



UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001

January 31, 2012

2012 AMERICAN COMMUNITY SURVEY RESEARCH AND EVALUATION REPORT
MEMORANDUM SERIES #ACS12-RER-06

MEMORANDUM FOR ACS Research and Evaluation Steering Committee

From: Jennifer Day */Signed/*
Chief (Acting), Social, Economic, and Housing Statistics Division

Prepared by: Marisa Hotchkiss
Industry and Occupation Statistics Branch
Social, Economic, and Housing Statistics Division

Subject: 2010 ACS Content Test Evaluation Report Covering Period of
Military Service

Attached is the final 2010 ACS Content Test Evaluation Report Covering Period of Military Service. This report describes the results of proposed changes to the period of military service content of the American Community Survey (ACS) as tested in the 2010 ACS Content Test.

If you have any questions about this report, please contact Melissa Chiu at (301)763-2421.

Attachment: (2010 ACS Content Test Evaluation Report Covering Period of Military Service)

cc:

ACS Research and Evaluation Team

Donna Daily (ACSO)

Todd Hughes

Debbie Klein

David Raglin

Jennifer Tancreto (DSSD)

Tony Tersine

2010 ACS Content Test Evaluation Report Covering Period of Military Service

FINAL REPORT

Marisa Hotchkiss
Social, Economic, and
Housing Statistics Division

Intentionally Blank

TABLE OF CONTENTS

EXECUTIVE SUMMARY	I
1. BACKGROUND	1
1.1 Motivation for the 2010 ACS Content Test	1
1.2 Previous Testing or Analysis.....	1
1.3 Recommendations from Cognitive Testing.....	3
1.4 Recommendations from the Expert Review Panel.....	4
2. SELECTION CRITERIA	6
3. METHODOLOGY	6
3.1 Data Collection Methods.....	6
3.2 Sample Design.....	7
3.3 Methodology Specific to the Period of Military Service Question	8
4. LIMITATIONS.....	10
5. RESEARCH QUESTIONS AND RESULTS	11
5.1 Response to the Content Test and Content Follow-Up	11
5.2 Reliability of the Data	11
5.3 Estimates of Periods of Military Service.....	13
5.4 Period of Military Service Item Missing Data Rates	13
5.5 Reliability of the Data by Mode	14
5.6 Reliability of the Data by Mail Response Stratum.....	17
5.7 Behavior Coding.....	20
6. SUMMARY	21
Acknowledgements.....	23
References.....	24
Appendix A: Additional Tables	A-1
Appendix B: Images of the Mail Versions of the Control and Test Questions.....	B-1
Appendix C: CATI and CAPI Versions of the Control and Test Questions.....	C-1
Appendix D: Flow of the Content Follow-Up.....	D-1
Appendix E: Information Page.....	E-1

LIST OF TABLES

Table 1. Content Test Response Rate Comparisons Between the Control and Test Treatments	11
Table 2. Period of Military Service Question Gross Difference Rate (GDR)	12
Table 3. Period of Military Service Question Index of Inconsistency Values	12
Table 4. Period of Military Service Question Response Distribution	13
Table 5. Period of Military Service Question Response Distribution – Mail	14
Table 6. Period of Military Service Question Response Distribution – CATI	14
Table 7. Period of Military Service Question Response Distribution – CAPI	15
Table 8. Period of Military Service Question Gross Difference Rate (GDR) – Mail	15
Table 9. Period of Military Service Question Gross Difference Rate (GDR) – CATI	16
Table 10. Period of Military Service Question Gross Difference Rate (GDR) – CAPI	16
Table 11. Period of Military Service Question Index of Inconsistency Values – Mail	16
Table 12. Period of Military Service Question Index of Inconsistency Values – CATI	17
Table 13. Period of Military Service Question Index of Inconsistency Values – CAPI	17
Table 14. Period of Military Service Item Missing Data Rates – Data Collection Mode	17
Table 15. Period of Military Service Question Response Distribution – High Response Stratum	18
Table 16. Period of Military Service Question Response Distribution – Low Response Stratum	18
Table 17. Period of Military Service Question Gross Difference Rate (GDR) – High Response Stratum	19
Table 18. Period of Military Service Question Gross Difference Rate (GDR) – Low Response Stratum	19
Table 19. Period of Military Service Question Index of Inconsistency Values – High Response Stratum	20
Table 20. Period of Military Service Question Index of Inconsistency Values – Low Response Stratum	20
Table 21. Period of Military Service Item Missing Data Rates – Mail Response Strata	20
Table 22. Period of Military Service Question Behavior Coding Results: "Standard" Question-Asking	A-1
Table 23. Period of Military Service Question Behavior Coding Results: "Standard" Initial Respondent Behavior (exact reading/slight change or correct verification) by Mode	A-1

EXECUTIVE SUMMARY

Test Objective

In late August through mid-December 2010, the Census Bureau conducted a field test of new and revised content in the 2010 American Community Survey (ACS) Content Test. The results of that testing will help determine the content to be incorporated into production ACS in 2013.

The Department of Veterans Affairs (VA) proposed several revisions to the wording of the period of military service question to simplify the question and to categorize period of military service categories in ranges that are meaningful for the VA.

The VA does not need the ACS periods of service to make the distinction between in-country service between 1961 and 1964, and Vietnam Era service 1964 to 1975; nor is there any practical significance or legal basis for May 1975 to August 1980 being collected separately from September 1980 to July 1990.

As such, the VA proposed a new version of the period of military service question that collapsed the September 1980 to July 1990 and May 1975 to August 1980 response categories, and the March 1961 to July 1964 and February 1955 to February 1961 response categories.

Methodology

The Content Test compared two versions of the period of military service question. The control version replicated the wording and response categories used in the current ACS question. The test version included the following changes to the control version of the question.

- Collapsing September 1980 to July 1990, and May 1975 to August 1980 into one response option, May 1975 to July 1990
- Collapsing March 1961 to July 1964, and February 1955 to February 1961 into one response option, February 1955 to July 1964

Research Questions and Results

Do the changes to the period of military service question improve the reliability of the data?

No. The “Korean War (July 1950 to January 1955)” had a significantly higher gross difference rate on the test version than control. For all other categories, the differences in gross difference rates and index of inconsistency values between the test and control version were not statistically significant.

Do the changes to the period of military service question result in similar response distributions for the test and control versions?

Yes. The response distributions are generally similar. However, the test question results show a statistically significant decrease in “May 1975 to July 1990,” one of the changed response categories, compared to control. For all other categories, the differences between the test and control version response distributions were not statistically significant.

Do the changes to the period of military service question adversely affect the item missing data rate?

No. The difference in the item missing data rate between the test version and the control version was not statistically significant.

For each mode of data collection, do the changes to the period of military service question affect the item missing data rate, response distribution, or reliability of the data?

There were no significant differences in the response distributions for mail and CATI. However, the results show a statistically significant decrease in the “May 1975 to July 1990” response category on the test version in the CAPI response distribution.

The differences between gross difference rates and item missing data rates by mode of data collection were not statistically significant in the CATI or CAPI response. In the mail response, “Korean War (July 1950 to January 1955)” had a significantly higher gross difference rate on the test version.

For each mail response stratum, do the changes to the period of military service affect the item missing data rate, response distribution, or reliability of the data?

In the high response stratum there was a statistically significant decrease in the “May 1975 to July 1990” response category on the test version. For all other categories, the differences between the test and control version response distributions were not statistically significant.

Does either question version elicit respondent or interviewer behaviors that may contribute to interviewer or respondent error?

The test question performed significantly better than the control for interviewers’ reading of the question. The difference in respondent behavior between the control and test series was not significant.

1. BACKGROUND

1.1 Motivation for the 2010 ACS Content Test

To evaluate proposed changes to the content of the American Community Survey (ACS), the Census Bureau conducted the 2010 ACS Content Test. The objective of the ACS Content Test, for both new and existing questions, was to determine the impact of changing question wording, response categories, and redefinition of underlying constructs on the quality of data collected.

Through the Office of Management and Budget (OMB) Interagency Committee on the ACS, subject matter experts from the Census Bureau and key data users from other federal agencies collaborated in identifying revised and new questions for inclusion in the Content Test. The suggested new and revised questions affected both the housing and detailed person sections of the ACS questionnaire.

In the housing section, the food stamps question was altered to reflect a name change for the food stamps program. In addition, a series of new questions were added related to household computer ownership and Internet subscription.

Several changes were made in the detailed person section. First, a change in data needs for the veteran series led to a revised set of response categories for the period of military service and period of military service questions. Second, the question wording of the cash public assistance income question was modified to address under-reporting of assistance on behalf of children and single payment recipients. Third, to simplify the income questions related to wages (wages, salary, commissions, bonuses, or tips) and property income (interest, dividends, rental income, royalty income or income from estates and trust), these questions were broken up into smaller questions for the Computer-Assisted Telephone Interviewing (CATI) and Computer-Assisted Personal Interviewing (CAPI) instruments only. Fourth, a set of new questions on parental place of birth were added to allow data users to divide the population into “first generation” (the foreign born), “second generation” (the children of immigrants), and “third or higher generation” (native born with no foreign-born parents).

