears to be significantly related to response status, with MSAs responding at a lower rate than non-MSAs. The chi-square test for the percentage below 150 percent of poverty is significant at the $\alpha = 0.05$ level. The higher response rate in high poverty areas may have been due to the offered incentive. Age, sex, and race/ethnicity were also found to be significantly related to response status, with older age groups, males, and other race/ethnicity responding at lower levels. All variables found to be significant in the bivariate analysis were used in nonresponse weighting adjustments. The effects of the adjustments are presented in the section on potential for nonresponse bias remaining after weighting procedures.

Table A-3 contains estimates of bias and relative bias for the subgroup percentages of the variables described in table 2. Bias can be expressed as follows:

$$Bias(\bar{y}_R) = (1 - W_R)(\bar{Y}_R - \bar{Y}_N),$$

where W_R is the weighted unit response rate, \bar{Y}_R is the weighted estimate of the subgroup percentage for respondents, and \bar{Y}_N is the weighted estimate of the subgroup percentage for nonrespondents. In practice, bias was estimated as the difference between the respective estimates for the respondents and the eligible sample. A t test was performed to determine whether the bias was significantly different from zero. The results of the t tests are consistent with the above chi-square analysis, in addition, the biases for the percentage of average household size between 2.2 and 2.8, for ages 16-25, and non-Hispanic Blacks and Hispanics are significant at the $\alpha = 0.05$ level. Another measure, the relative bias, was computed as the bias divided by the estimate from the eligible sample. The relative bias is a measure of the size of the bias compared to the eligible sample estimate.

The standard error used in the calculation of the test statistic is small relative to the standard error of the difference between two independent samples, since the standard error computation accounts for the covariance between two highly overlapping samples (respondents and eligible samples). Although a valid statistical test, this may result in statistical significance for seemingly small differences. Therefore

in summary, even though there are statistically significant results as described above, the relative magnitude of the differences was small, resulting in minimal potential for bias. The relative bias estimates are all less than 10 percent, except in the Northeast, which is 15 percent.

Multivariate Analysis

The bivariate analysis, described in the preceding section, is useful in explaining the relationship of response status to each variable individually. A multivariate analysis is useful in showing relationships among a number of variables. One approach is to provide a Chi-squared Automatic Interaction Detector (CHAID) analysis. CHAID is a classification algorithm that uses chi-square tests to divide a sample into subgroups that best explain differential response rates.

The analysis in CHAID begins by dividing the sample into two or more groups based on the categories of the best predictor of differential response rates. Each of these groups is divided into smaller subgroups based on the best available predictor at each level. The splitting process continues until either no significant predictor remains or the minimum cell size requirement is met. The CHAID software displays the final subgroups in the form of a tree diagram whose branches (nodes) correspond to the groups. The resulting classification tree reveals the domains, as defined by combinations of variables, with the most differential response rates, thereby identifying domains with the highest potential for nonresponse bias.

CHAID was run with response status as the dependent variable and the variables described in table 2 as the independent variables. The resulting tree is presented in exhibit A-1 and summarized in table A-4. Twenty-two cells were formed with weighted response rates ranging from 58.7 percent to 97.4 percent. The tree in exhibit A-1 shows the lowest response rates are in MSAs in the Northeast. For Northeast MSAs, the subgroup with the greatest potential for bias due to nonresponse is in areas with average household size greater than 2.8 (58.7 percent response rate). While the CHAID tree is useful to dissect the sample into fine groups of persons with response patterns as different as possible, one should be cautious when using it since CHAID does not take into account the complex design of the sample. Consequently, the significance level of the test may be lower than indicated. If the appropriate significance level could be used, then the tree may result in fewer significant response cells. Thus, the tree shown in exhibit A-1 is a conservative picture since any indication of nonresponse bias shown by the CHAID results may be overstated.

Logistic regression models are also useful in identifying significant effects on response propensity. Response status was used as the binary dependent variable and the variables described in table 2 as the predictors. The model fitted had the following form:

$$\log\left[\frac{P(\text{Response})}{1-P(\text{Response})}\right] = A + \sum B_i X_{ij} + \sum C_i Y_{ij} ,$$

where the X_{ij} 's are indicator variables for variables used in the BQ nonresponse weighting adjustments: age, race/ethnicity, sex, MSA status, percentage below 150 percent of poverty, and region; and the Y_{ij} 's are indicator variables for variables not used in the BQ nonresponse weighting adjustments: average household size, majority own/rent, percentage less than high school, percentage with limited English proficiency, and locale. An F test was performed for the Y variables as a group to determine if they were significantly related to response propensity, after accounting for the weighting variables.

Results of the logistic regression are presented in table A-5. The overall fit of the model is significant at the 5 percent α level, with p value equal to 0.016. The test of the nonweighting variables does indicate a significant relationship to response propensity (p value equals 0.009), however when average household size was excluded from the test, the remaining nonweighting variables were not significant (p value equals 0.596).

The results of CHAID and the logistic regression were fairly consistent, with the nonweighting variables majority own/rent, percentage with limited English proficiency, locale, and percentage less than high school entering the CHAID model at a later stage than the weighting variables or not at all. Although significant in the logistic regression, household size also entered CHAID at a later stage but was only used in forming 5 of the 22 cells.

Potential for Nonresponse Bias Remaining After Weighting Procedures

The weighting procedures, as described in appendix B, attempt to reduce the potential for nonresponse bias by creating nonresponse adjustment classes for which the respondents' literacy-related characteristics are similar to the nonrespondents. Subsequent to the nonresponse adjustment, the raking

adjustment also has some impact on reducing the bias due to nonresponse. The extent of the reduction in nonresponse bias depends on the correlation of the weighting class variables with literacy scores.

Table A-6 shows and tests the change in the percentage distribution of the sample cases after each BQ weighting stage, using the analysis variables in table 2. The following *t* test comparisons were made:

- Comparison of percentage distributions from BQ base weights for the total eligible sample of persons with the BQ base weights for the BQ respondents only to check for differences due to nonresponse to the BQ;
- Comparison of percentage distributions from BQ base weights for the total eligible sample of persons with that from the BQ nonresponse adjusted weights for respondents to check for differences even after the nonresponse adjustment process to the BQ;
- Comparison of percentage distributions from BQ nonresponse adjusted weights for respondents with that from the BQ raked weights for respondents to check for differences that may have been introduced through the initial raking procedure; and
- Comparison of percentage distributions from BQ base weights for the total eligible sample of persons with that from the BQ final weights for respondents, resulting after a trimming adjustment and another round of raking to recalibrate the weights. This comparison checks for differences that may have been introduced through the BQ weighting steps.

When base weights (prior to nonresponse adjustments) are used, potential for nonresponse bias appears for the following domains: Northeast region, West region, non-MSAs, MSAs, greater than 30 percent below 150 percent of the poverty threshold, age group 16 to 25, age group 46 to 65, males, females, Hispanics, non-Hispanic Blacks, and other race/ethnicity. Through the nonresponse adjustment, the potential for bias was reduced (p-values moved from significant to nonsignificant) in all but the following domains: age group 16 to 25, males, females, and other race/ethnicity. However, two domains relating to average household size show significant bias created by the nonresponse adjustments. While the differences in the percentage distributions of these domains remain statistically significant, they are not large enough to be meaningful (see the bivariate analysis section for an explanation). All significant differences with the base weights were eliminated after the raking adjustment, that is, weight adjustments eliminated significant differences between the initial sample selected and the final sample of completed cases, indicating negligible amount of bias remaining in the ALL data on the set of variables analyzed.

Summary

Without incorporating the weighting adjustments, the unit-level analyses results still show little difference between respondents and nonrespondents. Subgroups relating to average household size, age, sex, MSA status, and race/ethnicity were found to have significant bias. However, even where there are statistically significant results, the relative magnitudes of the differences were small, resulting in minimal potential for bias, and thus minimal impact on resulting literacy scores. The analysis also showed that the weighting adjustments were highly effective in reducing the bias. The general conclusion is that the potential amount of nonresponse bias due to unit nonresponse in ALL is likely to be negligible on the set of variables analyzed.

3. Evaluating Bias Due to Item Nonresponse

There are numerous reasons for item nonresponse: the unit respondent did not know the answer for the item or did not wish to respond, the interview was interrupted leaving items in the later part of the questionnaires blank, or responses were later found to be internally inconsistent during the editing process and were blanked out. Response rates for each item in the BQ were computed, and a discussion of response rates is provided in the first section of this chapter. For items with less than 85 percent response rate, a nonresponse bias analysis was conducted. The second section of this chapter presents a bivariate analysis, comparing the distribution of item respondents and eligibles. The third section contains a multivariate analysis of the relationship between item response status and variables known for both respondents and nonrespondents.

Weighted Response Rates

Response rates were calculated for all 349 items in the BQ. In accordance with NCES standard 1-3-5, the numerator of the item response rate consisted of item respondents; the denominator contained all unit respondents, excluding those with a valid skip for the item. Weighted item response rates ranged from 21.7 to 100.0 percent, with a median of 99.9 percent. Eight items had response rates under 85 percent. The set of items is the same whether unweighted or weighted response rates are used. The low response rate items are displayed in table A-7. Seven of the eight items are income related.

Bivariate Analysis

For each low item response rate in table A-7, the distribution of item respondents and eligibles was compared on related BQ items with response rates of 99.9 percent or higher. To test for independence between the response indicator and high response rate variable, a RS3 chi-square test was computed using WesVar. Base weights were used in the analysis. The results are in table A-8.

The variables Enjoyed math in school and sex are not significantly related to response status of any of the low response rate items. The variable Read letters or e-mails for job is significantly related to the response status of five of the eight low response rate variables, and age is significant for four variables. While the response indicator for D42 (gross monthly salary) and D43 (annual personal net income) do not appear to be significantly related to the selected items, the other low response rate items show some slight potential for nonresponse bias, particularly F4 (type of certificate, degree, or diploma) and K11 (total household income). However, the sample size for many of these items is quite small due to skip patterns. The sample sizes for each item are provided in table A-7.

To understand the magnitude of the potential bias, estimates of bias were computed as described in the bivariate analysis section in chapter 2. The bias estimates are displayed in table A-9. *T* tests were performed to determine whether the bias was significantly different from zero. The results of the *t* tests are consistent with the chi-square analysis.

Multivariate Analysis

For the multivariate analysis of item nonresponse, CHAID was used to divide the sample into subgroups that best explain differential response rates. The resulting classification trees reveal the domains, as defined by combinations of variables, with the most differential response rates, thereby leading to domains with the highest potential for nonresponse bias. Item response status was used as the dependent variable, and table 2 variables were the predictors.

The trees for the low response rate items are displayed in exhibits A-2 to A-8. No tree was produced for item D42 (gross monthly salary), since the total sample size is only 44 for this item. For item K9 (personal income range - \$20,000 or more), differential response rates are revealed with a response rate of 44.4 percent for the less than or equal to 10 percent below 150 percent of poverty subgroup compared to 76.0 percent for the greater than 10 percent below 150 percent of poverty subgroup. For item K12 (total income range), age, region, and locale indicator were used to form the response cells, with response rates ranging from 52.4 percent to 100.0 percent. Item K7 (personal income range) shows differential response rates by age, percentage of persons 5 to 64 years old with limited English proficiency, and percentage of persons 25 years or older with less than a high school diploma. Age category 46 to 65, with less than 2 percent with limited English proficiency, and greater than 20 percent with less than a high school diploma shows the greatest potential for nonresponse bias for this item, with a response rate of only 38.0 percent.

Item K14 (total income range - \$40,000 or more) response rates differ by majority own or rent, age, and percentage below 150 percent of poverty. The following variables were used to form response cells for item F4 (type of certificate, degree, or diploma): age, region, percentage of persons 25 years or older with less than a high school diploma, race/ethnicity, locale indicator, and percentage below 150 percent of poverty. Higher age groups show lower response rates. The response rate for age category 46 to 65 is only 39.8 percent. For item D43 (annual personal net income), response rates range from 97.2 percent in the Northeast to 65.5 percent in non-MSAs in the Midwest/South/West where the majority own. Finally, item K11 (total household income) shows a low response rate for non-Hispanic Blacks ages 16 to 25 (51.7 percent). Although not typically done, weighting adjustments at the item-level could be

used to reduce the bias. Item-level weights are usually created when the potential for bias for a critical item is not negligible. Imputation is another alternative to reduce item nonresponse bias.

