2010 ACS Content Test Evaluation Report Covering Veteran Status

Marisa Hotchkiss
Social, Economic, and
Housing Statistics Division

United States™ Census Bureau

Intentionally Blank

TABLE OF CONTENTS

LIST OF TABLES

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 veteran identification question to simplify the question and to generate more reliable and accurate estimates of veterans. Results from the 2006 ACS Content Test had suggested that the complexity of the existing question may result in the undercounting of veterans.

The VA's proposal was based on several factors. First, the VA no longer has a need to distinguish between veterans who served in the past 12 months and those who served over 12 months ago. Secondly, previous analysis of this question suggested that respondents may not fully read the three consecutive "yes" response options, but rather check the first "yes" checkbox to indicate that they had served in the military. This results in some respondents being classified as "on active duty now" rather than "on active duty in the past" and potentially undercounts the number of veterans. For this reason, removal of the lead-in "yes" and "no" to the response options is tested. Finally, the instructions for this question refer to service in the Persian Gulf War of 1991 rather than the more recent wars in Iraq and Afghanistan, and is modified and tested accordingly.

Methodology

The 2010 ACS Content Test compared two versions of the veteran status question. The control version is a modification of the current ACS question, but collapses the response options for veterans who served in the past 12 months and veterans who served over 12 months ago into one category, veterans on active duty in the past, but not now. Additionally, the instructions for this question were modified to refer to the more recent wars in Iraq and Afghanistan. (See Appendix B for question wording.)

The test version included the following changes to the control version of the question.

- Added a question instruction, "Mark (X) ONE box," but removed the other test instructions
- Removed the lead-in "yes" and "no" from the response options
- Reordered response options

Research Questions and Results

Do the changes to the veteran status question result in a similar or higher estimate of veterans?

Yes. The differences in response distributions between the test and control version were not statistically significant.

Do the changes to the veteran status question decrease the number of respondents that classified themselves as "Yes, on active duty now" but marked a period of service that was not "September 2001 or later"?

Yes. The test version showed a statistically significant decrease (by 12.7 percentage points) in the percent of respondents that classified themselves as active duty, but did not mark the most recent period of service.

Do the changes to the veteran status question adversely affect the item missing data rate?

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

Do the changes to the veteran status question result in a lower percentage of people who mark "now on active duty," and consequently, a higher percentage of people who mark "on active duty in the past, but not now"?

No. The differences in response distributions between the test and control version were not statistically significant.

Do the changes to the veteran status question improve the reliability of the data?

Yes. The "Never Served" and "Training Only" response categories had significantly lower gross difference rates, and index values on the indices of inconsistency on the test version. The L-fold index value was also significantly lower on the test version (7.4) than on the control (8.9). 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 veteran status question affect the item missing data rate for service-connected disability?

Yes. The test version showed a statistically significant decrease in the item missing data rates for service-connected disability.

Do the changes to veteran status question reduce the occurrence of multiple answers?

Yes. The test version showed a statistically significant decrease in the percent of multiple answers.

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

Mail response, for the test version, showed a statistically significant decrease in the percent indicating "On Active Duty Now" and a statistically significant decrease in the gross difference rates and index of inconsistency values for "Never Served" and "Training Only" compared to the control version. CAPI response showed a statistically significant increase in respondents indicating "Training Only" on the test version. CATI response showed a statistically significant increase in the item missing data rate on the test version.

For each mail response stratum, do the changes to the veteran status question affect the item missing data rate, response distribution, or reliability of the data?

The low response stratum showed a statistically significant increase in the percent indicating "Training Only" on the test version. The high response stratum showed a statistically significant decrease in the gross difference rates for "Never Served" and "Training Only" on the test version.

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

The test series of veteran status questions performed significantly better for both interviewer and respondent behavior than the control series.

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 veteran status 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 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 veteran status questions.

1.2 Previous Testing or Analysis

A military service status topic was tested in the 2006 American Community Survey (ACS) Content Test. The objective of this test was to improve the count of civilian veterans and test whether the ACS could produce an accurate count of people who are currently on active duty in the Armed Forces and have a prior discharge that

qualifies them as a veteran. Comparisons of veteran-status estimates from the ACS with those from the Department of Veterans Affairs (VA) suggested that the ACS may be underestimating the number of civilian veterans, and the VA was concerned that not having accurate counts of military veterans leads to underestimates of the VA's potential clientele and the future demand for VA services.