To meet the test objective of the 2010 ACS Content Test, analysts evaluated changes to question wording, response categories, instructions, and examples relative to a control version of the question or another version for new questions. Specifically, this report discusses changes to the period of military service questions.

1.2 Previous Testing or Analysis

Certain VA benefits such as disability or death pensions require that a veteran have “wartime service,” or be a “veteran of a war.” For VA purposes, it is not required that the veteran have actually served in combat, or have even served overseas; it is only required that he or she have served at least one day during a time period designated as “wartime.”

United States Code, Title 38 defines "period of war" to include many non-declared wars, including Korean Conflict, Vietnam Era, and the Persian Gulf. Such conflicts entitle a veteran to VA benefits under Title 38. Thus it is critically important to use the correct definitions in determining eligibility for specific rights and benefits in employment. There are only six periods currently recognized by the VA as constituting "wartime service." Beginning after 1900, these periods consist of the following:

- a. *Mexican Border Period*: May 9, 1916 to April 5, 1917
- b. *World War I*: April 6, 1917 through November 11, 1918, inclusive
- c. *World War II*: December 7, 1941 through December 31, 1946, inclusive
- d. *Korean Conflict*: June 27, 1950 through January 31, 1955, inclusive
- e. *Vietnam Era*: February 28, 1961 through May 7, 1975, inclusive, if the veteran served in-country in Vietnam during that period. In all other cases August 5, 1964 through May 7, 1975, inclusive
- f. *Persian Gulf War*: August 2, 1990 through a (future) date to be prescribed by Presidential proclamation or law.

In terms of this statute, a veteran is assumed to have served in the Vietnam Era over the period March 1961 to July 1964 only if it can be ascertained that the service was "in-country" during that period meaning "in the Republic of Vietnam" during that period. However, for statistical purposes such as ACS, period of military service is only about measurement with respect to time periods. Because we cannot determine that service actually occurred in country, VA does not need the ACS periods of service to make the distinction between in-country service between 1961 and 1964, and Vietnam Era service 1964 to 1975.

According to the VA, there is also no practical significance or legal basis for May 1975 to August 1980 being collected separately from September 1980 to July 1990. Historically, these categories were created to ensure there were no date-breakups in the list of period of military service categories, and these two categories were separate in order to refine a universe for a question that is no longer asked.

As such, the Department of Veterans Affairs has proposed collapsing the categories that have either no practical significance or no legal basis for being collected separately. The VA also proposed that the periods of military service should alternate between war and peacetime periods. This collapsing would also make the ACS response categories identical to the Current Population Survey and Re-engineered Survey of Income and Program Participation response categories.

A change to the period of military service topic was tested in the 2006 American Community Survey (ACS) Content Test. The objective for including this topic was to

improve the quality of the data. The Census 2000 Content Reinterview Survey (CRS) showed moderate levels of inconsistency in the data for several period of military service categories. Anecdotal evidence suggested that respondents tended to omit reporting service in periods in which they only served briefly. Additionally, the VA had informed the Census Bureau that the existing question had more categories than were needed.

The Content Test used the existing question as the control question. The test question was a modified version of the existing question in that the parenthetical material for the categories was altered. For example, the control reads, “Vietnam era (August 1964 to April 1975)); the test reads: “August 1964 to April 1975 (Vietnam era).”

Control Version

26 When did this person serve on active duty in the U.S. Armed Forces? Mark (X) a box for EACH period in which this person served, even if just for part of the period.

- September 2001 or later
- August 1990 to August 2001 (including Persian Gulf War)
- September 1980 to July 1990
- May 1975 to August 1980
- Vietnam era (August 1964 to April 1975)
- March 1961 to July 1964
- February 1955 to February 1961
- Korean War (July 1950 to January 1955)
- January 1947 to June 1950
- World War II (December 1941 to December 1946)
- November 1941 or earlier

Test Version

26 In which period(s) did this person serve on active duty in the U.S. Armed Forces? Mark (X) all that apply.

- September 2001 or later
- August 1990 to August 2001
- May 1975 to July 1990
- August 1964 to April 1975 (Vietnam Era)
- February 1955 to July 1964
- July 1950 to January 1955 (Korean War)
- January 1947 to June 1950
- December 1941 to December 1946 (World War II)
- November 1941 or earlier

The selection criteria for choosing the test question were that, compared with the control question: (1) there was no difference in the distributions of veterans by periods of military service; (2) the item nonresponse rates were the same or lower; (3) the simple response variances were the same or lower; and (4) the number of multiple responses were the same or higher.

With respect to these criteria, the test question failed to meet the first criteria, and met the other criteria only because the results showed no statistical differences between the control and test questions (indicating that one version was not better at meeting the criteria than the other). Therefore, the test version of the question was not added to the 2008 ACS.

1.3 Recommendations from Cognitive Testing

Prior to conducting the Content Test, the Research Triangle Institute (RTI), Westat, and Research Support Services (RSS) conducted cognitive interviewing, under contract, to assist in identifying a final set of questions for the field test. Three versions of each question topic were tested with the goal of choosing the best one for the revised questions and the best two for the new questions. The questions were pretested in the three modes used in the ACS data collection (paper, telephone interview, and personal interview) in English and Spanish. Cognitive interviews consisted of one-on-one interviews using the

proposed questions in the context of the ACS survey. Survey methodologists also conducted respondent debriefings.

Of the 47 respondents, 38 understood the question as intended and answered accurately. Nine respondents demonstrated some type of confusion or misunderstanding with the question. Of these, five answered incorrectly because they included times when they were in training only; two answered incorrectly because they could not accurately recall the dates; and two answered incorrectly because they did not review the answer choices thoroughly enough before answering.

Overall, respondents reacted very favorably toward the response categories. One respondent especially liked that the response categories included “nonwar” periods. She contrasted these answer options with the typical options she sees that cover only war periods: “Again because it’s kind of hard, because my time period, when I was in, no one really pays much attention to that time period. They only mark times where actually the country’s at war.” Respondents also liked that the categories showed dates and labels. Nonetheless, two respondents initially answered incorrectly because they focused on the “era” rather than on the dates. During probing, they realized that years for another category fit as well. In addition, most respondents indicated that recalling when they or a household member served was easy; however, two respondents did have difficulty recalling the specific dates of their service, simply because the service occurred quite some time ago.

Although the response categories were not problematic for the remaining respondents, two respondents found that the presentation of the dates and time periods was somewhat confusing and suggested that they be better organized, for example, by listing the dates first and then the eras in parentheses. Currently, for some options, the date is first; for other options, the era is first.

A consistent problem for the Veteran Status question was whether training was considered active duty; this was also an issue for the Veteran Period of Military Service question. Five respondents indicated that they did not exclude time spent in training only. This was more common for respondents who were in the Reserves or National Guard (or answering about someone who was) and included the time the person was in the Reserves or National Guard, even when the person was not on active duty. Several other respondents indicated that they were thinking about their time in training when they answered this question, but that it was during a time period when they were also active.

1.4 Recommendations from the Expert Review Panel

Following the cognitive testing, an expert review panel, composed of government survey methodology experts, reviewed and added changes to the final question versions proposed to move forward from the cognitive testing into the field test. The proposed changes for each question topic were approved by the corresponding OMB interagency subcommittee responsible for initiating the research. The OMB provided final approval of the proposed changes.

The final, expert panel approved question for the paper questionnaire is as follows:

29 When did this person serve on active duty in the U.S. Armed Forces? Mark (X) a box for EACH period in which this person served, even if just for part of the period.

- September 2001 or later
- August 1990 to August 2001 (including Persian Gulf War)
- May 1975 to July 1990
- Vietnam era (August 1964 to April 1975)
- February 1955 to July 1964
- Korean War (July 1950 to January 1955)
- January 1947 to June 1950
- World War II (December 1941 to December 1946)
- November 1941 or earlier

The final, expert panel approved questions for the CATI instrument is as follows (the second line after the question was an interview instruction not read to respondents):

Did (<name>/you) serve on active duty at any time during the following periods?

[Enter all that apply, even if the person served for only part of the period. Separate with commas.](#)

- 11. September 2001 or later
- 12. August 1990 to August 2001 (including Persian Gulf War)
- 13. May 1975 to July 1990
- 15. Vietnam Era (August 1964 to April 1975)
- 16. February 1955 to July 1964
- 18. Korean War (July 1950 to January 1955)
- 19. January 1947 to June 1950
- 20. World War II (December 1941 to December 1946)
- 21. November 1941 or earlier

VETP

The final, expert panel approved questions for the CAPI instrument is as follows (the bulleted item was an interview instruction not read to respondents):

FF ? [F1]

Using Card FF, tell me each period in which [name/you] served on active duty, even if it was just for part of the period.

◆ Enter all that apply, even if the person served for only part of the period. Separate with commas.