Summary

The item nonresponse analysis revealed some low response rates in certain domains and some indication that a potential for bias may exist. There were no weighting adjustments at the item level or imputation procedures completed on these items that may have reduced the bias due to item nonresponse. Users of these data items need to be cautious when interpreting results, most notably with the following categories with the greatest potential for bias for each item:

- Item K9 (personal income range \$20,000 or more). The greatest potential for bias is for geographic areas with less than or equal to 10 percent below 150 percent of the poverty threshold.
- Item K12 (total income range). The greatest potential for bias is among ages 26-65 in rural areas.
- Item K7 (personal income range). The greatest potential for bias is among ages 45-65 in geographic areas with less than 2 percent limited English proficiency, especially in areas with greater than 20 percent with less than a high school education.
- Item K14 (total income range \$40,000 or more). The greatest potential for bias is among homeowners, age 26-65 in areas with less than or equal to 10 percent below 150 percent of the poverty threshold.
- Item F4 (type of certificate, degree, or diploma). The greatest potential for bias is among ages 46-65, especially in geographic areas with less than or equal to 10 percent with less than a high school education, or more than 20 percent.
- Item K11 (total household income). The greatest potential for bias is among non-Hispanic Black adults aged 16-25.

4. Conclusions

The general conclusion from the unit-level analyses is that the potential amount of nonresponse bias due to unit nonresponse in ALL is likely to be negligible even where there is statistically significant bias. The standard error involved in the calculation of the test statistic is small relative to the standard error that would be computed for the difference between two independent samples, since the standard error computation accounts for the covariance between two highly overlapping samples (respondent and eligible samples). Although a valid statistical test, this may result in statistical significance for seemingly small differences. The magnitude of the relative bias was small, resulting in minimal potential for bias, and thus minimal impact on resulting literacy scores.

The unit-level analysis also showed that the weighting adjustments were highly effective in reducing the bias. All significant differences between the full sample and respondents in the nonresponse bias analysis variables were eliminated through the weighting adjustments.

The item nonresponse analysis revealed some low response rates in certain domains and some indication that a potential for bias may exist. There were no weighting adjustments at the item level or imputation procedures completed on these items that may have reduced the bias due to item nonresponse. Users of these data items need to be cautious when interpreting results, most notably with the following categories with the greatest potential for bias for each item:

- Item K9 (personal income range \$20,000 or more). The greatest potential for bias is for geographic areas with less than or equal to 10 percent below 150 percent of the poverty threshold.
- Item K12 (total income range). The greatest potential for bias is among ages 26-65 in rural areas.
- Item K7 (personal income range). The greatest potential for bias is among ages 45-65 in geographic areas with less than 2 percent limited English proficiency, especially in areas with greater than 20 percent with less than a high school education.
- Item K14 (total income range \$40,000 or more). The greatest potential for bias is among homeowners, age 26-65 in geographic areas with less than 10 percent below 150 percent of the poverty threshold.
- Item F4 (type of certificate, degree, or diploma). The greatest potential for bias is among ages 46-65, especially in geographic areas with less than or equal to 10 percent with less than a high school education, or more than 20 percent.
- Item K11 (total household income). The greatest potential for bias is among non-Hispanic Black adults aged 16-25.

References

- Kirsch, I., Yamamoto, K., Norris, N., Rock, D., Jungeblut, A, O'Reilly, P., Berlin, M., Mohadjer, L., Waksberg, J., Goksel, H., Burke, J., Rieger, S., Green, J., Klein, M., Campbell, A, Jenkins, L., Kolstad, A., Mosenthal, P., and Baldi, S. (2001). *Technical Report and Data File User's Manual for the 1992 National Adult Literacy Survey*. U.S. Department of Education, Office of Educational Research and Improvement.
- Lessler, J., and Kalsbeek, W. (1992). *Nonsampling Error in Surveys*. New York: John Wiley & Sons.
- NCES (2003). *NCES Statistical Standards*. NCES 2003-601. Washington, DC.
- Rao, J.N.K., and Thomas, D.R. (2003). Analysis of Categorical Response Data from Complete Survey: An Appraisal and Update, in R. L. Chambers and C. J. Skinner (Eds.) *Analysis of Survey Data* (pp. 85 – 108). West Sussex, England: John Wiley and Sons.
- U.S. Census Bureau (2006). *Summary File 3 Technical Documentation*. Washington, DC.
- Westat (2002). *WesVar 4.2 User's Guide*. Rockville, MD: Westat, Inc.

Appendix A – Tables and Figures

Table A-1. Number of respondents, eligibles, weighted unit response rates for the ALL Background Questionnaire, by key characteristics: 2003

Analysis variable	Number of respondents	Number of eligibles	Weighted response rate
Total	3,420	4,243	80.0
Region			
Northeast	534	782	67.7
Midwest	793	959	82.8
South	1,199	1,468	80.4
West	894	1,034	86.0
Metropolitan Statistical Area status			
Non-Metropolitan Statistical Area	786	933	83.5
Metropolitan Statistical Area	2,634	3,310	79.0
Locale ¹			
Urban	2,226	2,786	79.1
Suburban	450	530	83.9
Rural	744	927	80.3
Majority own/rent			
Rent	936	1,138	81.5
Own	2,484	3,105	79.5
Average household size			
less than or equal to 2.2	680	864	77.8
greater than 2.2 and less than or equal to 2.8	1,806	2,189	82.1
greater than 2.8	934	1,190	77.7
Percent less than high school			
less than or equal to 10%	1,085	1,368	79.0
greater than 10% and less than or equal to 20%	1,062	1,343	78.8
greater than 20%	1,273	1,532	82.0
Percent limited English proficient			
0%	1,025	1,278	79.8
greater than 0% and less than 2%	1,168	1,447	80.2
greater than or equal to 2%	1,227	1,518	79.9
Percent below 150% of poverty			
less than or equal to 10%	974	1,245	77.9
greater than 10% and less than or equal to 30%	1,566	1,976	78.9
greater than 30%	880	1,022	84.8
Age			
16–25	710	837	84.2
26–35	754	923	81.4
36–45	768	976	77.8
46–65	1,188	1,507	77.8
Sex			
Male	1,582	2,031	77.6
Female	1,838	2,212	82.3
Race/ethnicity ²			
Hispanic	468	544	85.6
Black, non-Hispanic	414	470	86.9
Other, non-Hispanic	2,538	3,229	78.1

¹This indicator was set to 'urban' if the largest percentage in the segment was inside urbanized areas; 'suburban' if the largest percentage was inside urban clusters; 'Rural' if the largest percentage was the rural population. The terms 'Urbanized areas', 'urban clusters', and 'rural' are Census Bureau defined terms.

²All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as Other include non-Hispanics of all other races, including White, Asian, American Indian/Alaska Native, or multiracial.

NOTE: Cases with hearing impaired, physical disability, or other disability status are excluded from response rate calculations.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Adult Literacy and Lifeskills Survey (ALL), 2003.

Table A-2. Sample distribution of unit respondents vs. eligibles for the ALL Background Questionnaire, by key characteristics: 2003

Analysis variable	Respondents		Eligibles		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	p value
Region						
Northeast	15.5	2.50	18.3	2.36	31.95	0.000
Midwest	22.9	2.54	22.1	2.24	†	†
South	35.0	1.66	34.9	1.50	†	†
West	26.6	2.47	24.7	2.21	†	†
Metropolitan Statistical Area status						
Non-Metropolitan Statistical Area	22.7	2.63	21.8	2.40	7.26	0.007
Metropolitan Statistical Area	77.3	2.63	78.2	2.40	†	†
Locale¹						
Urban	64.6	2.67	65.3	2.45	3.78	0.120
Suburban	13.0	2.00	12.4	1.83	†	†
Rural	22.4	2.11	22.3	2.02	†	†
Majority own/rent						
Rent	25.7	2.48	25.2	2.40	1.14	0.287
Own	74.3	2.48	74.8	2.40	†	†
Average household size						
less than or equal to 2.2	18.1	1.87	18.6	1.81	4.73	0.081
greater than 2.2 and less than or equal to 2.8	52.9	2.50	51.5	2.80	†	†
greater than 2.8	29.1	2.92	29.9	3.10	†	†
Percent less than high school						
less than or equal to 10%	33.1	2.71	33.4	2.73	2.70	0.257
greater than 10% and less than or equal to 20%	30.9	1.81	31.4	1.75	†	†
greater than 20%	36.1	3.01	35.2	3.09	†	†
Percent limited English proficient						
0%	29.9	2.68	29.9	2.51	0.06	0.963
greater than 0% and less than 2%	34.7	3.19	34.6	3.09	†	†
greater than or equal to 2%	35.4	3.78	35.5	3.64	†	†
Percent below 150% of poverty						
less than or equal to 10%	30.7	2.65	31.5	2.79	8.59	0.011
greater than 10% and less than or equal to 30%	44.9	3.27	45.4	3.27	†	†
greater than 30%	24.4	2.95	23.1	2.76	†	†
Age						
16–25	23.7	1.38	22.4	1.25	17.92	0.000
26–35	21.2	0.80	20.8	0.77	†	†
36–45	22.2	0.84	22.9	0.79	†	†
46–65	32.9	1.29	33.9	1.16	†	†
Sex						
Male	48.1	1.08	49.6	0.99	22.77	0.000
Female	51.9	1.08	50.4	0.99	†	†

See notes at end of table.

Table A-2. Sample distribution of unit respondents vs. eligibles for the ALL Background Questionnaire, by key characteristics: 2003—Continued

Analysis variable	Respondents		Eligibles		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	p value
Race/ethnicity ²						
Hispanic	14.3	1.79	13.3	1.61	19.11	0.000
Black, non-Hispanic	11.1	1.82	10.2	1.62	†	†
Other, non-Hispanic	74.6	2.08	76.4	1.94	†	†

† Not applicable.

¹This indicator was set to 'urban' if the largest percentage in the segment was inside urbanized areas; 'suburban' if the largest percentage was inside urban clusters; 'Rural' if the largest percentage was the rural population. The terms 'Urbanized areas', 'urban clusters', and 'rural' are Census Bureau defined terms.

²All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as Other include non-Hispanics of all other races, including White, Asian, American Indian/Alaska Native, or multiracial.

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Adult Literacy and Lifeskills Survey (ALL), 2003.

Table A-3. Estimates of unit nonresponse bias for the ALL Background Questionnaire, by key characteristics: 2003

Analysis variable	Respondents percent	Eligibles percent	Bias			
			Estimate	Standard error	Percent Relative Bias	
Region						
Northeast	15.5	18.3	-2.8	0.53	0.000	-15.3
Midwest	22.9	22.1	0.8	0.61	0.187	3.6
South	35	34.9	0.1	0.49	0.842	0.3
West	26.6	24.7	1.8	0.48	0.001	7.3
Metropolitan Statistical Area status						
Non-Metropolitan Statistical Area	22.7	21.8	0.9	0.34	0.010	4.1
Metropolitan Statistical Area	77.3	78.2	-0.9	0.34	0.010	-1.2
Locale ¹						
Urban	64.6	65.3	-0.7	0.38	0.074	-1.1
Suburban	13.0	12.4	0.6	0.37	0.108	4.8
Rural	22.4	22.3	0.1	0.31	0.747	0.4
Majority own/rent						
Rent	25.7	25.2	0.5	0.44	0.298	2.0
Own	74.3	74.8	-0.5	0.44	0.298	-0.7
Average household size						
less than or equal to 2.2	18.1	18.6	-0.5	0.39	0.215	-2.7
greater than 2.2 and less than or equal to 2.8	52.9	51.5	1.4	0.65	0.044	2.7
greater than 2.8	29.1	29.9	-0.9	0.6	0.155	-3.0
Percent less than high school						
less than or equal to 10%	33.1	33.4	-0.3	0.53	0.540	-0.9
greater than 10% and less than or equal to 20%	30.9	31.4	-0.5	0.48	0.284	-1.6
greater than 20%	36.1	35.2	0.9	0.52	0.112	2.6
Percent limited English proficient						
0%	29.9	29.9	#	0.56	0.977	#
greater than 0% and less than 2%	34.7	34.6	0.1	0.69	0.841	0.3
greater than or equal to 2%	35.4	35.5	-0.1	0.55	0.826	-0.3

See notes at end of table.