The existing ACS question served as the control question in the Content Test. The control question used a one-part question to collect several items of information: (1) the existence of current or prior military service; the type of service (active-duty /Reserve or National Guard Training only); and the timing of active-duty. The test version used a two-part question to collect the first two items, but instead of the third, it asked whether the respondent was a military veteran.

The primary goal of the test question was to provide more accurate estimates of the number of civilian veterans. Not only did the test produce a lower estimate of civilian veterans than the control, it also produced a higher item nonresponse rate, and generally higher net difference rates and simple response variances. These findings suggest that respondents had more difficulty answering the test question than the control.
The test question did not adequately identify and count military veterans either. A major impediment to evaluating the results was the exclusion of people in group quarters and the military population serving overseas from the sample. The test question did not meet its primary or secondary goal. 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 asked the veteran status question, 32 understood the question as intended and answered accurately. Sixteen respondents demonstrated some type of confusion or misunderstanding with the question. Of these 16 respondents, only 2 answered the question inaccurately for their situations. Both respondents answered that they were on active duty when they had only been in training for the National Guard. The remaining 14 had some type of difficulty with the question but were able to answer accurately for their situations.

The following revisions to the question were recommended:

- Respondents noted that the question was confusing because it says that "active duty" does not include training, yet respondents noted that training *is* active duty. In addition, the question says to exclude people who only had training, but only being in training is a response option. The clarification phrase did not appear to help. If the clarification phrase must be included, it was recommended that instead of clarifying "active duty," we tell respondents specifically what to exclude (e.g., "Do NOT include active duty for training") before indicating what to include.
- Some respondents focused on the locations "Iraq, Afghanistan, or elsewhere" and were not thinking of the United States or other locations overseas. The recommendation was to change the location to refer to "the United States or overseas."
- Although most respondents seemed to understand the term "activation," they frequently used other terms, including "called up," "mobilization," and "deployment." The recommendation was to use these additional terms to ensure that the question is understood by as many respondents as possible.
- Several respondents did not read the clarification for active duty because they felt confident that they knew what "active duty" meant. To increase the likelihood that the instruction are read, the recommendation was that it not be italicized but that it instead uses the same typeface as the rest of the question.
- Particularly in Spanish, respondents preferred the term "military Reserves" to "Reserves." We recommend using the phrase "la Reserva Militar" in all instances for the Spanish question. In Version 3 of the English question, the first reference is to "military Reserves," and the remaining references are to "Reserves." The recommendation was to preserve this approach because it clarifies the term "Reserves" upon introduction but then shortens it for the rest of the question.
- A recommendation was made to change the response options for the self-administered questionnaire in the following ways:
    - Change the option, "no, training for the Reserves or National Guard only," to "only active duty for training (ADT) for the National Guard or Reserves." Placing the "only active duty for training (ADT) for the National Guard or Reserves" option before either of the "active duty" options also may encourage more respondents to appropriately select the "active duty training" response.
    - Change the sequence of response options by placing the two "no" responses together, followed by the two "yes" responses, but remove the "yes" and "no" because they may encourage respondents to skip reading some of the options.

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.

See Appendices B and C for the final versions of the question that were tested.

2. SELECTION CRITERIA

The following criteria were used to determine whether the test version of the veteran status question would be recommended.

Criterion 1 (research questions 1 and 2):
The estimate of veterans from the test version is comparable to or higher than the estimate from the control version. Additionally, there are fewer inconsistencies between the veteran status question and the period of service question, thus requiring less editing.

Criterion 2 (research questions 3, 4 and 5):
The item missing data rates and reliability measures, along with seeing an increase in the number of respondents who mark that they are veterans and seeing a decrease in the number of respondents currently on active duty when comparing the control and test versions, will be considered together when determining which question version performs better.

Criterion 3 (research questions 6 and 7):
The item missing data rates for service-connected disability along with the occurrence of multiple answers 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 Veteran Status Question

The control version of the veteran status question was a modification of the current production question. The Department of Veterans Affairs (VA) no longer needs the 12-month distinction in the veteran classifications, resulting in the collapsing of response categories from five categories in the current question to four categories in the control question (see Appendix B and C). This modified control version was used in production prior to 2003. The only difference between the modified control version and the 2002 question is in the instructions for the question, which have been updated to reference the wars in Iraq and Afghanistan while the 2002 instructions refer to the Persian Gulf War.