- 11. September 2001 or later
- 12. August 1990 to August 2001 (including Persian Gulf War)
- 13. May 1975 to July 1990
- 15. Vietnam Era (August 1964 to April 1975)
- 16. February 1955 to July 1964
- 18. Korean War (July 1950 to January 1955)
- 19. January 1947 to June 1950
- 20. World War II (December 1941 to December 1946)
- 21. November 1941 or earlier

Period of Service

VETP

2. SELECTION CRITERIA

The following criteria were used to determine whether the test version of the period of military service question would be recommended.

Criterion 1 (research questions 1 and 2):

The reliability for the test version should be the same or greater than the control version and the distributions of periods of military service between the control and test versions should have minimal to no differences.

Criterion 2 (research question 3):

The item missing data rates for the test version should be the same or lower than the control version.

3. METHODOLOGY

3.1 Data Collection Methods

The initial stages of the Content Test consisted of content determination, cognitive laboratory pretesting, and expert reviews for the purpose of developing alternate versions of question content. The field test portion of the ACS Content Test used the data collection methodology currently used in the production ACS (i.e., mail questionnaire, follow-up CATI, and follow-up CAPI) with an added reinterview conducted via a CATI instrument known as the Content Follow-Up (CFU). Additional data were collected on respondent and interviewer behavior during the field test via Computer Audio Recorded

Interviewing (CARI) technologies for a subset of respondents during the CATI and CAPI follow-up modes of data collection.

The Content Test followed the same schedule and procedures for the mail, CATI, and CAPI operations as the September 2010 ACS production panel. Questionnaires were mailed to sampled households at the end of August 2010. The Content Test used an English-only mail form but the automated instruments (CATI, CAPI, and CFU) included both English and Spanish versions. Households not responding by mail and for which we had a phone number were contacted for a CATI interview during the month of October 2010. In November 2010, Census Bureau field representatives visited a sample of households that did not respond by mail or CATI to attempt a CAPI interview. The CAPI operations ended December 2, 2010.

The field test included a CATI CFU reinterview to collect additional measures for the study of response error. This operation started approximately two weeks after the initial mail out of questionnaires and ended two weeks after the end of the CAPI follow-up data collection operation. The CFU included all occupied households for which we received a response in the original interview and had a telephone number. A response was defined as a case where the household provided data through at least the first person's place of birth question for mail cases or at least a sufficient partial interview for CATI/CAPI interviews. The reinterview was conducted about 2 to 4 weeks after the original interview and with the original respondent when possible. Note that the CFU CATI interview was an abbreviated version of the original Content Test interview. The CFU instrument included the basic demographic section and only those questions preceding the questions being tested in the housing and the detailed person sections to provide context (see Appendix D for the flow of the CFU instrument).

The ACS Content Test did not include all of the production data collection operations and processes. First, while the Telephone Questionnaire Assistance program's toll-free number was available to Content Test respondents for assistance, the CATI instrument did not include content changes from the Content Test. Therefore data collected from Content Test respondents via TQA CATI interview were not included in our analysis. Second, since our objective was to study response error using unedited data, the Content Test excluded the Failed Edit Follow-up (FEFU) CATI operation and the edit and imputation data processes.

3.2 Sample Design

The 2010 Content Test consisted of a national sample of 70,000 residential addresses in the contiguous United States (the sample universe did not include Puerto Rico, Alaska, and Hawaii). The sample design for the Content Test was largely based on the ACS production sample design with some modifications to meet the test objectives. The modifications included adding an additional level of stratification by stratifying addresses into high and low mail response areas, over-sampling addresses from the low mail response areas to ensure equal response from both strata, and sampling units as pairs. The high and low mail response strata were defined based on ACS mail response rates at

the tract-level. The paired sample selection formed pairs by first systematically sampling an address within the defined sampling strata and then pairing that address with the address listed next in the geographically sorted list. However, the pair was not likely comprised of neighboring addresses. One member of the pair was randomly assigned to the control group and the other member was assigned to the test group. Those addresses assigned to the test group received the revised ACS questions and the questions new to the ACS. The control group received the current questions on the production ACS as well as different versions of the new questions.

Another modification to the production ACS sample design included adding a third sampling stage. At the first stage, the production 2010 ACS first stage sample was used as the Content Test first stage sample. At the second stage, all housing units in the ACS first stage sample not selected in the production 2010 ACS second-stage sample were selected as the Content Test second-stage sample. In addition, any units that were selected to be in other operations (e.g., training, other tests, etc.) were not selected in the Content Test second stage sample. At the third stage, addresses were selected using a sampling method similar to the production ACS second stage sample design with the exception of adding the high and low mail response stratification.

3.3 Methodology Specific to the Period of Military Service Question

The control version of the period of military service question replicated the current production question. There were two differences between the control version and the test version of the question: September 1980 to July 1990 and May 1975 to August 1980 have been collapsed to read May 1975 to July 1990, and March 1961 to July 1964 and February 1955 to February 1961 have been collapsed to read February 1955 to July 1964.

The universe for the period of military service question evaluation is the population 18 years who indicated having served in active military duty on the Veteran Identification question (i.e., ACS question 28). Veterans were defined by the response category “On active duty in the past, but not now.” Statistical significance between versions was determined using a t-test.

Reliability of the Data

To evaluate whether the changes to this question improve the reliability of the data, data from the Content Test and CFU were compared to produce the simple response variance, as measured by gross difference rates, indices of inconsistency, and the L-fold index of inconsistency between the control and test versions. Those respondents 18 years and older with a response for both the original interview and the CFU interview were included in the analyses.

The gross difference rate is the percent of inconsistent answers between the original interview and the CFU. The simple response variance, which is half of the GDR, measured the average variability, across respondents, between the responses to the period of military service question in the original interview and in the CFU. The GDR, and

subsequently the simple response variance, were calculated using the following table and formula:

CFU Response (reinterview)	Content Test Response		
	yes	no	Total
yes	a	b	a+b
no	c	d	c+d
Total	a+c	b+d	n = a+b+c+d

$$GDR = \frac{b + c}{n}$$

Statistical significance between the GDRs of each version was determined using a t-test.

The index of inconsistency (I) provides an estimate of the magnitude of response variability for the period of military service question. It is the percent of total variance due to simple response variance and was calculated as:

$$I = \frac{b + c}{\frac{1}{n} [(a+c)(c+d) + (b+d)(c+d)]}$$

For the period of military service question, an index of inconsistency was computed for each response category and an overall index of inconsistency, called the L-fold index of inconsistency, was reported for the entire distribution. The L-fold index is a weighted average of the individual indexes computed for each response category. Per the Census Bureau's general rule, index values of less than 20 percent indicate low inconsistency, 20 to 50 percent indicate moderate inconsistency, and over 50 percent indicate high inconsistency.

To evaluate response inconsistencies related to the periods of service, The 11 periods in the control version were collapsed to match the 9 periods in the test version. For the control version, the third (September 1980 to July 1990) and fourth (May 1975 to August 1980) response options and the sixth (March 1961 to July 1964) and seventh (February 1955 to February 1961) response options to the period of military service question were collapsed for this comparison (see Appendix B and C).

Missing Data

The item missing data rate is the percentage of people who were eligible for the period of military service question but did not provide a response. The formula used for computing item missing data rates is:

$$\text{item missing data rate} = \frac{\text{\# of person records with missing data for this question}}{\text{total number of respondents that are over the age of 18}} * 100$$

4. LIMITATIONS

Control and test CATI-CAPI workload assignments were not assigned using an interpenetrated experimental design. That is, interviewers were allowed to administer interviews for both control and test cases, in addition to production ACS cases. The potential risk of this approach is the introduction of a cross-contamination or carry-over effect due to the interviewer administering multiple versions of the same question item. Interviewers are trained to read the questions verbatim to minimize this risk, but there still exists the possibility that an interviewer may deviate from the scripted wording of one question version to another. This could potentially mask a treatment effect from the data collected.

The CFU reinterview was not conducted in the same mode of data collection for households that responded by mail or CAPI in the original interview since CFU interviews were only administered using a CATI mode of data collection. As a result, the data quality measures derived from the reinterview may include some bias due to the differences in mode of data collection.

Respondents needed to provide a telephone number in the original Content Test interview or the Census Bureau had to be able to find a telephone number for that unit through reverse address look-up to be included in the CFU interview. As a result, 18.4 percent of the responding households from the original interview were not eligible for the CFU reinterview.

We did not have the same respondent in the CFU that we had in the original interview for 9.1 percent of the CFU cases. This means that differences between the original interview and the CFU for these cases could be due in part to having different people answering the questions.

The Content Test does not include the production weighting adjustments for seasonal variations in ACS response patterns, nonresponse bias, and under-coverage bias. The CFU portion of the Content Test did include a unit nonresponse adjustment for those Content Test cases that responded to the Content Test, but failed to respond to the CFU. As a result, the statistics derived from the Content Test data do not provide the same level of inference as the production ACS to the entire population of housing units and persons in the contiguous United States.