Table A-3. Estimates of unit nonresponse bias for the ALL Background Questionnaire, by key characteristics: 2003—Continued

Analysis variable	Respondents percent	Eligibles percent	Bias			Percent relative bias
			Estimate	Standard error	<i>p</i> value	
Percent below 150% of poverty						
less than or equal to 10%	30.7	31.5	-0.8	0.52	0.129	-2.5
greater than 10% and less than or equal to 30%	44.9	45.4	-0.5	0.58	0.379	-1.1
greater than 30%	24.4	23.1	1.3	0.38	0.001	5.6
Age						
16–25	23.7	22.4	1.3	0.35	0.001	5.8
26–35	21.2	20.8	0.4	0.25	0.106	1.9
36–45	22.2	22.9	-0.6	0.36	0.088	-2.6
46–65	32.9	33.9	-1	0.41	0.016	-2.9
Sex						
Male	48.1	49.6	-1.5	0.31	0.000	-3.0
Female	51.9	50.4	1.5	0.31	0.000	3.0
Race/ethnicity ²						
Hispanic	14.3	13.3	0.9	0.29	0.003	6.8
Black, non-Hispanic	11.1	10.2	0.9	0.31	0.009	8.8
Other, non-Hispanic	74.6	76.4	-1.8	0.32	0.000	-2.4

Rounds to zero

¹This indicator was set to 'urban' if the largest percentage in the segment was inside urbanized areas; 'suburban' if the largest percentage was inside urban clusters; 'Rural' if the largest percentage was the rural population. The terms 'Urbanized areas', 'urban clusters', and 'rural' are Census Bureau defined terms.

²All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as Other include non-Hispanics of all other races, including White, Asian, American Indian/Alaska Native, or multiracial.

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Adult Literacy and Lifeskills Survey (ALL), 2003.

Table A-4. Weighted response rates for the ALL Background Questionnaire, by response cell: 2003

Response cell	Description	Weighted response rate (%)
Overall		79.7
1	Northeast, Non-Metropolitan Statistical Area (MSA)	88.9
2	Northeast, MSA, Average household size less than or equal to 2.8	68.1
3	Northeast, MSA, Average household size greater than 2.8	58.7
4	Midwest and South, Hispanic or non-Hispanic Black, Male, Average household size less than or equal to 2.2	79.8
5	Midwest and South, Hispanic or non-Hispanic Black, Male, Average household size greater than 2.2 and less than or equal to 2.8	91.1
6	Midwest and South, Hispanic or non-Hispanic Black, Male, Average household size greater than 2.8	78.0
7	Midwest and South, Hispanic or non-Hispanic Black, Female, Age 16-45	94.4
8	Midwest and South, Hispanic or non-Hispanic Black, Female, Age 46-65	82.0
9	Midwest and South; Other, non-Hispanic; Age 16-35; Male; Rent	70.0
10	Midwest and South; Other, non-Hispanic; Age 16-35; Male; Own	82.1
11	Midwest; Other, non-Hispanic; Female; Age 16-25	97.4
12	Midwest; Other, non-Hispanic; Female; Age 26-35	90.2
13	South; Other, non-Hispanic; Female; Percent limited English proficient less than 2%	86.0
14	South; Other, non-Hispanic; Female; Percent limited English proficient greater than or equal to 2%	70.3
15	Midwest; Other, non-Hispanic; Female; Age 36-65; Percent below 150% of poverty less than or equal to 30%	77.3
16	South; Other, non-Hispanic; Female; Age 36-65; Percent below 150% of poverty less than or equal to 30%	71.5
17	Midwest and South; Other, non-Hispanic; Age 36-65; Percent below 150% of poverty greater than 30%	84.7
18	West, Rent	88.6
19	West, Own, Age 16-25	88.7
20	West, Own, Age 26-45	79.6
21	West, Own, Age 46-65, Urban or Suburban	90.5
22	West, Own, Age 46-65, Rural	78.5

NOTE: The response cells were formed using the Chi-square Automated Interaction Detector, as shown in exhibit A-1.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Adult Literacy and Lifeskills Survey (ALL), 2003.

Table A-5. Weighted response rates for the ALL Background Questionnaire: 2003

Test	<i>F</i> statistic	Numerator degrees of freedom	Denominator degrees of freedom	<i>p</i> value
Overall fit	3.89	21	10	0.016
Average household size	6.77	2	29	0.004
Majority own/rent	0.32	1	30	0.578
Percent less than high school	0.44	2	29	0.647
Percent limited English proficient	2.08	2	29	0.143
Locale ¹	0.12	2	29	0.887
Age	4.47	3	28	0.011
Race/ethnicity ²	10.79	2	29	0.000
Sex	19.43	1	30	0.000
Metropolitan Statistical Area status	0.03	1	30	0.871
Percent below 150% of poverty	3.22	2	29	0.055
Region	6.65	3	28	0.002
$Y_{ij} = 0^3$	3.41	9	22	0.009
$Y_{ij'} = 0^4$	0.80	7	24	0.596

¹This indicator was set to 'urban' if the largest percentage in the segment was inside urbanized areas; 'suburban' if the largest percentage was inside urban clusters; 'Rural' if the largest percentage was the rural population. The terms 'Urbanized areas', 'urban clusters', and 'rural' are Census Bureau defined.

²All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as Other include non-Hispanics of all other races, including White, Asian, American Indian/Alaska Native, or multiracial.

³ Y_{ij} includes all nonweighting variables.

⁴ $Y_{ij'}$ includes all nonweighting variables except household size.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Adult Literacy and Lifeskills Survey (ALL), 2003.

Table A-6. Weighted response rates for the ALL Background Questionnaire, by key characteristics: 2003

Analysis variable	Total sample		BQ respondents		SE ¹	p value ¹
	Population 16-65 (BQ base weight)	Percentage	Population 16-65 (BQ base weight)	Percentage		
Total	169,957,000	100.0	135,519,000	100.0	†	†
Region						
Northeast	31,183,000	18.4	21,069,000	15.6	0.54	0.0000
Midwest	37,500,000	22.1	31,036,000	22.9	0.61	0.1800
South	59,265,000	34.9	47,391,000	35.0	0.49	0.8418
West	42,009,000	24.7	36,024,000	26.6	0.48	0.0005
Metropolitan Statistical Area status						
Non-Metropolitan Statistical Area	37,038,000	21.8	30,820,000	22.7	0.36	0.0122
Metropolitan Statistical Area	132,920,000	78.2	104,699,000	77.3	0.36	0.0122
Locale ³						
Urban	111,023,000	65.3	87,550,000	64.6	0.40	0.0799
Suburban	21,037,000	12.4	17,610,000	13.0	0.37	0.1025
Rural	37,897,000	22.3	30,359,000	22.4	0.32	0.7488
Majority own/rent						
Rent	42,805,000	25.2	34,774,000	25.7	0.44	0.2940
Own	127,152,000	74.8	100,746,000	74.3	0.44	0.2940
Average household size						
less than or equal to 2.2	31,556,000	18.6	24,480,000	18.1	0.41	0.2245
greater than 2.2 and less than or equal to 2.8	87,504,000	51.5	71,654,000	52.9	0.68	0.0507
greater than 2.8	50,897,000	30.0	39,385,000	29.1	0.61	0.1570
Percent less than high school						
less than or equal to 10%	56,750,000	33.4	44,801,000	33.1	0.54	0.5424
greater than 10% and less than or equal to 20%	53,387,000	31.4	41,847,000	30.9	0.50	0.2916
greater than 20%	59,820,000	35.2	48,871,000	36.1	0.54	0.1227
Percent limited English proficient						
0%	50,849,000	29.9	40,523,000	29.9	0.57	0.9768
greater than 0% and less than 2%	58,740,000	34.6	47,029,000	34.7	0.70	0.8401
greater than or equal to 2%	60,368,000	35.5	47,967,000	35.4	0.56	0.8254
Percent below 150% of poverty						
less than or equal to 10%	53,538,000	31.5	41,575,000	30.7	0.54	0.1367
greater than 10% and less than or equal to 30%	77,190,000	45.4	60,836,000	44.9	0.58	0.3718
greater than 30%	39,229,000	23.1	33,108,000	24.4	0.38	0.0015
Age						
16-25	38,102,000	22.4	32,099,000	23.7	0.35	0.0012
26-35	35,298,000	20.8	28,727,000	21.2	0.26	0.1033
36-45	38,910,000	22.9	30,149,000	22.2	0.36	0.0831
46-65	57,647,000	33.9	44,544,000	32.9	0.41	0.0165
Sex						
Male	84,219,000	49.6	65,146,000	48.1	0.31	0.0000
Female	85,738,000	50.5	70,373,000	51.9	0.31	0.0000
Race/ethnicity ⁴						
Hispanic	22,663,000	13.3	19,359,000	14.3	0.30	0.0038
Black, non-Hispanic	17,379,000	10.2	15,024,000	11.1	0.31	0.0088
Other, non-Hispanic	129,915,000	76.4	101,137,000	74.6	0.34	0.0000

See notes at end of table.

Table A-6. Weighted response rates for the ALL Background Questionnaire, by key characteristics:
2003—Continued

Analysis variable	Nonresponse adjustment				Initial raking			
	Estimate	Percentage	SE ¹	p value ¹	Estimate	Percentage	SE ²	p value ²
Total	169,957,000	100.0	†	†	188,173,000	100.0	†	†
Region								
Northeast	31,533,000	18.6	0.22	0.3545	35,730,000	19.0	2.56	0.8662
Midwest	37,151,000	21.9	0.22	0.3545	42,734,000	22.7	2.43	0.7282
South	59,265,000	34.9	0.00	0.8248	66,654,000	35.4	1.50	0.7159
West	42,009,000	24.7	0.00	0.8946	43,055,000	22.9	2.21	0.4133
Metropolitan Statistical Area status								
Non-Metropolitan Statistical Area	37,200,000	21.9	0.16	0.5578	33,472,000	17.8	2.42	0.1002
Metropolitan Statistical Area	132,757,000	78.1	0.16	0.5578	154,701,000	82.2	2.42	0.1002
Locale								
Urban	110,599,000	65.1	0.32	0.4484	128,698,000	68.4	1.94	0.0970
Suburban	21,126,000	12.4	0.27	0.8485	21,247,000	11.3	0.87	0.1983
Rural	38,232,000	22.5	0.34	0.5661	38,228,000	20.3	1.18	0.0746
Majority own/rent								
Rent	42,428,000	25.0	0.31	0.4857	48,485,000	25.8	0.69	0.2543
Own	127,529,000	75.0	0.31	0.4857	139,688,000	74.2	0.69	0.2543
Average household size								
less than or equal to 2.2	30,711,000	18.1	0.39	0.2125	34,067,000	18.1	0.50	0.9458
greater than 2.2 and less than or equal to 2.8	90,544,000	53.3	0.57	0.0036	98,844,000	52.5	0.64	0.2566
greater than 2.8	48,703,000	28.7	0.43	0.0050	55,262,000	29.4	0.80	0.3811
Percent less than high school								
less than or equal to 10%	56,267,000	33.1	0.49	0.5685	61,434,000	32.7	0.58	0.4312
greater than 10% and less than or equal to 20%	53,285,000	31.4	0.42	0.8869	58,928,000	31.3	0.45	0.9367
greater than 20%	60,405,000	35.5	0.36	0.3490	67,810,000	36.0	0.77	0.5235
Percent limited English proficient								
0%	50,896,000	30.0	0.46	0.9527	54,100,000	28.8	0.68	0.0879
greater than 0% and less than 2%	60,077,000	35.4	0.55	0.1617	66,808,000	35.5	0.49	0.7559
greater than or equal to 2%	58,984,000	34.7	0.40	0.0519	67,265,000	35.8	1.02	0.3140
Percent below 150% of poverty								
less than or equal to 10%	53,623,000	31.6	0.45	0.9123	60,171,000	32.0	0.87	0.6287
greater than 10% and less than or equal to 30%	76,899,000	45.3	0.41	0.6821	85,354,000	45.4	0.48	0.8135
greater than 30%	39,436,000	23.2	0.16	0.4518	42,648,000	22.7	0.98	0.5845
Age								
16–25	38,603,000	22.7	0.13	0.0317	39,430,000	21.0	1.25	0.1694
26–35	35,559,000	20.9	0.12	0.2186	39,412,000	21.0	0.78	0.9775
36–45	38,696,000	22.8	0.20	0.5379	44,182,000	23.5	0.82	0.3897
46–65	57,099,000	33.6	0.23	0.1705	65,149,000	34.6	1.21	0.4035
Sex								
Male	83,001,000	48.8	0.22	0.0028	92,467,000	49.1	1.01	0.7658
Female	86,956,000	51.2	0.22	0.0028	95,705,000	50.9	1.01	0.7658
Race/ethnicity ⁴								
Hispanic	22,533,000	13.3	0.21	0.7168	25,293,000	13.4	1.68	0.9140
Black, non-Hispanic	17,553,000	10.3	0.21	0.6274	22,069,000	11.7	1.67	0.4087
Other, non-Hispanic	129,871,000	76.4	0.01	0.0394	140,810,000	74.8	1.94	0.4199

See notes at end of table.