There are three differences between the modified control version and the test version of the question. First, the current production version of the veteran status question includes multiple "yes" response options. Analysis of this question suggested that the respondents may not fully read the "Yes" response options, and mistakenly choose the first yes response "Yes, now on active duty" rather than "Yes, on active duty in the past, but not now." To correct this problem, the test version of the question eliminated the lead-in "yes" and "no" for each response option. Second, the response categories in the test version were ordered differently so that the first response option presented to respondents was not "Yes, now on active duty." Third, the response category for Reservists in the test version incorporated the information from the instruction in the control version. There were no instructions for the test version of the question.

The universe for the veteran status question evaluation is the population 18 years and older. Veterans are 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 response inconsistencies related to the estimate of veterans, the percentage of respondents in the control version that classified themselves as "Yes, on active duty now" but marked a period of service that was not "September 2001 or later" were compared to

the percentage of respondents in the test version that classified themselves as "On active duty now" and did not mark the most recent period of service. This comparison was based on the belief that respondents may be more likely to mark the veteran status question incorrectly than the period of service question. A higher amount of inconsistencies between these two variables in the control version could be an indicator that veterans were marking the first "yes" category of the veteran status question in error. 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 are 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 veteran status question in the original interview and in the CFU. We calculated the GDR, and subsequently the simple response variance, 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 veteran status question. It is the percent of total variance due to simple response variance and is calculated as:

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

For the veteran status question, an index of inconsistency was computed for each response category and an overall index of inconsistency, called the L-fold index of inconsistency, is 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.

Missing Data

For the veteran status question and the service-connected disability status rating, the item missing data rate is the percentage of people who were eligible for the veteran status question but did not provide a response. The formula used for computing item missing data rates is:

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

Statistical significance between versions was determined using a t-test.

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 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 Estimate of Veterans

Do the changes to the veteran status question result in a similar or higher estimate of veterans?

Table 2 shows the response distribution for each veteran status category by the control and test groups. The differences in response distributions between the test and control

version were not statistically significant. The test version generated a similar estimate of veterans.

Table 2. Veteran Status Question Response Distribution

Category	Test (%)	Standard Error (%)	Control (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significant
Unweighted sample size	34,027		34,029				
Never Served on Active Duty	89.3	0.2	88.9	0.2	0.4	0.3	No
Training Only	1.6	0.1	1.3	0.1	0.2	0.1	No
On Active Duty Now	0.5	0.1	0.6	0.1	-0.1	0.1	No
Active Duty in the Past, but Not Now	8.7	0.2	9.2	0.2	-0.5	0.3	No
Total	100.0		100.0				

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 multiple comparison method at the $\alpha = 0.10$ level.

5.3 Veteran Status and Period of Service Inconsistencies

Do the changes to the veteran status question decrease the number of respondents that classified themselves as "Yes, on active duty now" but marked a period of service that was not "September 2001 or later"?

Table 3 shows the percent of inconsistent responses related to the estimate of veterans. It compares the percentage of respondents in the control version that classified themselves as "Yes, on active duty now" but marked a period of service that was not "September 2001 or later" to the percentage of respondents in the test version that classified themselves as "On active duty now" and did not mark the most recent period of service. The test version shows a statistically significant decrease (by 12.7 percentage points) in the percent of respondents that classified themselves as active duty, but did not mark the most recent period of service, reducing a major source of inconsistency.

Table 3. Percent of Inconsistent Responses

Category	Test (%)	Standard Error (%)	Control (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significant
Unweighted cases	144		181				
"On active duty now" but did not choose "September 2001 or later"	7.6	2.7	20.2	3.6	-12.7	4.1	Yes

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

*Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a one-sided test

5.4 Veteran Status Item Missing Data Rates

Do the changes to the veteran status question adversely affect the item missing data rate?

The difference in the item missing data rate between the test version (6.0) and the control version (5.9) was not statistically significant, suggesting no adverse effect on the item missing data rate.

5.5 Estimate of Veteran and Active-duty

Do the changes to the veteran status question result in a lower percentage of people who mark "now on active duty," and consequently, a higher percentage of people who mark "on active duty in the past, but not now"?

Table 2, above, shows the response distribution for each veteran status category by the control and test groups. The difference in response distributions between the test and control version was not statistically significant. The changes to the veteran status question did not result in a lower percentage of people who mark “now on active duty,” nor a higher percentage of people who mark "on active duty in the past, but not now," but the changes categories did not adversely affect either category, either.

5.6 Reliability of the Data

Do the changes to the veteran status question improve the reliability of the data?

Table 4 shows the gross difference rates, i.e. the percent of inconsistent answers between the original interview and the CFU, for the veteran status question by the control and test groups. The “Never Served” and “Training Only” response categories had significantly lower gross difference rates, and index values on the indices of inconsistency on the test version, demonstrating a lower percent of inconsistent answers on the test version compared to control. For all other categories, the differences in gross difference rates between the test and control version were not statistically significant.