The sample for the Content Test was not chosen with veterans in mind and does not include a group quarters sample. Additionally, the Content Test data is not edited as normal production data would be to account for age, employment, and period of military service inconsistencies.

5. RESEARCH QUESTIONS AND RESULTS

5.1 Response to the Content Test and Content Follow-Up

Table 1 shows the unit response rates for each of the modes of data collection and all modes combined (excluding CFU) by the control and test groups. The comparison between control and test shows that respondent participation was similar for both control and test for each of the modes of data collection and all modes combined, with the exception of the CATI mode. The test treatment produces a CATI rate of response that is 3 percentage points higher compared to that of the control. We are not able to explain the increase in response due to the test treatment for the CATI mode of data collection other than by random occurrence given that the conditions affecting unit response were equivalent between the test and control groups.

Table 1. Content Test Response Rate Comparisons Between the Control and Test Treatments

Mode	Test (%)	Standard Error (%)	Control (%)	Standard Error (%)	Test - Control (%)	Standard Error (%)	Significant
All Modes (CFU excluded)	95.4	0.2	95.7	0.2	-0.3	0.3	No
Mail	58.1	0.5	57.7	0.5	0.5	0.7	No
CATI	52.6	1.2	49.6	1.0	3.0	1.5	Yes
CAPI	90.4	0.5	91.5	0.5	-1.1	0.7	No
CFU	54.3	0.5	53.5	0.6	0.8	0.7	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test

5.2 Reliability of the Data

Do the changes to the period of military service question improve the reliability of the data?

Table 2 shows the gross difference rates, i.e. the percent of inconsistent answers between the original interview and the CFU, for the period of military service question by the control and test groups. The Korean War (July 1950 to January 1955) category had a significantly higher gross difference rate on the test version compared to the control version. However, this is not one of the modified response categories, and appears to be a result of small cell sizes in the Korean War category in the CAPI mode. For all other categories, the differences in gross difference rates between the test and control version were not statistically significant.

Table 2. Period of Military Service Question Gross Difference Rate (GDR)

Category	Test GDR	Standard Error	Control GDR	Standard Error	Test- Control (%)	Standard Error (%)	Significant
Unweighted cases	1705		1806				
September 2001 or later	2.3	0.5	3.1	0.5	-0.7	0.8	No
August 1990 to August 2001	4.2	0.8	5.6	1.0	-1.4	1.3	No
May 1975 to July 1990	8.3	0.8	7.9	1.0	0.4	1.2	No
Vietnam Era (August 1964 to April 1975)	6.4	0.9	3.9	0.5	2.5	1.0	No
February 1955 to July 1964	6.8	0.8	6.2	0.7	0.6	1.1	No
Korean War (July 1950 to January 1955)	3.0	0.6	1.4	0.3	1.6	0.6	Yes
January 1947 to June 1950	3.0	0.5	1.5	0.4	1.5	0.7	No
World War II (December 1941 to December 1946)	0.8	0.3	0.7	0.2	0.1	0.3	No
November 1941 or earlier	0.7	0.2	0.6	0.2	0.1	0.3	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

* For this family of one-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

Table 3 shows the index of inconsistency values, the estimate of the magnitude of response variability for the period of military service question, by the control and test groups. The differences in index of inconsistency values between the test and control version were not statistically significant. The Vietnam Era (August 1964 to April 1975) category had a significantly higher index of inconsistency value on the test version compared to the control version. This is not one of the modified response categories, and does not appear to be significant by data collection mode (see Tables 12, 13, and 14).

Table 3. Period of Military Service Question Index of Inconsistency Values

Category	Test Index Value	Standard Error	Control Index Value	Standard Error	Test- Control (%)	Standard Error (%)	Significant
Unweighted cases	1705		1806				
September 2001 or later	9.5	2.2	11.9	2.4	-2.4	3.3	No
August 1990 to August 2001	17.5	3.0	20.3	3.1	-2.8	4.6	No
May 1975 to July 1990	25.9	2.8	22.1	2.6	3.8	3.6	No
Vietnam Era (August 1964 to April 1975)	14.1	2.0	8.5	1.0	5.6	2.3	Yes
February 1955 to July 1964	23.0	2.7	20.3	2.1	2.7	3.4	No
Korean War (July 1950 to January 1955)	11.4	2.2	5.7	1.4	5.7	2.2	No
January 1947 to June 1950	43.7	6.8	26.1	7.9	17.6	10.1	No
World War II (December 1941 to December 1946)	4.2	1.4	4.8	1.5	-0.6	1.9	No
November 1941 or earlier	65.3	21.1	75.6	18.6	-10.3	28.0	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

* For this family of one-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

The two categories that show significant differences in reliability measures are not the same, and neither one is a category that was changed in the test version. The differences in gross difference rates and index of inconsistency values between the test and control versions of the changed categories were not statistically significant. While there is no observable improvement, the changes to the period of military service question do not seem to adversely affect the reliability of the data.

5.3 Estimates of Periods of Military Service

Do the changes to the period of military service question result in similar response distributions for the test and control versions?

Table 4 shows the response distribution for each period of military service category by the control and test groups. The response distributions are generally similar, though the test question results show a statistically significant decrease in the percent of veterans indicating service during “May 1975 to July 1990,” compared to control. Though this is one of the modified response categories, this difference appears in the CAPI data collection mode only (see tables 5, 6 and 7). Considering that the difference between the test and control version of this category appears in only one mode, and the differences between test and control distributions in other period of military service categories were not statistically significant, the test and control versions seem to produce similar response distributions.

Table 4. Period of Military Service Question Response Distribution

Category	Test (%)	Standard Error (%)	Control (%)	Standard Error (%)	Test- Control (%)	Standard Error (%)	Significant
Unweighted cases	3024		3190				
September 2001 or later	13.9	0.9	15.0	0.9	-1.2	1.2	No
August 1990 to August 2001 (including Persian Gulf War)	16.7	1.0	16.6	0.9	0.1	1.3	No
May 1975 to July 1990	21.9	1.0	26.1	1.1	-4.2	1.4	Yes
Vietnam era (August 1964 to April 1975)	35.5	1.2	34.3	1.0	1.2	1.7	No
February 1955 to July 1964	14.9	0.8	16.6	0.8	-1.7	1.2	No
Korean War (July 1950 to January 1955)	13.1	0.7	11.4	0.7	1.7	1.1	No
January 1947 to June 1950	2.4	0.3	2.6	0.4	-0.2	0.5	No
World War II (December 1941 to December 1946)	9.6	0.6	7.8	0.6	1.9	0.8	No
November 1941 or earlier	0.4	0.1	0.4	0.1	0.0	0.2	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Since multiple periods could be chosen, the cumulative frequency does not equal 100%

*For this family of two-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

5.4 Period of Military Service Item Missing Data Rates

Do the changes to the period of military service question adversely affect the item missing data rate?

The difference in the item missing data rate between the test version (2.23) and the control version (2.33) was not statistically significant, suggesting no adverse effect on the item missing data rate.

5.5 Reliability of the Data by Mode

For each mode of data collection, do the changes to the period of military service question affect the item missing data rate, response distribution, or reliability of the data?

Tables 5, 6, and 7 show the response distribution for each period of military service category by the control and test groups for mail, CATI, and CAPI. Though, there were no significant differences in the response distributions for mail and CATI, the CAPI response distribution shows a statistically significant decrease in the “May 1975 to July 1990” response category on the test version compared to control. This research question was not included in the selection criteria, but does help explain the significant decrease in the same category in Table 4.

Table 5. Period of Military Service Question Response Distribution – Mail

Category	Test (%)	Standard Error (%)	Control (%)	Standard Error (%)	Test- Control (%)	Standard Error (%)	Significant
September 2001 or later	9.7	0.9	10.6	0.7	-0.9	1.1	No
August 1990 to August 2001	14.1	0.9	13.2	0.8	0.9	1.2	No
May 1975 to July 1990	20.0	1.1	22.8	1.1	-2.8	1.6	No
Vietnam Era (August 1964 to April 1975)	38.1	1.2	36.9	1.2	1.2	1.6	No
February 1955 to July 1964	15.9	0.9	19.5	1.0	-3.6	1.4	No
Korean War (July 1950 to January 1955)	16.2	0.9	14.0	0.8	2.3	1.4	No
January 1947 to June 1950	2.8	0.4	3.0	0.5	-0.2	0.6	No
World War II (December 1941 to December 1946)	12.3	0.8	10.0	0.8	2.2	1.3	No
November 1941 or earlier	0.4	0.2	0.4	0.2	0.1	0.2	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Since multiple periods could be chosen, the cumulative frequency does not equal 100%