Table A-6. Weighted response rates for the ALL Background Questionnaire, by key characteristics:
2003—Continued

Analysis variable	Trimming and Final Raking					
	Final estimate	Percentage	SE^2	p value ²	SE^1	p value ¹
Total	188,173,000	100.0	†	†	†	†
Region						
Northeast	35,730,000	19.0	#	0.9256	2.36	0.7883
Midwest	42,734,000	22.7	#	0.9719	2.24	0.7755
South	66,654,000	35.4	#	0.9961	1.50	0.7159
West	43,055,000	22.9	#	0.9519	2.21	0.4133
Metropolitan Statistical Area status						
Non-Metropolitan Statistical Area	33,472,000	17.8	#	0.9404	2.40	0.1062
Metropolitan Statistical Area	154,701,000	82.2	#	0.9404	2.40	0.1062
Locale ³						
Urban	128,627,000	68.4	0.03	0.2676	1.93	0.1272
Suburban	21,245,000	11.3	0.01	0.9477	0.80	0.1814
Rural	38,301,000	20.4	0.03	0.1607	1.22	0.1215
Majority own/rent						
Rent	48,491,000	25.8	0.09	0.9693	0.66	0.3866
Own	139,681,000	74.2	0.09	0.9693	0.66	0.3866
Average household size						
less than or equal to 2.2	34,014,000	18.1	0.08	0.7213	0.59	0.4087
greater than 2.2 and less than or equal to 2.8	98,897,000	52.6	0.08	0.7293	0.70	0.1342
greater than 2.8	55,262,000	29.4	0.06	0.9957	0.85	0.4999
Percent less than high school						
less than or equal to 10%	61,389,000	32.6	0.08	0.7559	0.78	0.3364
greater than 10% and less than or equal to 20%	58,924,000	31.3	0.07	0.9742	0.63	0.8772
greater than 20%	67,860,000	36.1	0.07	0.6939	0.96	0.3733
Percent limited English proficient						
0%	54,225,000	28.8	0.04	0.0729	0.70	0.1248
greater than 0% and less than 2%	66,767,000	35.5	0.07	0.7718	0.66	0.1707
greater than or equal to 2%	67,180,000	35.7	0.09	0.6105	1.13	0.8732
Percent below 150% of poverty						
less than or equal to 10%	60,169,000	32.0	0.08	0.9924	0.98	0.6308
greater than 10% and less than or equal to 30%	85,377,000	45.4	0.08	0.8699	0.58	0.9380
greater than 30%	42,626,000	22.7	0.08	0.8855	0.95	0.6542
Age						
16–25	39,430,000	21.0	#	0.9036	1.25	0.2494
26–35	39,412,000	21.0	#	0.9393	0.77	0.8203
36–45	44,182,000	23.5	#	0.7865	0.78	0.4618
46–65	65,149,000	34.6	#	0.9750	1.16	0.5488
Sex						
Male	92,467,000	49.1	#	0.9294	0.99	0.6794
Female	95,705,000	50.9	#	0.9294	0.99	0.6794

See notes at end of table.

Table A-6. Weighted response rates for the ALL Background Questionnaire, by key characteristics: 2003—Continued

Analysis variable	Trimming and Final Raking					
	Final estimate	Percentage	SE ²	p value ²	SE ¹	p value ¹
Race/ethnicity ⁴						
Hispanic	25,293,000	13.4	#	0.9143	1.61	0.9476
Black, non-Hispanic	22,069,000	11.7	#	0.9987	1.62	0.3606
Other, non-Hispanic	140,810,000	74.8	#	0.9405	1.94	0.4123

rounds to zero.

† Not applicable.

¹ The p value and SE are for tests of difference between estimated percentage of current step and total sample (BQ base weights).

² The p value and SE are for tests of difference between estimated percentage of current and previous weighting steps.

³ This indicator was set to 'urban' if the largest percentage in the segment was inside urbanized areas; 'suburban' if the largest percentage was inside urban clusters; 'Rural' if the largest percentage was the rural population. The terms 'Urbanized areas', 'urban clusters', and 'rural' are Census Bureau defined.

⁴ All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as Other include non-Hispanics of all other races, including White, Asian, American Indian/Alaska Native, or multiracial.

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Adult Literacy and Lifeskills Survey (ALL), 2003.

Table A-7. Weighted response rates for the ALL Background Questionnaire, by BQ items: 2003

Variable	Description	Number of eligibles	Weighted item response rate
D42	Gross monthly salary	44	21.7
K9	Personal income range (\$20,000 or more)	137	61.5
K12	Total income range	498	71.8
K7	Personal income range	385	72.4
K14	Total income range (\$40,000 or more)	173	76.3
F4	Type of certificate, degree or diploma	760	81.4
D43	Annual personal net income	226	81.1
K11	Total household income	2,618	80.3

SOURCE: U.S. Department of Education, National Center for Education Statistics, Adult Literacy and Lifeskills Survey (ALL), 2003.

Table A-8. Weighted response rates for the ALL Background Questionnaire, by key characteristics: 2003

Survey item	D42: Gross monthly salary				Chi-square	
	Respondents		Eligibles			
	Percent	Standard error	Percent	Standard error	Statistic	<i>p</i> value
Born in US						
Yes	60.0	14.61	76.9	9.13	1.01	0.315
No	40.0	14.61	23.1	9.13	†	†
Highest education						
Less than high school	27.5	12.23	23.4	4.51	0.08	0.944
High school	23.1	11.18	23.6	5.71	†	†
More than high school	49.4	16.26	53.0	5.96	†	†
Remedial help						
Yes	11.2	9.83	15.9	6.64	0.26	0.613
No	88.8	9.83	84.1	6.64	†	†
Enjoyed math in school						
Agree or no opinion	69.5	15.29	67.3	7.93	0.02	0.880
Disagree	30.5	15.29	32.7	7.93	†	†
Age						
16–25	#	#	14.1	7.62	1.68 ¹	0.359 ¹
26–35	23.6	13.38	27.6	8.38	†	†
36–45	48.9	18.14	29.2	7.80	†	†
46–65	27.5	14.81	29.1	9.70	†	†
Language spoken at home						
English	60.0	14.61	78.2	10.71	1.24	0.446
Spanish	33.2	14.03	16.4	7.40	†	†
Other	6.8	6.20	5.4	4.24	†	†
Work in past 12 months						
Yes	87.5	9.09	70.4	5.76	2.61	0.106
No	12.5	9.09	29.6	5.76	†	†
Read letters or e-mails for job						
At least once a week	68.6	13.85	66.3	8.28	0.05	0.830
Less than once a week	31.4	13.85	33.7	8.28	†	†
Sex						
Male	52.4	15.24	43.8	8.78	0.40	0.526
Female	47.6	15.24	56.2	8.78	†	†
Use computer in job						
Yes	61.5	19.97	73.0	9.46	0.41	0.524
No	38.5	19.97	27.0	9.46	†	†

See notes at end of table.

Table A-8. Weighted response rates for the ALL Background Questionnaire, by key characteristics:
2003—Continued

Survey item	K9: Personal income range (\$20,000 or more)					
	Respondents		Eligibles		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	<i>p</i> value
Born in US						
Yes	79.8	5.80	84.5	4.33	4.19	0.041
No	20.2	5.80	15.5	4.33	†	†
Highest education						
Less than high school	3.8	2.81	2.8	1.71	4.85	0.088
High school	28.5	5.72	22.3	4.38	†	†
More than high school	67.7	6.19	74.9	4.77	†	†
Remedial help						
Yes	11.6	4.08	8.0	2.85	6.92	0.009
No	88.4	4.08	92.0	2.85	†	†
Enjoyed math in school						
Agree or No opinion	67.0	5.99	67.2	4.22	0.00	0.956
Disagree	33.0	5.99	32.8	4.22	†	†
Age						
16–25	7.7	2.95	6.0	2.44	2.38	0.404
26–35	28.5	5.24	25.0	4.02	†	†
36–45	28.4	5.57	30.7	5.10	†	†
46–65	35.5	5.63	38.4	4.34	†	†
Language spoken at home						
English	83.5	4.33	89.0	2.86	9.59 ²	0.002 ²
Spanish	3.5	2.11	2.1	1.27	†	†
Other	13.0	3.64	8.9	2.64	†	†
Work in past 12 months						
Yes	91.2	2.67	92.4	2.14	0.73	0.394
No	8.8	2.67	7.6	2.14	†	†
Read letters or e-mails for job						
At least once a week	71.5	6.63	78.3	5.25	8.08	0.004
Less than once a week	28.5	6.63	21.7	5.25	†	†
Sex						
Male	42.5	6.04	44.8	3.72	0.41	0.520
Female	57.5	6.04	55.2	3.72	†	†
Use computer in job						
Yes	76.1	6.74	80.1	4.19	1.00	0.318
No	23.9	6.74	19.9	4.19	†	†

See notes at end of table.

Table A-8. Weighted response rates for the ALL Background Questionnaire, by key characteristics:
2003—Continued

Survey item	K12: Total income range					
	Respondents		Eligibles		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	<i>p</i> value
Born in US						
Yes	86.3	1.74	87.1	1.73	0.49	0.486
No	13.7	1.74	12.9	1.73	†	†
Highest education						
Less than high school	33.3	3.41	30.8	2.62	4.21	0.103
High school	26.3	3.72	25.8	2.97	†	†
More than high school	40.4	3.48	43.4	3.12	†	†
Remedial help						
Yes	17.3	2.23	15.7	1.65	1.94	0.164
No	82.7	2.23	84.3	1.65	†	†
Enjoyed math in school						
Agree or no opinion	68.9	2.35	70.9	1.97	1.55	0.213
Disagree	31.1	2.35	29.1	1.97	†	†
Age						
16–25	47.8	3.63	40.3	3.36	24.98	0.000
26–35	15.2	2.02	16.7	1.85	†	†
36–45	14.3	2.45	15.6	2.17	†	†
46–65	22.7	2.13	27.4	2.53	†	†
Language spoken at home						
English	88.0	1.83	88.7	1.87	1.07	0.583
Spanish	6.6	1.30	6.0	1.15	†	†
Other	5.4	1.27	5.3	1.36	†	†
Work in past 12 months						
Yes	74.9	2.47	74.9	2.55	0.00	0.977
No	25.1	2.47	25.1	2.55	†	†
Read letters or e-mails for job						
At least once a week	56.1	3.88	60.1	2.97	7.33	0.007
Less than once a week	43.9	3.88	39.9	2.97	†	†
Sex						
Male	43.7	3.26	44.3	2.66	0.23	0.633
Female	56.3	3.26	55.7	2.66	†	†
Use computer in job						
Yes	59.0	3.14	61.1	2.62	1.71	0.191
No	41.0	3.14	38.9	2.62	†	†

See notes at end of table.