*For this family of two-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

Table 6. Period of Military Service Question Response Distribution – CATI

Category	Test (%)	Standard Error (%)	Control (%)	Standard Error (%)	Test- Control (%)	Standard Error (%)	Significant
September 2001 or later	9.2	2.0	6.9	1.3	2.3	2.5	No
August 1990 to August 2001	14.8	2.4	16.7	2.6	-1.8	3.6	No
May 1975 to July 1990	25.1	2.8	22.3	2.3	2.8	3.7	No
Vietnam Era (August 1964 to April 1975)	41.2	3.4	40.7	3.3	0.6	4.6	No
February 1955 to July 1964	19.0	2.2	19.5	2.2	-0.5	2.9	No
Korean War (July 1950 to January 1955)	10.3	1.7	11.7	2.0	-1.4	2.6	No
January 1947 to June 1950	3.2	1.1	2.3	1.0	0.8	1.4	No
World War II (December 1941 to December 1946)	8.4	1.5	8.8	1.7	0.4	2.3	No
November 1941 or earlier	1.0	0.5	1.7	0.9	-0.7	1.0	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Since multiple periods could be chosen, the cumulative frequency does not equal 100%

*For this family of two-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

Table 7. Period of Military Service Question Response Distribution – CAPI

Category	Test (%)	Standard Error (%)	Control (%)	Standard Error (%)	Test- Control (%)	Standard Error (%)	Significant
September 2001 or later	27.0	2.8	28.3	2.9	-1.3	3.7	No
August 1990 to August 2001	24.5	2.9	24.7	2.9	-0.2	3.9	No
May 1975 to July 1990	25.8	3.2	35.3	2.7	-9.6	3.7	Yes
Vietnam Era (August 1964 to April 1975)	26.1	3.1	26.2	2.8	-0.1	4.1	No
February 1955 to July 1964	10.5	2.0	8.7	1.6	1.8	2.4	No
Korean War (July 1950 to January 1955)	6.0	1.6	5.2	1.5	0.8	2.1	No
January 1947 to June 1950	1.0	0.6	1.7	0.8	-0.7	1.0	No
World War II (December 1941 to December 1946)	3.1	1.2	2.0	0.8	1.2	1.3	No
November 1941 or earlier	0.0	0.0	0.1	0.1	-0.1	0.1	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Since multiple periods could be chosen, the cumulative frequency does not equal 100%

*For this family of two-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

Tables 8, 9, and 10 show the gross difference rates by the control and test groups for mail, CATI, and CAPI. Mail response, had a significantly higher gross difference rate for “January 1947 to June 1950” on the test version compared to control. This period of military service category was not modified. However, the older individuals likely to choose this category may be more likely to ask a friend or relative to complete the form or the CFU interview by proxy which could affect the gross difference rates for this category. For all other categories and modes, the differences in gross difference rates between the test and control version were not statistically significant and do not seem to affect the reliability of the data.

Table 8. Period of Military Service Question Gross Difference Rate (GDR) – Mail

Category	Test GDR	Standard Error	Control GDR	Standard Error	Test- Control (%)	Standard Error (%)	Significant
September 2001 or later	2.1	0.5	1.7	0.4	0.4	0.6	No
August 1990 to August 2001	3.0	0.6	3.1	0.5	-0.1	0.8	No
May 1975 to July 1990	7.3	0.8	6.5	0.8	0.9	1.2	No
Vietnam Era (August 1964 to April 1975)	6.0	0.7	4.5	0.6	1.5	0.8	No
February 1955 to July 1964	8.2	1.0	7.7	1.0	0.5	1.5	No
Korean War (July 1950 to January 1955)	2.8	0.6	2.2	0.5	0.6	0.7	No
January 1947 to June 1950	3.6	0.7	1.5	0.4	2.1	0.8	Yes
World War II (December 1941 to December 1946)	1.0	0.4	1.1	0.3	-0.1	0.4	No
November 1941 or earlier	0.8	0.2	0.5	0.2	0.3	0.3	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

* For this family of one-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

Table 9. Period of Military Service Question Gross Difference Rate (GDR) – CATI

Category	Test	Standard	Control	Standard	Test-	Standard	Significant
	GDR	Error	GDR	Error	Control (%)	Error (%)	
September 2001 or later	3.3	1.6	3.7	2.1	-0.3	2.6	No
August 1990 to August 2001	7.6	2.6	4.9	2.0	2.7	3.5	No
May 1975 to July 1990	11.4	3.1	6.2	2.0	5.2	3.5	No
Vietnam Era (August 1964 to April 1975)	8.5	2.4	4.1	1.8	4.4	3.0	No
February 1955 to July 1964	5.8	2.0	5.4	1.9	0.4	2.8	No
Korean War (July 1950 to January 1955)	3.6	1.5	0.2	0.2	3.3	1.6	No
January 1947 to June 1950	2.4	1.0	1.8	1.1	0.7	1.5	No
World War II (December 1941 to December 1946)	0.9	0.9	0.2	0.2	0.7	0.9	No
November 1941 or earlier	1.5	0.9	2.3	1.3	-0.8	1.8	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

* For this family of one-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

Table 10. Period of Military Service Question Gross Difference Rate (GDR) – CAPI

Category	Test	Standard	Control	Standard	Test-	Standard	Significant
	GDR	Error	GDR	Error	Control (%)	Error (%)	
September 2001 or later	2.4	1.4	6.4	1.9	-4.0	2.7	No
August 1990 to August 2001	5.9	2.3	12.6	3.7	-6.7	4.7	No
May 1975 to July 1990	9.3	2.4	12.3	3.3	-3.1	3.8	No
Vietnam Era (August 1964 to April 1975)	6.3	2.4	2.2	0.7	4.1	2.7	No
February 1955 to July 1964	3.7	1.6	2.6	1.3	1.0	2.0	No
Korean War (July 1950 to January 1955)	3.3	1.8	0.0	0.0	3.3	1.8	No
January 1947 to June 1950	1.5	1.3	1.3	1.3	0.2	1.8	No
World War II (December 1941 to December 1946)	0.3	0.3	0.0	0.0	0.3	0.3	No
November 1941 or earlier	0.0	0.0	0.2	0.2	-0.2	0.2	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

* For this family of one-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

Tables 11, 12, and 13 show the index of inconsistency values by the control and test groups for mail, CATI, and CAPI. For all modes and categories, the differences in index of inconsistency values between the test and control version were not statistically significant.

Table 11. Period of Military Service Question Index of Inconsistency Values – Mail

Category	Test		Control		Test-	Standard	Significant
	Index Value	Standard Error	Index Value	Standard Error	Control (%)	Error (%)	
September 2001 or later	12.7	3.0	11.6	2.8	1.1	3.7	No
August 1990 to August 2001	14.3	2.8	14.6	2.6	-0.3	3.8	No
May 1975 to July 1990	24.3	2.6	19.7	2.4	4.6	3.6	No
Vietnam Era (August 1964 to April 1975)	12.9	1.6	9.6	1.2	3.3	1.8	No
February 1955 to July 1964	26.9	3.3	21.7	2.6	5.2	4.2	No
Korean War (July 1950 to January 1955)	8.9	1.9	7.2	1.7	1.7	2.2	No
January 1947 to June 1950	42.8	7.0	22.1	6.9	20.7	9.3	No
World War II (December 1941 to December 1946)	4.1	1.5	6.0	1.8	-1.8	2.2	No
November 1941 or earlier	75.1	20.6	67.9	25.9	7.2	31.4	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

* For this family of one-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

Table 12. Period of Military Service Question Index of Inconsistency Values – CATI

Category	Index		Standard		Test- Control (%)	Standard Error (%)	Significant
	Value	Error	Value	Error			
September 2001 or later	19.0	9.2	27.0	14.2	-7.9	17.1	No
August 1990 to August 2001	31.1	9.3	18.5	7.8	12.6	12.5	No
May 1975 to July 1990	31.3	8.5	20.5	6.5	10.8	10.1	No
Vietnam Era (August 1964 to April 1975)	17.3	5.0	8.5	3.7	8.8	6.2	No
February 1955 to July 1964	20.0	6.9	18.1	5.9	1.9	8.7	No
Korean War (July 1950 to January 1955)	16.9	6.8	0.9	0.9	16.0	6.9	No
January 1947 to June 1950	49.1	21.6	24.8	16.7	24.3	27.8	No
World War II (December 1941 to December 1946)	5.1	5.2	0.9	1.0	4.2	5.4	No
November 1941 or earlier	47.0	38.4	84.8	16.2	-37.9	45.3	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

* For this family of one-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

Table 13. Period of Military Service Question Index of Inconsistency Values – CAPI

Category	Test		Control		Test- Control (%)	Standard Error (%)	Significant
	Index Value	Standard Error	Index Value	Standard Error			
September 2001 or later	5.9	3.3	13.8	4.3	-7.9	6.1	No
August 1990 to August 2001	19.0	7.3	30.9	8.3	-11.8	11.9	No
May 1975 to July 1990	27.4	8.0	28.5	7.4	-1.2	10.0	No
Vietnam Era (August 1964 to April 1975)	17.3	6.5	5.3	1.8	12.0	7.1	No
February 1955 to July 1964	13.4	5.8	17.0	9.0	-3.7	10.1	No
Korean War (July 1950 to January 1955)	22.7	11.9	0.0	0.0	22.7	11.9	No
January 1947 to June 1950	47.3	27.8	55.2	36.9	-7.9	46.9	No
World War II (December 1941 to December 1946)	5.4	6.4	0.0	0.0	5.4	6.4	No
November 1941 or earlier	0.0	0.0	100.0	0.0	-100.0	0.0	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

* For this family of one-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

Table 14 shows the item missing data rate for the period of military service question by the control and test groups for mail, CATI, and CAPI. For all modes, the differences in the item missing data rates between the test and control version were not statistically significant. The changes to the period of military service question do not seem to affect the item missing data rate.

Table 14. Period of Military Service Item Missing Data Rates – by Data Collection Mode

Item Missing Data Rate	Test		Control		Test- Control (%)	Standard Error (%)	Significant
	(%)	Standard Error (%)	(%)	Standard Error (%)			
Mail	2.0	0.3	1.4	0.3	0.7	0.5	No
CATI	2.3	0.8	3.5	1.1	-1.2	1.4	No
CAPI	2.7	1.0	4.2	1.3	-1.5	1.5	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

* For this family of one-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

5.6 Reliability of the Data by Mail Response Stratum

For each mail response stratum, do the changes to the period of military service affect the item missing data rate, response distribution, or reliability of the data?

Tables 15 and 16 show the response distribution for each period of military service category by the control and test groups for high and low mail response strata. The high response stratum showed a statistically significant decrease in the “May 1975 to July 1990” response category on the test version compared to the control. Though this is one of the modified response categories, this difference does not appear in the overall mail data collection mode (see table 5). For all other categories and strata, the differences between the test and control version response distributions were not statistically significant. This research question was not included in the selection criteria, but the effect on the high response stratum’s response distribution is not concerning.

Table 15. Period of Military Service Question Response Distribution – High Response Stratum

Category	Test (%)	Standard Error (%)	Control (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significant
September 2001 or later	13.1	1.1	15.1	1.1	-2.0	1.4	No
August 1990 to August 2001	16.2	1.2	16.6	1.1	-0.3	1.6	No
May 1975 to July 1990	21.3	1.2	26.3	1.3	-5.0	1.7	Yes
Vietnam Era (August 1964 to April 1975)	35.9	1.5	34.4	1.2	1.5	2.0	No
February 1955 to July 1964	15.3	1.0	16.6	1.0	-1.3	1.5	No
Korean War (July 1950 to January 1955)	13.6	0.8	11.9	0.9	1.7	1.3	No
January 1947 to June 1950	2.6	0.4	2.8	0.5	-0.2	0.6	No
World War II (December 1941 to December 1946)	10.1	0.7	7.8	0.7	2.3	1.1	No
November 1941 or earlier	0.5	0.2	0.4	0.2	0.0	0.2	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Since multiple periods could be chosen, the cumulative frequency does not equal 100%

*For this family of two-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

Table 16. Period of Military Service Question Response Distribution – Low Response Stratum

Category	Test (%)	Standard Error (%)	Control (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significant
September 2001 or later	16.9	1.3	14.5	1.1	2.4	1.7	No
August 1990 to August 2001	18.6	1.3	16.8	1.2	1.8	1.5	No
May 1975 to July 1990	24.4	1.3	25.6	1.2	-1.2	1.7	No
Vietnam Era (August 1964 to April 1975)	34.0	1.5	34.2	1.3	-0.2	2.0	No
February 1955 to July 1964	13.1	1.0	16.3	0.9	-3.2	1.3	No
Korean War (July 1950 to January 1955)	11.0	0.9	9.3	0.7	1.8	1.1	No
January 1947 to June 1950	1.7	0.3	1.7	0.4	0.0	0.5	No
World War II (December 1941 to December 1946)	7.6	0.7	7.4	0.7	0.1	0.9	No
November 1941 or earlier	0.1	0.1	0.3	0.2	-0.2	0.2	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

Note: Since multiple periods could be chosen, the cumulative frequency does not equal 100%

*For this family of two-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

Tables 17 and 18 show the gross difference rates by the control and test groups for high and low mail response strata. For all strata and categories, the differences in gross difference rates between the test and control version were not statistically significant.

Table 17. Period of Military Service Question Gross Difference Rate (GDR) – High Response Stratum

Category	Test	Standard	Control	Standard	Test-	Standard	Significant
	GDR	Error	GDR	Error	Control (%)	Error (%)	
September 2001 or later	2.3	0.6	2.6	0.6	-0.3	0.9	No
August 1990 to August 2001	3.9	0.9	5.6	1.1	-1.7	1.6	No
May 1975 to July 1990	7.6	1.1	7.2	1.1	0.4	1.5	No
Vietnam Era (August 1964 to April 1975)	6.0	1.1	3.1	0.5	2.9	1.2	No
February 1955 to July 1964	6.8	1.0	6.1	0.8	0.7	1.3	No
Korean War (July 1950 to January 1955)	3.0	0.7	1.5	0.4	1.5	0.7	No
January 1947 to June 1950	3.4	0.7	1.6	0.5	1.8	0.9	No
World War II (December 1941 to December 1946)	0.9	0.3	0.8	0.3	0.1	0.4	No
November 1941 or earlier	0.7	0.2	0.5	0.2	0.2	0.3	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

* For this family of one-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

Table 18. Period of Military Service Question Gross Difference Rate (GDR) – Low Response Stratum

Category	Test	Standard	Control	Standard	Test-	Standard	Significant
	GDR	Error	GDR	Error	Control (%)	Error (%)	
September 2001 or later	2.5	0.7	5.2	1.3	-2.7	1.5	No
August 1990 to August 2001	5.4	1.1	5.6	1.2	-0.2	1.6	No
May 1975 to July 1990	10.5	1.3	10.7	1.3	-0.1	1.7	No
Vietnam Era (August 1964 to April 1975)	7.7	1.1	7.3	1.1	0.4	1.6	No
February 1955 to July 1964	6.6	1.0	6.7	0.9	-0.1	1.5	No
Korean War (July 1950 to January 1955)	2.9	0.6	1.1	0.3	1.8	0.7	No
January 1947 to June 1950	1.6	0.5	1.2	0.4	0.4	0.7	No
World War II (December 1941 to December 1946)	0.7	0.2	1.0	0.4	-0.4	0.5	No
November 1941 or earlier	0.5	0.3	0.5	0.2	0.0	0.4	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

* For this family of one-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

Tables 19 and 20 show the index of inconsistency values for by the control and test groups for high and low mail response strata. The high response stratum had a significantly lower index of inconsistency value for “August 1990 to August 2001” on the test version compared to the control. This period of military service category was not modified, but a lower index of inconsistency value is still encouraging. For all other categories, the differences in index of inconsistency values between the test and control version were not statistically significant.

Table 19. Period of Military Service Question Index of Inconsistency Values – High Response Stratum

Category	Test		Control		Test- Control (%)	Standard Error (%)	Significant
	Index Value	Standard Error	Index Value	Standard Error			
September 2001 or later	9.6	2.5	9.8	2.5	-0.1	3.7	No
August 1990 to August 2001	17.1	4.0	20.2	3.8	-3.2	5.9	Yes
May 1975 to July 1990	24.2	3.6	20.3	3.1	4.0	4.4	No
Vietnam Era (August 1964 to April 1975)	13.3	2.4	6.8	1.2	6.5	2.7	No
February 1955 to July 1964	22.7	3.3	20.2	2.4	2.5	4.1	No
Korean War (July 1950 to January 1955)	11.1	2.6	5.8	1.6	5.4	2.5	No
January 1947 to June 1950	46.6	8.1	26.7	9.3	19.8	11.2	No
World War II (December 1941 to December 1946)	4.4	1.7	5.2	1.7	-0.8	2.2	No
November 1941 or earlier	59.4	23.2	72.4	25.7	-13.0	33.9	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

* For this family of one-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

Table 20. Period of Military Service Question Index of Inconsistency Values – Low Response Stratum

Category	Test		Control		Test- Control (%)	Standard Error (%)	Significant
	Index Value	Standard Error	Index Value	Standard Error			
September 2001 or later	9.0	2.4	22.2	5.3	-13.2	5.9	No
August 1990 to August 2001	19.0	3.8	20.7	4.4	-1.8	5.8	No
May 1975 to July 1990	31.7	3.6	30.3	3.4	1.4	4.8	No
Vietnam Era (August 1964 to April 1975)	16.9	2.4	15.6	2.4	1.3	3.5	No
February 1955 to July 1964	24.2	3.3	20.6	2.9	3.6	4.7	No
Korean War (July 1950 to January 1955)	12.6	2.8	5.2	1.5	7.5	3.2	No
January 1947 to June 1950	29.6	9.7	23.1	8.2	6.5	13.6	No
World War II (December 1941 to December 1946)	3.2	1.8	3.0	1.3	0.2	2.2	No
November 1941 or earlier	100.2	0.2	83.4	16.0	16.8	16.0	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

* For this family of one-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

Table 21 shows the item missing data rate for the veteran status question by the control and test groups for high and low mail response strata. For both strata, the differences in the item missing data rates between the test and control version were not statistically significant.