Table A-8. Weighted response rates for the ALL Background Questionnaire, by key characteristics:
2003—Continued

Survey item	K7: Personal income range					
	Respondents		Eligibles		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	<i>p</i> value
Born in US						
Yes	83.5	3.09	84.2	2.37	0.24	0.621
No	16.5	3.09	15.8	2.37	†	†
Highest education						
Less than high school	25.7	3.16	23.9	2.86	1.81	0.402
High school	25.1	3.29	25.4	2.84	†	†
More than high school	49.2	3.53	50.7	3.31	†	†
Remedial help						
Yes	15.3	2.55	13.2	1.98	3.42	0.064
No	84.7	2.55	86.8	1.98	†	†
Enjoyed math in school						
Agree or no opinion	68.1	2.98	69.0	2.42	0.23	0.632
Disagree	31.9	2.98	31.0	2.42	†	†
Age						
16–25	24.1	3.78	20.4	3.02	11.72	0.005
26–35	20.7	2.86	20.6	2.51	†	†
36–45	22.8	3.24	21.0	2.98	†	†
46–65	32.5	2.82	37.9	2.69	†	†
Language spoken at home						
English	86.8	2.60	87.6	2.41	1.13	0.374
Spanish	7.4	2.16	7.4	1.66	†	†
Other	5.8	1.59	5.0	1.48	†	†
Work in past 12 months						
Yes	72.7	2.92	72.8	2.78	0.02	0.874
No	27.3	2.92	27.2	2.78	†	†
Read letters or e-mails for job						
At least once a week	65.2	3.09	68.3	2.25	3.77	0.052
Less than once a week	34.8	3.09	31.7	2.25	†	†
Sex						
Male	39.8	2.93	41.5	2.43	2.07	0.150
Female	60.2	2.93	58.5	2.43	†	†
Use computer in job						
Yes	65.8	3.77	69.0	3.19	4.16	0.041
No	34.2	3.77	31.0	3.19	†	†

See notes at end of table.

Table A-8. Weighted response rates for the ALL Background Questionnaire, by key characteristics:
2003—Continued

Survey item	K14: Total income range (\$40,000 or more)					
	Respondents		Eligibles		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	<i>p</i> value
Born in US						
Yes	87.2	2.62	88.8	2.14	1.84	0.175
No	12.8	2.62	11.2	2.14	†	†
Highest education						
Less than high school	25.4	4.74	22.3	3.96	6.20	0.032
High school	27.8	4.60	25.2	4.35	†	†
More than high school	46.8	3.85	52.5	4.15	†	†
Remedial help						
Yes	19.9	3.88	18.2	3.09	0.96	0.327
No	80.1	3.88	81.8	3.09	†	†
Enjoyed math in school						
Agree or no opinion	71.2	4.79	70.5	3.41	0.11	0.736
Disagree	28.8	4.79	29.5	3.41	†	†
Age						
16–25	52.7	3.73	46.8	3.93	5.54	0.132
26–35	10.6	2.69	11.5	2.29	†	†
36–45	16.5	3.28	18.0	3.05	†	†
46–65	20.2	3.39	23.7	3.32	†	†
Language spoken at home						
English	91.4	2.11	92.6	1.83	2.33	0.306
Spanish	2.0	1.05	1.8	0.88	†	†
Other	6.6	1.62	5.7	1.41	†	†
Work in past 12 months						
Yes	83.0	3.41	84.9	2.67	2.51	0.113
No	17.0	3.41	15.1	2.67	†	†
Read letters or e-mails for job						
At least once a week	57.4	6.47	63.5	5.64	6.86	0.009
Less than once a week	42.6	6.47	36.5	5.64	†	†
Sex						
Male	51.2	4.50	50.9	4.26	0.07	0.789
Female	48.8	4.50	49.1	4.26	†	†
Use computer in job						
Yes	63.9	4.99	67.6	4.01	2.24	0.134
No	36.1	4.99	32.4	4.01	†	†

See notes at end of table.

Table A-8. Weighted response rates for the ALL Background Questionnaire, by key characteristics:
2003—Continued

Survey item	F4: Type of certificate, degree or diploma					
	Respondents		Eligibles		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	<i>p</i> value
Born in US						
Yes	90.1	1.84	90.8	1.63	1.96	0.162
No	9.9	1.84	9.2	1.63	†	†
Highest education						
Less than high school	26.6	3.51	22.2	2.78	24.82	0.000
High school	15.4	4.01	15.3	3.39	†	†
More than high school	58.0	3.27	62.5	2.80	†	†
Remedial help						
Yes	16.8	1.63	15.6	1.31	4.01	0.045
No	83.2	1.63	84.4	1.31	†	†
Enjoyed math in school						
Agree or no opinion	67.5	2.41	67.0	2.18	0.57	0.450
Disagree	32.5	2.41	33.0	2.18	†	†
Age						
16–25	66.0	2.40	55.4	2.51	68.15	0.000
26–35	17.7	1.69	18.7	1.51	†	†
36–45	9.2	1.29	11.5	1.23	†	†
46–65	7.0	0.92	14.4	1.46	†	†
Language spoken at home						
English	93.6	1.36	93.8	1.31	0.69	0.660
Spanish	2.9	0.99	3.0	0.91	†	†
Other	3.5	1.25	3.2	1.01	†	†
Work in past 12 months						
Yes	83.7	1.57	85.9	1.21	17.66	0.000
No	16.3	1.57	14.1	1.21	†	†
Read letters or e-mails for job						
At least once a week	67.8	2.25	72.9	1.81	57.13	0.000
Less than once a week	32.2	2.25	27.1	1.81	†	†
Sex						
Male	47.1	2.45	47.3	2.02	0.06	0.803
Female	52.9	2.45	52.7	2.02	†	†
Use computer in job						
Yes	65.3	1.81	69.7	1.64	37.69	0.000
No	34.7	1.81	30.3	1.64	†	†

See notes at end of table.

Table A-8. Weighted response rates for the ALL Background Questionnaire, by key characteristics:
2003—Continued

Survey item	D43: Annual personal net income					
	Respondents		Eligibles		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	<i>p</i> value
Born in US						
Yes	86.3	2.45	86.6	2.14	0.08	0.776
No	13.7	2.45	13.4	2.14	†	†
Highest education						
Less than high school	6.0	2.39	6.3	2.43	0.18	0.865
High school	25.5	3.50	25.3	3.41	†	†
More than high school	68.5	3.86	68.3	3.59	†	†
Remedial help						
Yes	10.3	2.31	10.6	2.25	0.09	0.768
No	89.7	2.31	89.4	2.25	†	†
Enjoyed math in school						
Agree or no opinion	72.0	3.59	71.3	3.12	0.14	0.705
Disagree	28.0	3.59	28.7	3.12	†	†
Age						
16–25	5.2	2.03	4.5	1.65	2.13	0.493
26–35	14.7	3.00	14.2	2.82	†	†
36–45	20.8	3.66	21.5	3.74	†	†
46–65	59.4	4.01	59.9	4.29	†	†
Language spoken at home						
English	89.3	2.50	90.6	2.04	2.31	0.281
Spanish	5.1	2.17	4.7	1.81	†	†
Other	5.5	2.14	4.7	1.76	†	†
Work in past 12 months						
Yes	85.3	2.89	84.5	2.68	0.32	0.574
No	14.7	2.89	15.5	2.68	†	†
Read letters or e-mails for job						
At least once a week	75.3	2.67	74.7	2.61	0.29	0.589
Less than once a week	24.7	2.67	25.3	2.61	†	†
Sex						
Male	62.2	3.98	59.8	3.64	3.33	0.068
Female	37.8	3.98	40.2	3.64	†	†
Use computer in job						
Yes	71.7	2.77	73.0	2.77	1.20	0.273
No	28.3	2.77	27.0	2.77	†	†

See notes at end of table.

Table A-8. Weighted response rates for the ALL Background Questionnaire, by key characteristics: 2003—Continued

Survey item	K11: Total household income					
	Respondents		Eligibles		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	<i>p</i> value
Born in US						
Yes	86.3	1.18	86.4	1.05	0.16	0.685
No	13.7	1.18	13.6	1.05	†	†
Highest education						
Less than high school	17.2	1.13	19.9	1.16	27.52	0.000
High school	24.7	1.15	24.9	1.08	†	†
More than high school	58.0	1.57	55.2	1.42	†	†
Remedial help						
Yes	13.3	0.92	13.8	0.92	2.26	0.132
No	86.7	0.92	86.2	0.92	†	†
Enjoyed math in school						
Agree or no opinion	67.5	1.14	68.1	1.01	2.44	0.118
Disagree	32.5	1.14	31.9	1.01	†	†
Age						
16–25	21.6	1.39	25.3	1.52	39.17	0.000
26–35	22.3	0.98	21.2	0.89	†	†
36–45	24.3	1.01	22.6	0.88	†	†
46–65	31.8	1.46	30.9	1.43	†	†
Language spoken at home						
English	89.8	1.29	89.6	1.27	3.81	0.127
Spanish	7.2	1.35	7.0	1.19	†	†
Other	3.0	0.57	3.5	0.61	†	†
Work in past 12 months						
Yes	82.6	0.88	81.1	1.00	9.97	0.002
No	17.4	0.88	18.9	1.00	†	†
Read letters or e-mails for job						
At least once a week	74.3	1.12	71.7	1.10	21.44	0.000
Less than once a week	25.7	1.12	28.3	1.10	†	†
Sex						
Male	49.1	1.16	48.2	1.24	3.66	0.056
Female	50.9	1.16	51.8	1.24	†	†
Use computer in job						
Yes	73.0	1.03	70.9	0.94	15.10	0.000
No	27.0	1.03	29.1	0.94	†	†

† Not applicable.

¹ Chi-squared test is performed after collapsing age categories 16–25 and 26–35.

² Chi-squared test is performed after collapsing language categories Spanish and Other.

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Adult Literacy and Lifeskills Survey (ALL), 2003.

Table A-9. Weighted response rates for the ALL Background Questionnaire, by key characteristics: 2003

Survey item	D42: Gross monthly salary					
	Respondents percent	Eligibles percent	Bias			Percent relative bias
			Estimate	Standard error	<i>p</i> value	
Born in US						
Yes	60	76.9	-16.9	13.66	0.227	-22.0
No	40	23.1	16.9	13.66	0.227	73.2
Highest education						
Less than high school	27.5	23.4	4	11.96	0.738	17.1
High school	23.1	23.6	-0.4	12.1	0.972	-1.7
More than high school	49.4	53.0	-3.6	17	0.833	-6.8
Remedial help						
Yes	11.2	15.9	-4.8	8.56	0.582	-30.2
No	88.8	84.1	4.8	8.56	0.582	5.7
Enjoyed math in school						
Agree or no opinion	69.5	67.3	2.2	13.5	0.874	3.3
Disagree	30.5	32.7	-2.2	13.5	0.874	-6.7
Age						
16–25	#	14.1	14.1	†	†	†
26–35	23.6	27.6	-4	12.09	0.745	-14.5
36–45	48.9	29.2	19.7	14.14	0.173	67.5
46–65	27.5	29.1	-1.7	10.98	0.881	-5.8
Language spoken at home						
English	60	78.2	-18.2	15.15	0.240	-23.3
Spanish	33.2	16.4	16.8	12.28	0.183	102.4
Other	6.8	5.4	1.4	6.55	0.830	25.9
Work in past 12 months						
Yes	87.5	70.4	17.1	8.5	0.053	24.3
No	12.5	29.6	-17.1	8.5	0.053	-57.8
Read letters or e-mails for job						
At least once a week	68.6	66.3	2.4	10.45	0.823	3.6
Less than once a week	31.4	33.7	-2.4	10.45	0.823	-7.1
Sex						
Male	52.4	43.8	8.6	12.46	0.496	19.6
Female	47.6	56.2	-8.6	12.46	0.496	-15.3
Use computer in job						
Yes	61.5	73.0	-11.3	16.63	0.501	-15.5
No	38.5	27.0	11.3	16.63	0.501	41.9

See notes at end of table.