Table 21. Period of Military Service Item Missing Data Rates – Mail Response Strata

Item Missing Data Rate	Test		Control		Test- Control (%)	Standard Error (%)	Significant
	(%) Error	(%) Error	(%) Error	(%) Error			
High Response Stratum	2.0	0.4	2.1	0.4	-0.1	0.6	No
Low Response Stratum	3.2	0.5	3.5	0.6	-0.2	0.7	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010

* For this family of one-sided hypothesis tests, the family-wise error rate has been controlled using the Bonferroni-Holm multiple comparison method at the $\alpha = 0.10$ level.

5.7 Behavior Coding

Does either question version elicit respondent or interviewer behaviors that may contribute to interviewer or respondent error?

The test question performed significantly better than the control for interviewers' reading of the question, by 28 percentage points. On the respondent side, the difference between test and control was not significant but the magnitude of the difference was quite high and in the opposite direction – in the control version respondents gave a standard answer 75% of the time versus in the test where the rate of a standard answer was only 55%.

See Tables 22 and 23 in Appendix A.

6. SUMMARY

The response distributions and gross difference rates in the test version are not significantly different from the response distributions and gross difference rates in the control version in the majority of the categories. Additionally, the three categories that showed significant differences were not the same throughout:

- The “May 1975 to July 1990” category, one of the changed response categories, showed a statistically significant decrease in the percent serving on the test version compared to the control version, a difference of 4.2 percentage points. Only the CAPI data collection mode showed this significant difference. There were no significant differences in the response distributions for mail and CATI.
- The “Korean War (July 1950 to January 1955)” category had a significantly higher gross difference rate on the test version than on the control version, a difference of 1.6. However, this significant difference in the gross difference rate appears to be a result of small cell sizes in the CAPI mode.
- The “Vietnam Era (August 1964 to April 1975)” had a significantly higher index of inconsistency value on the test version than on the control, a difference of 5.6. This difference was not repeated by data collection mode or mail response strata.

Per the VA, these unexplained differences are trivial compared to the larger question of whether information should be collected that is not useful. Their justification for recommending the new test version is based on the fact that the test version eliminates two response options, putting an end to collecting useless data. The usefulness of the information generated is their first criteria in recommending any modification to the ACS questions.

Importantly, the test question also performed significantly better than the control for interviewers' reading of the question, and the difference in respondent behavior between the control and test series was not significant.

Overall, accepting the test version of these response categories would mean that ACS categories would collect meaningful period of military service information for the VA,

would match CPS and Re-engineered SIPP categories, and could also reduce respondent burden by reducing the number of response categories.

Acknowledgements

We would like to thank the following Census Bureau staff for their valuable contributions and assistance to the development and analysis of this project: Donna Daily, Patricia Goerman, Kelly Holder, Todd Hughes, Debra Klein, Joanne Pascale, David Raglin, Michelle Ruitter, Jennifer Tancreto, Anthony G. Tersine Jr., and Mary Frances Zelenak.

References

Clark S. and Raglin D. (2006). "Evaluation Report Covering Period of Military Service," 2006 American Community Survey Content Test Report P.5.b, U.S. Census Bureau, Washington

Pascale J. and Goerman P. (2010) "ACS 2010 Content Test Behavior Coding Report" U.S. Census Bureau, Washington

RTI International (2009) "Cognitive Testing of the American Community Survey Content Test Items" RTI International, Research Triangle Park

Singer P. and Ennis S. (2003), "Census 2000 Content Reinterview Survey: Accuracy of Data for Selected Population and Housing Characteristics as Measured by Reinterview", Census 2000 Evaluation B.5, U.S. Census Bureau, Washington

Appendix A: Additional Tables

Table 22. Period of Military Service Question Behavior Coding Results: "Standard" Question-Asking

Category	Test		Control		Test- Control		Significant
	% Standard	Standard Error	% Standard	Standard Error	(%)	(%)	
Overall	61.0	0.1	33.0	0.1	28.0	0.1	Yes
CAPI	54.0	0.1	31.0	0.1	23.0	0.2	No
CATI	70.0	0.2	36.0	0.2	34.0	0.2	No

Table 23. Period of Military Service Question Behavior Coding Results: "Standard" Initial Respondent Behavior

Category	Test		Control		Test- Control		Significant
	% Standard	Standard Error	% Standard	Standard Error	(%)	(%)	
Overall	55.0	0.1	75.0	0.1	-20.0	0.1	No
CAPI	33.0	0.1	69.0	0.1	-36.0	0.2	No
CATI	80.0	0.1	82.0	0.1	-2.0	0.2	No

Appendix B: Images of the Mail Versions of the Control and Test Questions

Figure B-1. Control Version of the Period of military service Question

29 When did this person serve on active duty in the U.S. Armed Forces? Mark (X) a box for EACH period in which this person served, even if just for part of the period.

- September 2001 or later
- August 1990 to August 2001 (including Persian Gulf War)
- September 1980 to July 1990
- May 1975 to August 1980
- Vietnam era (August 1964 to April 1975)
- March 1961 to July 1964
- February 1955 to February 1961
- Korean War (July 1950 to January 1955)
- January 1947 to June 1950
- World War II (December 1941 to December 1946)
- November 1941 or earlier

Figure B-2. Test Version of the Period of military service Question

29 When did this person serve on active duty in the U.S. Armed Forces? Mark (X) a box for EACH period in which this person served, even if just for part of the period.

- September 2001 or later
- August 1990 to August 2001 (including Persian Gulf War)
- May 1975 to July 1990
- Vietnam era (August 1964 to April 1975)
- February 1955 to July 1964
- Korean War (July 1950 to January 1955)
- January 1947 to June 1950
- World War II (December 1941 to December 1946)
- November 1941 or earlier

Appendix C: CATI and CAPI Versions of the Control and Test Questions

CATI Control Question:

MILP

Did (<Name>/you) serve on active duty during:	
<ul style="list-style-type: none">◆ Read all answer categories.◆ Enter all that apply, even if the person served for only part of the period.◆ Separate with commas.	
<input type="checkbox"/> 11. September 2001 or later	<input type="checkbox"/> 16. March 1961 to July 1964
<input type="checkbox"/> 12. August 1990 to August 2001 (including Persian Gulf War)	<input type="checkbox"/> 17. February 1955 to February 1961
<input type="checkbox"/> 13. September 1980 to July 1990	<input type="checkbox"/> 18. Korean War (July 1950 to January 1955)
<input type="checkbox"/> 14. May 1975 to August 1980	<input type="checkbox"/> 19. January 1947 to June 1950
<input type="checkbox"/> 15. Vietnam Era (August 1964 to April 1975)	<input type="checkbox"/> 20. World War II (December 1941 to December 1946)
	<input type="checkbox"/> 21. November 1941 or earlier
When Serve?	

CAPI Control Question:


MILP

<input type="checkbox"/> <input type="checkbox"/> F	
Using Card F, please tell me each period in which (<name>/you) served on active duty, even if it was just for part of the period.	
<ul style="list-style-type: none">◆ Enter all that apply, even if the person served for only part of the period.◆ Separate with commas.	
<input type="checkbox"/> 11. September 2001 or later	<input type="checkbox"/> 16. March 1961 to July 1964
<input type="checkbox"/> 12. August 1990 to August 2001 (including Persian Gulf War)	<input type="checkbox"/> 17. February 1955 to February 1961
<input type="checkbox"/> 13. September 1980 to July 1990	<input type="checkbox"/> 18. Korean War (July 1950 to January 1955)
<input type="checkbox"/> 14. May 1975 to August 1980	<input type="checkbox"/> 19. January 1947 to June 1950
<input type="checkbox"/> 15. Vietnam Era (August 1964 to April 1975)	<input type="checkbox"/> 20. World War II (December 1941 to December 1946)
	<input type="checkbox"/> 21. November 1941 or earlier
When Serve?	