Table A-9. Weighted response rates for the ALL Background Questionnaire, by key characteristics: 2003—Continued

Survey item	K9: Personal income range (\$20,000 or more)					
	Respondents percent	Eligibles percent	Bias			Percent relative bias
			Estimate	Standard error	<i>p</i> value	
Born in US						
Yes	79.8	84.5	-4.7	2.26	0.047	-5.6
No	20.2	15.5	4.7	2.26	0.047	30.3
Highest education						
Less than high school	3.8	2.8	1	1.19	0.391	35.7
High school	28.5	22.3	6.1	3.22	0.066	27.4
More than high school	67.7	74.9	-7.2	3.26	0.036	-9.6
Remedial help						
Yes	11.6	8.0	3.6	1.32	0.010	45.0
No	88.4	92.0	-3.6	1.32	0.010	-3.9
Enjoyed math in school						
Agree or no opinion	67	67.2	-0.2	3.98	0.957	-0.3
Disagree	33	32.8	0.2	3.98	0.957	0.6
Age						
16–25	7.7	6.0	1.6	1.19	0.179	26.7
26–35	28.5	25.0	3.5	2.67	0.201	14.0
36–45	28.4	30.7	-2.3	3.84	0.558	-7.5
46–65	35.5	38.4	-2.9	3.83	0.461	-7.6
Language spoken at home						
English	83.5	89.0	-5.5	1.71	0.003	-6.2
Spanish	3.5	2.1	1.4	0.86	0.131	66.7
Other	13	8.9	4.1	1.27	0.003	46.1
Work in past 12 months						
Yes	91.2	92.4	-1.2	1.39	0.400	-1.3
No	8.8	7.6	1.2	1.39	0.400	15.8
Read letters or e-mails for job						
At least once a week	71.5	78.3	-6.8	2.34	0.007	-8.7
Less than once a week	28.5	21.7	6.8	2.34	0.007	31.3
Sex						
Male	42.5	44.8	-2.3	3.63	0.526	-5.1
Female	57.5	55.2	2.3	3.63	0.526	4.2
Use computer in job						
Yes	76.1	80.1	-3.7	3.68	0.325	-4.6
No	23.9	19.9	3.7	3.68	0.325	18.6

See notes at end of table.

Table A-9. Weighted response rates for the ALL Background Questionnaire, by key characteristics: 2003—Continued

Survey item	K12: Total income range					
	Respondents percent	Eligibles percent	Bias			Percent relative bias
			Estimate	Standard error	<i>p</i> value	
Born in US						
Yes	86.3	87.1	-0.7	1.05	0.497	-0.8
No	13.7	12.9	0.7	1.05	0.497	5.4
Highest education						
Less than high school	33.3	30.8	2.5	1.27	0.055	8.1
High school	26.3	25.8	0.5	1.55	0.756	1.9
More than high school	40.4	43.4	-3	1.3	0.027	-6.9
Remedial help						
Yes	17.3	15.7	1.6	1.08	0.161	10.2
No	82.7	84.3	-1.6	1.08	0.161	-1.9
Enjoyed math in school						
Agree or no opinion	68.9	70.9	-2	1.66	0.231	-2.8
Disagree	31.1	29.1	2	1.66	0.231	6.9
Age						
16–25	47.8	40.3	7.5	1.12	0.000	18.6
26–35	15.2	16.7	-1.5	1.35	0.266	-9.0
36–45	14.3	15.6	-1.3	0.95	0.198	-8.3
46–65	22.7	27.4	-4.7	1.21	0.000	-17.2
Language spoken at home						
English	88	88.7	-0.8	0.77	0.336	-0.9
Spanish	6.6	6.0	0.6	0.6	0.333	10.0
Other	5.4	5.3	0.2	0.58	0.781	3.8
Work in past 12 months						
Yes	74.9	74.9	-0.0	1.17	0.977	0.0
No	25.1	25.1	0.0	1.17	0.977	0.0
Read letters or e-mails for job						
At least once a week	56.1	60.1	-4.1	1.51	0.011	-6.8
Less than once a week	43.9	39.9	4.1	1.51	0.011	10.3
Sex						
Male	43.7	44.3	-0.6	1.32	0.637	-1.4
Female	56.3	55.7	0.6	1.32	0.637	1.1
Use computer in job						
Yes	59	61.1	-2.1	1.65	0.204	-3.4
No	41	38.9	2.1	1.65	0.204	5.4

See notes at end of table.

Table A-9. Weighted response rates for the ALL Background Questionnaire, by key characteristics:
2003—Continued

Survey item	K7: Personal income range					
	Respondents percent	Eligibles percent	Bias			Percent relative bias
			Estimate	Standard error	<i>p</i> value	
Born in US						
Yes	83.5	84.2	-0.7	1.5	0.631	-0.8
No	16.5	15.8	0.7	1.5	0.631	4.4
Highest education						
Less than high school	25.7	23.9	1.8	1.34	0.184	7.5
High school	25.1	25.4	-0.3	1.4	0.846	-1.2
More than high school	49.2	50.7	-1.6	1.67	0.362	-3.2
Remedial help						
Yes	15.3	13.2	2.2	1.15	0.065	16.7
No	84.7	86.8	-2.2	1.15	0.065	-2.5
Enjoyed math in school						
Agree or no opinion	68.1	69.0	-1	2.05	0.639	-1.4
Disagree	31.9	31.0	1	2.05	0.639	3.2
Age						
16–25	24.1	20.4	3.7	1.28	0.007	18.1
26–35	20.7	20.6	#	1.37	0.991	#
36–45	22.8	21.0	1.8	1.17	0.145	8.6
46–65	32.5	37.9	-5.5	1.88	0.007	-14.5
Language spoken at home						
English	86.8	87.6	-0.8	1.02	0.448	-0.9
Spanish	7.4	7.4	-0.1	1.03	0.941	-1.4
Other	5.8	5.0	0.9	0.32	0.012	18.0
Work in past 12 months						
Yes	72.7	72.8	-0.2	1.14	0.876	-0.3
No	27.3	27.2	0.2	1.14	0.876	0.7
Read letters or e-mails for job						
At least once a week	65.2	68.3	-3.2	1.62	0.061	-4.7
Less than once a week	34.8	31.7	3.2	1.62	0.061	10.1
Sex						
Male	39.8	41.5	-1.7	1.19	0.160	-4.1
Female	60.2	58.5	1.7	1.19	0.160	2.9
Use computer in job						
Yes	65.8	69.0	-3.4	1.67	0.054	-4.9
No	34.2	31.0	3.4	1.67	0.054	11.0

See notes at end of table.

Table A-9. Weighted response rates for the ALL Background Questionnaire, by key characteristics:
2003—Continued

Survey item	K14: Total income range (\$40,000 or more)					
	Respondents percent	Eligibles percent	Bias			Percent relative bias
			Estimate	Standard error	<i>p</i> value	
Born in US						
Yes	87.2	88.8	-1.6	1.15	0.171	-1.8
No	12.8	11.2	1.6	1.15	0.171	14.3
Highest education						
Less than high school	25.4	22.3	3.1	2.09	0.155	13.9
High school	27.8	25.2	2.7	1.25	0.042	10.7
More than high school	46.8	52.5	-5.7	2.12	0.012	-10.9
Remedial help						
Yes	19.9	18.2	1.8	1.88	0.355	9.9
No	80.1	81.8	-1.8	1.88	0.355	-2.2
Enjoyed math in school						
Agree or no opinion	71.2	70.5	0.7	2.11	0.743	1.0
Disagree	28.8	29.5	-0.7	2.11	0.743	-2.4
Age						
16–25	52.7	46.8	5.9	2.51	0.025	12.6
26–35	10.6	11.5	-0.9	1.67	0.591	-7.8
36–45	16.5	18.0	-1.5	1.91	0.444	-8.3
46–65	20.2	23.7	-3.5	2.19	0.116	-14.8
Language spoken at home						
English	91.4	92.6	-1.2	0.80	0.148	-1.3
Spanish	2.0	1.8	0.2	0.38	0.576	11.1
Other	6.6	5.7	1	0.68	0.165	17.5
Work in past 12 months						
Yes	83.0	84.9	-1.9	1.32	0.160	-2.2
No	17.0	15.1	1.9	1.32	0.160	12.6
Read letters or e-mails for job						
At least once a week	57.4	63.5	-5.9	2.17	0.011	-9.3
Less than once a week	42.6	36.5	5.9	2.17	0.011	16.2
Sex						
Male	51.2	50.9	0.4	1.46	0.794	0.8
Female	48.8	49.1	-0.4	1.46	0.794	-0.8
Use computer in job						
Yes	63.9	67.6	-3.5	2.23	0.127	-5.2
No	36.1	32.4	3.5	2.23	0.127	10.8

See notes at end of table.

Table A-9. Weighted response rates for the ALL Background Questionnaire, by key characteristics:
2003—Continued

Survey item	F4: Type of certificate, degree or diploma					
	Respondents percent	Eligibles percent	Bias			Percent relative bias
			Estimate	Standard error	<i>p</i> value	
Born in US						
Yes	90.1	93.8	-0.7	0.47	0.160	-0.7
No	9.9	6.2	0.7	0.47	0.160	11.3
Highest education						
Less than high school	26.6	3.4	4.3	0.66	0.000	126.5
High school	15.4	14.5	0.2	0.84	0.833	1.4
More than high school	58.0	82.2	-4.5	0.85	0.000	-5.5
Remedial help						
Yes	16.8	10.5	1.2	0.61	0.063	11.4
No	83.2	89.5	-1.2	0.61	0.063	-1.3
Enjoyed math in school						
Agree or no opinion	67.5	64.7	0.5	0.68	0.447	0.8
Disagree	32.5	35.3	-0.5	0.68	0.447	-1.4
Age						
16–25	66.0	9.0	10.6	0.67	0.000	117.8
26–35	17.7	22.8	-0.9	0.83	0.265	-3.9
36–45	9.2	21.6	-2.3	0.59	0.000	-10.6
46–65	7.0	46.6	-7.4	0.84	0.000	-15.9
Language spoken at home						
English	93.6	95.0	-0.3	0.47	0.583	-0.3
Spanish	2.9	3.2	-0.1	0.4	0.900	-3.1
Other	3.5	1.8	0.3	0.29	0.290	16.7
Work in past 12 months						
Yes	83.7	95.8	-2.3	0.48	0.000	-2.4
No	16.3	4.2	2.3	0.48	0.000	54.8
Read letters or e-mails for job						
At least once a week	67.8	92.5	-4.6	0.49	0.000	-5.0
Less than once a week	32.2	7.5	4.6	0.49	0.000	61.3
Sex						
Male	47.1	48.0	-0.2	0.65	0.806	-0.4
Female	52.9	52.0	0.2	0.65	0.806	0.4
Use computer in job						
Yes	65.3	87.5	-4.1	0.61	0.000	-4.7
No	34.7	12.5	4.1	0.61	0.000	32.8

See notes at end of table.

Table A-9. Weighted response rates for the ALL Background Questionnaire, by key characteristics:
2003—Continued

Survey item	D43: Annual personal net income					
	Respondents percent	Eligibles percent	Bias			Percent relative bias
			Estimate	Standard error	<i>p</i> value	
Born in US						
Yes	86.3	87.7	-0.3	0.94	0.779	-0.3
No	13.7	12.3	0.3	0.94	0.779	2.4
Highest education						
Less than high school	6.0	7.9	-0.4	0.64	0.572	-5.1
High school	25.5	24.4	0.2	1.62	0.893	0.8
More than high school	68.5	67.7	0.2	1.77	0.935	0.3
Remedial help						
Yes	10.3	11.6	-0.2	0.81	0.771	-1.7
No	89.7	88.4	0.2	0.81	0.771	0.2
Enjoyed math in school						
Agree or no opinion	72.0	68.1	0.8	1.94	0.702	1.2
Disagree	28.0	31.9	-0.8	1.94	0.702	-2.5
Age						
16–25	5.2	1.3	0.7	0.47	0.126	53.8
26–35	14.7	12.0	0.5	0.91	0.582	4.2
36–45	20.8	24.7	-0.7	1.23	0.551	-2.8
46–65	59.4	62.0	-0.5	1.51	0.740	-0.8
Language spoken at home						
English	89.3	95.9	-1.2	0.74	0.105	-1.3
Spanish	5.1	2.8	0.4	0.66	0.507	14.3
Other	5.5	1.3	0.8	0.46	0.093	61.5
Work in past 12 months						
Yes	85.3	81.4	0.7	1.30	0.582	0.9
No	14.7	18.6	-0.7	1.30	0.582	-3.8
Read letters or e-mails for job						
At least once a week	75.3	71.7	0.7	1.32	0.606	1.0
Less than once a week	24.7	28.3	-0.7	1.32	0.606	-2.5
Sex						
Male	62.2	49.5	2.4	1.31	0.076	4.8
Female	37.8	50.5	-2.4	1.31	0.076	-4.8
Use computer in job						
Yes	71.7	79.2	-1.4	1.25	0.262	-1.8
No	28.3	20.8	1.4	1.25	0.262	6.7

See notes at end of table.

Table A-9. Weighted response rates for the ALL Background Questionnaire, by key characteristics:
2003—Continued

Survey item	K11: Total household income					Percent relative bias
	Respondents percent	Eligibles percent	Bias			
			Estimate	Standard error	<i>p</i> value	
Born in US						
Yes	86.3	87.1	-0.2	0.39	0.691	-0.2
No	13.7	12.9	0.2	0.39	0.691	1.6
Highest education						
Less than high school	17.2	30.8	-2.7	0.47	0.000	-8.8
High school	24.7	25.8	-0.2	0.63	0.748	-0.8
More than high school	58.0	43.4	2.9	0.63	0.000	6.7
Remedial help						
Yes	13.3	15.7	-0.5	0.31	0.136	-3.2
No	86.7	84.3	0.5	0.31	0.136	0.6
Enjoyed math in school						
Agree or no opinion	67.5	70.9	-0.7	0.44	0.134	-1.0
Disagree	32.5	29.1	0.7	0.44	0.134	2.4
Age						
16–25	21.6	40.3	-3.7	0.6	0.000	-9.2
26–35	22.3	16.7	1.1	0.41	0.013	6.6
36–45	24.3	15.6	1.7	0.46	0.001	10.9
46–65	31.8	27.4	0.9	0.48	0.086	3.3
Language spoken at home						
English	89.8	88.7	0.2	0.31	0.530	0.2
Spanish	7.2	6.0	0.2	0.28	0.401	3.3
Other	3.0	5.3	-0.4	0.26	0.098	-7.5
Work in past 12 months						
Yes	82.6	74.9	1.5	0.46	0.003	2.0
No	17.4	25.1	-1.5	0.46	0.003	-6.0
Read letters or e-mails for job						
At least once a week	74.3	60.1	2.8	0.59	0.000	4.7
Less than once a week	25.7	39.9	-2.8	0.59	0.000	-7.0
Sex						
Male	49.1	44.3	1.0	0.49	0.059	2.3
Female	50.9	55.7	-1.0	0.49	0.059	-1.8
Use computer in job						
Yes	73.0	61.1	2.3	0.57	0.000	3.8
No	27.0	38.9	-2.3	0.57	0.000	-5.9

† Not applicable.

Rounds to zero

NOTE: There were no item respondents to D42 for age 16-25. Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Adult Literacy and Lifeskills Survey (ALL), 2003.

Exhibit A-1. Chi-square Automated Interaction Detector (CHAID) multivariate analysis of response indicators for the ALL Sample: 2003

Region		MSA status											
Northeast	67.6%	Non-MSA	88.9%	Cell 1		Overall weighted response rate = 79.7% Total number of eligibles = 4,255							
		MSA	65.2%	Household (HH) size									
				<=2.8	68.1%					Cell 2			
>2.8	58.7%	Cell 3											
Midwest and South	81.1%	Race											
		Hispanic or Non-Hispanic Black	87.9%	Sex									
				Male	84.4%	HH size							
						<=2.2	79.8%	Cell 4					
						(2.2, 2.8]	91.1%	Cell 5					
				>2.8	80.0%	Cell 6							
				Female	91.0%	Age							
		16-45	94.4%			Cell 7							
		46-65	82.0%	Cell 8									
		Other	78.7%	Age									
				16-35	83.1%	Sex							
						Male	79.4%	Own/rent					
								Rent	70.0%	Cell 9			
Own	82.1%					Cell 10							
Female	86.8%			Region									
				Midwest	94.0%	Age							
						16-25	97.4%	Cell 11					
				26-35	90.2%	Cell 12							
				South	81.2%	% Limited English proficient							
		<2%	86.0%			Cell 13							
>=2%	70.3%	Cell 14											
36-65	75.9%	% Poverty											
		<=30%	74.2%	Region									
				Midwest	77.3%	Cell 15							
		South	71.5%	Cell 16									
>30%	84.7%	Cell 17											
West	85.8%	Own/rent											
		Rent	88.6%	Cell 18									
		Own	84.1%	Age									
				16-25	88.7%	Cell 19							
				26-45	79.6%	Cell 20							
		46-65	87.3%	Locale									
Urban/Suburban	90.5%			Cell 21									
Rural	78.5%	Cell 22											

SOURCE: U.S. Department of Education National Center for Education Statistics, Adult Literacy and Lifeskills Survey (ALL), 2003.

Exhibit A-2. Chi-square Automated Interaction Detector (CHAID) multivariate analysis of item nonresponse for item K9 for the ALL Background Questionnaire: 2003

% Poverty			
<=10%	44.4%	Cell 1	Overall weighted response rate = 61.5% Total number of eligibles = 137
>10%	76.0%	Cell 2	

NOTE: Item K9 is the personal income range (\$20,000 or more).

SOURCE: U.S. Department of Education, National Center for Education Statistics, Adult Literacy and Lifeskills Survey (ALL), 2003.

Exhibit A-3. Chi-square Automated Interaction Detector (CHAID) multivariate analysis of item nonresponse for item K12 for the ALL Background Questionnaire: 2003

Age					
16-25	85.1%	Region		Overall weighted response rate = 71.8% Total number of eligibles = 498	
		Northeast & West	93.9%	Locale	
				Urban	89.8%
		Suburban & Rural	100.0%	Cell 2	
Midwest & South	79.0%	Cell 3			
26-65	62.7%	Locale			
		Urban	64.0%	Cell 4	
		Suburban	81.1%	Cell 5	
Rural	52.4%	Cell 6			

NOTE: Item K12 is the total income range.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Adult Literacy and Lifeskills Survey (ALL), 2003.

Exhibit A-4. Chi-square Automated Interaction Detector (CHAID) multivariate analysis of item nonresponse for item K7 for the ALL Background Questionnaire: 2003

Age				Overall weighted response rate = 72.4%		
16-45	78.7%	Cell 1		Total number of eligibles = 385		
46-65	62.0%	% Limited English proficient				
		<2%	54.5%	% Less than high school		
				<=20%	62.8%	Cell 2
				>20%	38.0%	Cell 3
		>=2%	79.2%	Cell 4		

NOTE: Item K7 is the personal income range.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Adult Literacy and Lifeskills Survey (ALL), 2003.

Exhibit A-5. Chi-square Automated Interaction Detector (CHAID) multivariate analysis of item nonresponse for item K14 for the ALL Background Questionnaire: 2003

Own/rent				Overall weighted response rate = 76.3%		
Rent	95.6%	Cell 1		Total number of eligibles = 173		
Own	72.3%	Age				
		16-25	82.7%	Cell 2		
		26-65	63.5%	% Poverty		
				<=10%	52.9%	Cell 3
>10%	78.5%	Cell 4				

NOTE: Item K14 is the total income range (\$40,000 or more).

SOURCE: U.S. Department of Education, National Center for Education Statistics, Adult Literacy and Lifeskills Survey (ALL), 2003.

Exhibit A-6. Chi-square Automated Interaction Detector (CHAID) multivariate analysis of item nonresponse for item F4 for the ALL Background Questionnaire: 2003

Age											
16-25	97.0%	Region						Overall weighted response rate = 81.4%			
		Northeast	100.0%	Cell 1				Total number of eligibles = 760			
		Midwest, South, & West	96.5%	Race/ethnicity							
				Hispanic	85.9%	Cell 2					
				non-Hispanic Black	100.0%	Cell 3					
		Other	97.6%	Region							
				Midwest & South	96.3%			% Poverty			
						<=10%	100.0%	Cell 4			
						>10%	94.5%	Cell 5			
		West	100.0%	Cell 6							
26-35	77.3%	Region									
		Northeast, Midwest, & South	70.4%	Locale							
				Urban	65.1%	Cell 7					
				Suburban & Rural	82.8%	Cell 8					
West	92.2%	Cell 9									
36-45	65.1%	Cell 10									
46-65	39.8%	% Less than high school									
		<=10%	32.5%	Cell 11							
		(10%, 20%]	57.6%	Cell 12							
		>20%	30.5%	Cell 13							

NOTE: Item F4 is the type of certificate, degree, or diploma.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Adult Literacy and Lifeskills Survey (ALL), 2003.

Exhibit A-7. Chi-square Automated Interaction Detector (CHAID) multivariate analysis of item nonresponse for item D43 for the ALL Background Questionnaire: 2003

Region		Overall weighted response rate = 81.1%				
Northeast	97.2%	Total number of eligibles = 226				
Midwest, South, & West	77.5%	Own/rent		Cell 2		
		Rent	65.6%			
		Own	80.8%	MSA status		
				Non-MSA		65.5%
MSA	85.8%	Cell 4				

NOTE: Item D43 is the annual personal net income.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Adult Literacy and Lifeskills Survey (ALL), 2003.

Exhibit A-8. Chi-square Automated Interaction Detector (CHAID) multivariate analysis of item nonresponse for item K11 for the ALL Background Questionnaire: 2003

Age		Race/ethnicity		% Limited English proficient		Age	
16-25	68.7%	Hispanic & Other	70.7%	<2%	66.5%	Cell 1	
				>=2%	77.7%	Cell 2	
		non-Hispanic Black	51.7%	Cell 3			
26-65	84.3%	Region		% Poverty			
		Northeast & West	88.3%	<=30%	90.3%	Cell 4	
				>30%	79.5%	26-35	65.7%
						36-65	88.1%
		Midwest & South	81.4%	Cell 7			

Overall weighted response rate = 80.3%

Total number of eligibles = 2,618

NOTE: Item K11 is total household income.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Adult Literacy and Lifeskills Survey (ALL), 2003.

Appendix B – Technical Notes

Total Survey Error

There are two major components of total survey error: sampling error and nonsampling error. Sampling error is the error that occurs because population estimates are based on a sample rather than a census. Due to clustering effects typical of household surveys, the sample size can be misleading to users when judging the magnitude of sampling error. Therefore, precise measurement of sampling error is necessary and is facilitated in ALL through the replication method known as the stratified jackknife technique.

Nonsampling error contains all sources of error besides sampling error. According to Lessler and Kalsbeek (1992), there are three subcomponents of nonsampling error: (1) frame error, (2) nonresponse error, and (3) measurement error. This report is focused on the nonresponse error, which is the error arising from failure to obtain a response, whether it is unit nonresponse or item nonresponse. A key measure of the impact of nonresponse on total survey error is called nonresponse bias. Nonresponse bias can be substantial when two conditions hold: (1) when response rate is relatively low, and (2) when the difference between the characteristics of respondents and nonrespondents is relatively large. An estimate for nonresponse bias can be expressed as follows:

$$Bias(\bar{y}_R) = (1 - W_R)(\bar{Y}_R - \bar{Y}_N),$$

where W_R is the response rate and \bar{Y}_R and \bar{Y}_N are the mean values of the survey items estimated among the respondents and nonrespondents, respectively. Because survey values for nonrespondents are not available, nonresponse bias is not known and can only be estimated by using data available for both respondents and nonrespondents.

Response Rate Computations

The response rates were computed for the three components as follows. The numerator of the response rate consists of all cases successfully completed at the relevant stage. The denominator includes all sampled cases except those with hearing impaired, physical disability, or other disability

status, following Statistics Canada's guidelines for response rate calculation. The basic response rate calculations were as follows:

$$W_R = \frac{C}{C + NR + eU}$$

where,

C = number of completes;

NR = number of nonrespondents known to be eligible;

U = number of nonrespondents whose eligibility status is unknown; and

e = estimated proportion of U that are eligible.

The estimated proportion of U that are eligible was computed as the observed proportion of the sample that is known to be eligible:

$$e = \frac{C + NR}{C + NR + I}$$

Where,

I = number in the sample that are known to be ineligible

The term eU in the denominator was only applied to the screener response rate due to instances of unknown eligibility status. However, once cooperation was attained for the screener, the eligibility of adults was known, therefore the term eU was equal to zero in the calculations of response rates for the BQ and the assessment.