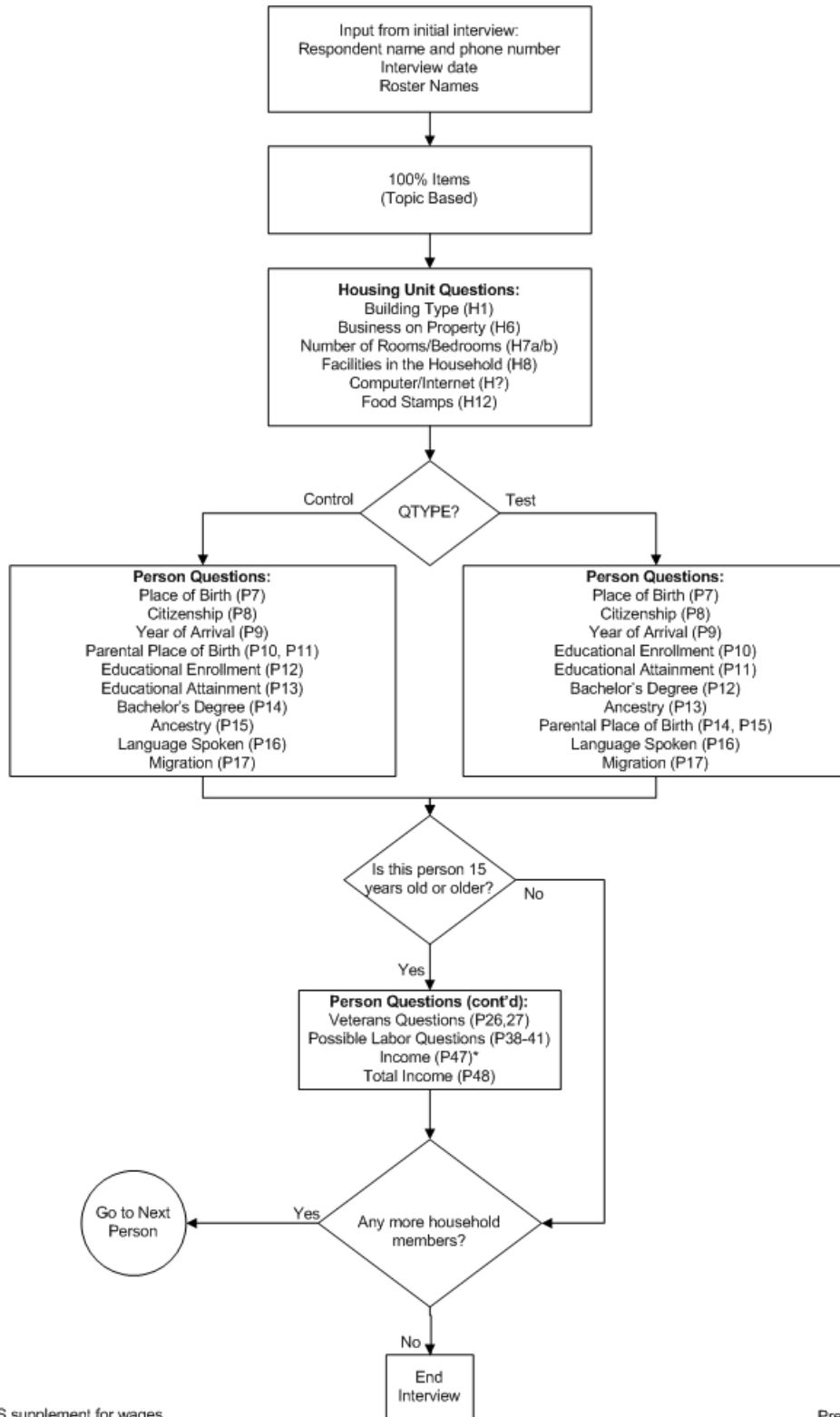
CATI Test Questions:

<p>Did (<name>/you) serve on active duty at any time during the following periods?</p> <p>Enter all that apply, even if the person served for only part of the period. Separate with commas.</p>
<p><input type="radio"/> 11. September 2001 or later</p> <p><input type="radio"/> 12. August 1990 to August 2001 (including Persian Gulf War)</p> <p><input type="radio"/> 13. May 1975 to July 1990</p> <p><input type="radio"/> 15. Vietnam Era (August 1964 to April 1975)</p> <p><input type="radio"/> 16. February 1955 to July 1964</p> <p><input type="radio"/> 18. Korean War (July 1950 to January 1955)</p> <p><input type="radio"/> 19. January 1947 to June 1950</p> <p><input type="radio"/> 20. World War II (December 1941 to December 1946)</p> <p><input type="radio"/> 21. November 1941 or earlier</p>
<p>VETP</p>

CAPI Test Questions:

<p> FF ? [F1]</p> <p>Using Card FF, tell me each period in which [name/you] served on active duty, even if it was just for part of the period.</p> <p>◆ Enter all that apply, even if the person served for only part of the period. Separate with commas.</p>
<p><input type="radio"/> 11. September 2001 or later</p> <p><input type="radio"/> 12. August 1990 to August 2001 (including Persian Gulf War)</p> <p><input type="radio"/> 13. May 1975 to July 1990</p> <p><input type="radio"/> 15. Vietnam Era (August 1964 to April 1975)</p> <p><input type="radio"/> 16. February 1955 to July 1964</p> <p><input type="radio"/> 18. Korean War (July 1950 to January 1955)</p> <p><input type="radio"/> 19. January 1947 to June 1950</p> <p><input type="radio"/> 20. World War II (December 1941 to December 1946)</p> <p><input type="radio"/> 21. November 1941 or earlier</p>
<p>Period of Service <input type="text"/></p>
<p>VETP</p>

Appendix D: Flow of the Content Follow-Up



* using CPS supplement for wages,
property income & cash public
assistance

Prepared by Michelle
Ruiter and Mary Davis

Appendix E: Information Page

PERIOD OF MILITARY SERVICE INFORMATION PAGE

2010 ACS CONTENT TEST

Megha Joshipura, Mary Frances Zelenak, DSSD

Baselined 11/16/2009

Test Design

Treatments	Two question versions with different wording (see page 3).
Sample Size	35,000 households per treatment (70,000 total)
Sample Design	Similar to production ACS with an additional level of stratification into high and low mail response areas.
Modes	Mail, CATI, and CAPI, with a CATI content follow-up (CFU) of all households. <i>CATI and CAPI interviews will be recorded using Computer-Assisted Recorded Interviewing (CARI) technology.</i>
Time Frame	Same schedule as the production September panel: mailout in late August, CATI in October, CAPI in November. CFU goes from mid-September to mid-December.

Research Questions & Evaluation Measures

No.	Research Questions	Evaluation Measures
1	Do the changes to the period of military service question improve the reliability of the data?	Using data from the Content Test and the Content Follow-up (CFU), compare the simple response variance, indices of inconsistency, and the L-fold index of inconsistency between the control and test versions.
2	Do the changes to the period of military service question result in similar response distributions for the test and control versions?	Compare the response distributions between the control and test versions.
3	Do the changes to the period of military service question lower the item missing data rate?	Compare the item nonresponse rate between the control and test versions.

No.	Research Questions	Evaluation Measures
4	For each mode of data collection, do the changes to the period of military service question affect the item missing data rate, response distribution, or reliability of the data?	<p>For each mode (mail, CATI, CAPI), compare the item missing data rates, response distributions, and reliability measures between the control and the test versions.</p> <p><i>Comparisons across modes of data collection cannot be made since measurable differences cannot be attributed strictly to the mode of data collection. Observed differences across modes may also be due to mode specific respondent characteristics and reinterview mode effects (CFU only).</i></p>
5	For each mail response stratum, do the changes to the period of military service affect the item missing data rate, response distribution, or reliability of the data?	For each mail response stratum (high and low), compare the item missing data rates, response distributions, and reliability measures between the control and the test versions.
6	Does either question version elicit respondent or interviewer behaviors that may contribute to interviewer or respondent error?	Compare the behavior coding results derived from the CARI recordings between the control and the test versions.

Selection Criteria (In order of priority)

Research Question(s)	Criteria
2	The reliability for the test version should be the same or greater than the control version and the distributions of periods of military service between the control and test versions should have minimal to no differences.
3	The item missing data rates for the test version should be the same or lower than the control version.

Supplemental Information

Research Question(s)	Criteria
4-6	Not part of the selection criteria. These data are presented to give additional information regarding how the questions performed.

Question Wording

Current ACS Version	Content Test Version
<p>Q.29 When did this person serve on active duty in the U.S. Armed Forces? Mark (X) a box for EACH period in which this person served, even if just for part of the period.</p> <ul style="list-style-type: none"><input type="checkbox"/> September 2001 or later<input type="checkbox"/> August 1990 to August 2001 (including Persian Gulf War)<input type="checkbox"/> September 1980 to July 1990<input type="checkbox"/> May 1975 to August 1980<input type="checkbox"/> Vietnam Era (August 1964 to April 1975)<input type="checkbox"/> March 1961 to July 1964<input type="checkbox"/> February 1955 to February 1961<input type="checkbox"/> Korean War (July 1950 to January 1955)<input type="checkbox"/> January 1947 to June 1950<input type="checkbox"/> World War II (December 1941 to December 1946)<input type="checkbox"/> November 1941 or earlier	<p>Q.29 When did this person serve on active duty in the U.S. Armed Forces? Mark (X) a box for EACH period in which this person served, even if just for part of the period.</p> <ul style="list-style-type: none"><input type="checkbox"/> September 2001 or later<input type="checkbox"/> August 1990 to August 2001 (including Persian Gulf War)<input type="checkbox"/> May 1975 to July 1990<input type="checkbox"/> Vietnam Era (August 1964 to April 1975)<input type="checkbox"/> February 1955 to July 1964<input type="checkbox"/> Korean War (July 1950 to January 1955)<input type="checkbox"/> January 1947 to June 1950<input type="checkbox"/> World War II (December 1941 to December 1946)<input type="checkbox"/> November 1941 or earlier