ALL Sample Design and Data Collection

ALL consists of a nationally representative sample of the noninstitutionalized civilian population of U.S. adults who, at the time of the interview, were between the ages of 16 and 65 years, inclusive. The ALL main survey sample design involved four stages of sample selection:

- The selection of 60 primary sampling units (PSUs), which consists of counties or groups of counties, using the 2000 Census data to form, stratify, and select;
- The selection of 505 secondary sampling units (segments), which consists of census blocks or block groups (BGs), using the 2000 Census data to form and select;
- The listing and selection of dwelling units (DUs) within segments; and
- The enumeration and selection of eligible individuals within DUs.

There were 3,420 completed BQs among the 4,255 sampled persons (SPs).

Instruments in the study included a Screener, a Background Questionnaire, a Core Assessment booklet and a Main Assessment booklet. Both the Screener and the Background Questionnaire were administered via computer-assisted personal interviewing (CAPI) using a laptop computer. The Screener and Background Questionnaire were available in English or Spanish. The Core Assessment and Main Assessment booklets were only administered in English.

The Screener enumerated the household residents, obtained the data necessary for sampling, and selected the study participants. The Background Questionnaire collected demographic data, information on educational experiences, health and well-being, and the use of media and technology. The Background Questionnaire took an average of 30 minutes to administer.

The Core Assessment contained six prose/document literacy and numeracy items and took approximately 5 minutes to administer. The assessment items resembled literacy-related activities that people commonly perform. Interviewers scored the respondents' responses on the laptop computer. Respondents who responded correctly to three or more Core Assessment items were administered the Main Assessment booklet items.

The Main Assessment was self-administered and consisted of sets of assessments measuring skills in prose/document literacy and numeracy. There were 18 versions of the Main Assessment booklet

and each contained between 40 and 53 assessment items. Interviewers assigned a booklet to a respondent by picking from the top of a stack of booklets prearranged to ensure random assignment and even distribution of each type of booklet among respondents. The Main Assessment took approximately 50 minutes to complete.

ALL Weighting Procedures

The purpose of calculating sample weights for ALL was to permit inferences from SPs to the populations from which they were drawn and to allow tabulations to reflect estimates of the population parameters. Sample weights were produced to accomplish the following four objectives:

- Permit unbiased estimates, taking account of the fact that all persons in the population will not have the same probability of selection;
- Minimize biases arising from differences between cooperating and noncooperating SPs;
- Bring data up to the dimensions of the population totals; and
- Use auxiliary data on known population characteristics in such a way as to reduce sampling errors.

The weighting process began with the creation of the screener base weights for the sampled DUs, and continued with a weight adjustment to account for screener nonresponse. Next, base weights were computed for SPs as the product of the screener nonresponse adjusted weight and the reciprocal of the within-household selection probability. The SP base weights were adjusted for BQ nonresponse. Subsequently, the BQ weights were adjusted through an initial raking procedure, so that their sums equal known totals from the 2003 March Supplement of the Current Population Survey. Lastly, outlier weights were trimmed (or reduced) and then recalibrated through a final raking procedure.

ALL Variance Estimation

The stratified jackknife method was implemented to estimate the variance (i.e., sampling error) for most statistics. Jackknifing estimates the sampling variability of any statistic Y as the sum of components of variability that may be attributed to individual pairs of first-stage sampling units (i.e., PSUs, or groups of sampled segments within certainty PSUs). The variance attributed to a particular pair

is measured by estimating how much the value of the statistic would change if only one unit in the pair had been sampled. When using replication techniques such as jackknifing to calculate standard errors, it is necessary to establish a number of subsamples (or replicates) from the full sample, calculate the estimate from each subsample, and sum the squared difference of each replicated estimate from the full-sample estimate.

The stratified jackknife method requires three steps:

1. Forming the replicates;
2. Constructing the replicate weights; and
3. Computing estimates of variance for survey statistics.

The estimate of the variance of a statistic Y is as follows:

$$v(Y) = \sum_{r=1}^G (Y_r - Y)^2,$$

where

- Y_r = the weighted estimate obtained using the r -th replicate weight, and
 Y = the weighted estimate obtained using the full-sample weight.

Tests of Significance

Comparisons made in the text of this report have been tested for statistical significance. For example, when comparing results obtained from the full sample with those obtained only from the responding sample units, tests of statistical significance were used to establish whether or not the observed differences are statistically significant. The estimation of the standard errors that are required in order to undertake the tests of significance is complicated by the complex sample and assessment designs which both generate error variance. Together they mandate a set of statistically complex procedures in order to estimate the correct standard errors. As a consequence, the estimated standard errors contain a sampling variance component estimated by the stratified jackknife method. Details on the procedures used can be found in the WesVar 4.2 User's Guide (Westat 2002).

Two kinds of statistical tests are included in the report: t tests and chi-square tests.

***t* Tests**

t tests were used for testing for the hypothesis that no difference exists between the estimated percentages of variables for two groups. In particular, comparisons were made between the full sample and respondents and for the sample at various weighting stages. Suppose that \bar{x}_A and \bar{x}_B are the percentages for two groups that are being compared and $se(\bar{x}_A - \bar{x}_B)$ is the standard error of the difference between the percentages which accounts for the complex survey design. Then the *t* test is defined as

$$t = \frac{|\bar{x}_A - \bar{x}_B|}{se(\bar{x}_A - \bar{x}_B)}$$

This statistic is then compared to the critical values of the appropriate Student *t*-distribution, to determine whether the difference is statistically significant. The appropriate number of degrees of freedom for the distribution is given by the number of primary sampling units in the design minus the number of sampling strata.

Note that this procedure took account of the fact that the two samples in question were not independent samples. The responding sample was a subsample of the full sample, and the sample at one stage of weighting overlapped with the sample at the previous stage. This effect was accounted for in calculating the standard error of the difference. Note also that, in those cases where both samples were weighted just using base weights the test is exactly equivalent to testing that the mean of the respondents was equal to the mean of the nonrespondents.

Chi-Square Tests

Chi-square tests are used for testing whether two distributions of a given categorical variable are different, conducted in a way that reflects the impact of the complex sample design on sampling variance. In this instance one distribution is for the full sample, and one for the responding sample. Suppose that the categorical variable in question has *c* levels, cross-tabulated producing weighted proportions *p*. The Pearson chi-square statistic is calculated as

$$X^2 = n \sum_{i=1}^2 \sum_{j=1}^c (p_{ij} - p_{i.} p_{.j})^2 / p_{i.} p_{.j}$$

where j denotes the categories of the categorical variable, and i indexes the samples (full sample and respondents), and n indicates the overall sample size. This statistic is not suitable for use directly in a statistical test with these data, for two reasons. First, the fact that the respondents are a subset of the full sample violates the standard assumptions for a chi-square test of this kind. Second, this statistic does not account for the complex sample design used to collect the data.

Thus the Pearson Chi-square statistic is modified appropriately to account for the impact of these two features. The resulting test statistic is referred to as the Rao-Scott Adjusted chi-square statistic. It is sometimes also referred to as the Satterthwaite-adjusted chi-square statistic. The number of degrees of freedom for the chi-square test, normally given as $(c - 1)$, where c is the number of categories of the categorical variable for each distribution, is also modified on account of the complex design. The modified test statistic is then compared to the chi-square distribution with the appropriate number of degrees of freedom, to determine whether the difference in the two distributions is statistically significant. A detailed description of the technique is provided in the following paragraphs (also see Rao and Thomas (2003)).

The first step in the calculation of the Satterthwaite-adjusted chi-square statistic is to form the following vector:

$$Y = \sqrt{n} \begin{pmatrix} p_{11} - p_{1.} p_{.1} \\ p_{12} - p_{1.} p_{.2} \\ \vdots \\ p_{rc} - p_{r.} p_{.c} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{rc} \end{pmatrix}$$

An $rc \times 1$ vector made up of the products of the marginal proportions is defined as

$$p = \begin{pmatrix} p_{1.} p_{.1} \\ p_{1.} p_{.2} \\ \vdots \\ p_{r.} p_{.c} \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_{rc} \end{pmatrix}$$

For each replicate, an $rc \times rc$ matrix is calculated whose ij -th element is made up of

$$(y_{ig} - y_i)(y_{jg} - y_j),$$

where y_{ig} and y_{jg} are the i -th and j -th elements of \mathbf{Y} calculated for the g -th replicate and y_i and y_j are the corresponding full-sample values. The ij -th element of the estimated covariance matrix for \mathbf{Y} , $\mathbf{B}=\text{cov}(\mathbf{Y})$, is calculated using the following formula:

$$B_{ij} = \sum_{g=1}^G (y_{ig} - y_i)(y_{jg} - y_j),$$

The Satterthwaite's approximation to degrees of freedom for the chi-square statistic to be calculated is

$$v = \frac{\left(\sum_{i=1}^{rc} \frac{B_{ii}}{p_i} \right)^2}{\sum_{i=1}^{rc} \sum_{j=1}^{rc} \frac{B_{ij}^2}{p_i p_j}}.$$

Since v will generally not be an integer, interpolation in standard chi-square tables is required.

Finally, the adjusted chi-square statistic is defined as

$$RS3 = \frac{X^2}{\left(\sum_{i=1}^{rc} \frac{B_{ii}}{p_i} \right)}.$$

Logistic Regression Models

A linear model for investigating the relationship between binary (dichotomous) outcomes and a set of explanatory variables is referred to as a *logistic regression model*. The data are assumed to follow a binomial distribution, with probabilities that depend on the independent variables. In this instance the binary outcome of interest is whether or not the sampled unit completed the ALL background questionnaire.

Let p_i denote the probability that the i th sampled person will participate. Under the logistic regression model, the log odds of response propensity (expressed in terms of the logarithm of $p_i/(1-p_i)$), is assumed to have the following linear form:

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi}$$

where $X_{1i}, X_{2i}, \dots, X_{pi}$ are p auxiliary variables associated with the i th sampled person, and $\beta_0, \beta_1, \dots, \beta_p$ are coefficients to be estimated. Asymptotic assumptions are used to develop statistical tests to determine which, if any, of the coefficients are significantly different from zero. In the analyses in this report the standard procedures for carrying out logistic regression analyses have been modified both to incorporate the sampling weights in the estimation of the coefficients, and to reflect the effect of the complex sample design on the variance-covariance matrix of the coefficients.

The Newton-Raphson algorithm is used to iteratively solve for parameter solutions in the logistic regression. Let $q(\boldsymbol{\beta}) = \partial L_n(\boldsymbol{\beta}) / \partial \boldsymbol{\beta}$ be the vector of first partial derivatives of the sample log-likelihood with respect to $\boldsymbol{\beta}$. Let $\mathbf{H}(\boldsymbol{\beta})$ be the matrix of second partial derivatives (or Hessian) of the sample log-likelihood having entries $\partial^2 L / \partial \beta_a \partial \beta_b$, where β_a and β_b are two separate components of $\boldsymbol{\beta}$. Denote by \mathbf{q}^t and \mathbf{H}^t the values of $q(\boldsymbol{\beta})$ and $\mathbf{H}(\boldsymbol{\beta})$ evaluated at \mathbf{b}^t , the value of the estimate \mathbf{b} at step t .

The general approach is to approximate the sample log-likelihood at the desired estimate, $L_n(b)$, at step t in the iterative process near the point \mathbf{b}^t by a second-order Taylor series expansion:

$$L_n^t(\mathbf{b}) \cong L_n(\mathbf{b}^t) + \mathbf{q}^{t'}(\mathbf{b} - \mathbf{b}^t) + \frac{1}{2}(\mathbf{b} - \mathbf{b}^t)' \mathbf{H}^t(\mathbf{b} - \mathbf{b}^t).$$

Solving $\partial L^t / \partial \mathbf{b} = \mathbf{q}^t + \mathbf{H}^t(\mathbf{b} - \mathbf{b}^t) = \mathbf{0}$ for \mathbf{b} yields the iteration equations

$$\mathbf{b}^{t+1} = \mathbf{b}^t - [\mathbf{H}^t]^{-1} \mathbf{q}^t,$$

assuming \mathbf{H}^t has an inverse. Given an initial value for $t = 0$, the set of iteration equations is solved for \mathbf{b}^1 , \mathbf{b}^1 is used to solve for \mathbf{b}^2 , and so on, until the convergence criterion is satisfied. The $se(\hat{\beta})$ is calculated using the stratified jackknife method and repeating the procedure for each replicate.