Table 4. Veteran Status Question Gross Difference Rate (GDR)

Category	Test GDR	Standard Error	Control GDR	Standard Error	Test-Control (%)	Standard Error (%)	Significant
Unweighted cases	17,484		17,402				
Never Served on Active Duty	0.6	0.1	1.1	0.1	-0.5	0.1	Yes
Training Only	0.8	0.1	1.3	0.1	-0.4	0.1	Yes
On Active Duty Now	0.3	0.1	0.2	0.1	0.1	0.1	No
Active Duty in the Past, but Not Now	1.3	0.1	1.1	0.1	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 multiple comparison method at the $\alpha = 0.10$ level.

Table 5 shows the index of inconsistency values, the estimate of the magnitude of response variability for the veteran status question, by the control and test groups. The "Never Served" and "Training Only" response categories had significantly lower index values on the indices of inconsistency on the test version compared to the control, demonstrating less response variability. For all other categories, the differences in index of inconsistency values between the test and control version were not statistically significant. The L-fold index value, an overall index of inconsistency, was also significantly lower on the test version (7.4) than on the control (8.9).

Table 5. Veteran Status Question Index of Inconsistency Values

Category	Test Index Value	Standard Error	Control Index Value	Standard Error	Test-Control (%)	Standard Error (%)	Significant
Unweighted cases	17,484		17,402				
Never Served on Active Duty	3.1	0.3	5.5	0.6	-2.4	0.6	Yes
Training Only	26.3	2.5	49.5	3.9	-23.2	4.5	Yes
On Active Duty Now	26.5	5.5	23.4	5.8	3.1	8.0	No
Active Duty in the Past, but Not Now	7.7	0.7	6.1	0.7	1.5	1.0	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 multiple comparison method at the $\alpha = 0.10$ level.

A lower percent of inconsistent answers and less response variability on the test versions of the "Never Served" and "Training Only" response categories seem to demonstrate that the changes to the veteran status question did improve the reliability of the test data.

5.7 Service-connected Disability Rating Item Missing Data Rate

Do the changes to the veteran status question affect the item missing data rate for service-connected disability?

The item missing data rate for service-connected disability status showed a statistically significant decrease between the control version (3.1) and the test version (1.8), of 1.3 points. This decrease shows that the changes to the veteran status question seem to improve the reliability of other questions as well.

5.8 Multiple Answers

Do the changes to veteran status question reduce the occurrence of multiple answers?

Table 6 shows the percent of multiple responses to the veteran status question. It compares the percentage of respondents in the control version that provided multiple answers to the question. Though there few instances of multiple answers in the control,

the test version showed a statistically significant decrease in the percent of multiple answers compared to control.

Table 6. Percent of multiple answers

	Test (%)	Standard Error (%)	Control (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significant
Unweighted cases	34,027		34,029				
Percent of Multiple Answers	0.0	0.0	0.2	0.0	-0.1	0.0	Yes

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

* Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a one-sided test

5.9 Reliability of the Data by Mode

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

Tables 7, 8, and 9 show the response distribution for each veteran status category by the control and test groups for mail, CATI, and CAPI. Mail response showed a statistically significant decrease in the percent indicating "On Active Duty Now" on the test version. CAPI response showed a statistically significant increase in respondents indicating "Training Only" on the test version compared to control. The differences in CATI response distributions between the test and control version were not statistically significant.

Table 7. Veteran Status Question Response Distribution – Mail

Category	Test (%)	Standard Error (%)	Control (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significant
Never Served on Active Duty	87.3	0.3	86.7	0.2	0.6	0.3	No
Training Only	1.9	0.1	1.9	0.1	0.0	0.2	No
On Active Duty Now	0.4	0.1	0.6	0.1	-0.2	0.1	Yes
Active Duty in the Past, but Not Now	10.5	0.2	10.8	0.2	-0.4	0.3	No
Total	100.0		100.0		0.0		

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 multiple comparison method at the $\alpha = 0.10$ level.

Table 8. Veteran Status Question Response Distribution – CATI

Category	Test (%)	Standard Error (%)	Control (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significant
Never Served on Active Duty	88.1	0.6	89.0	0.6	-1.0	0.8	No
Training Only	1.6	0.2	1.0	0.2	0.6	0.3	No
On Active Duty Now	0.3	0.1	0.3	0.1	0.0	0.2	No
Active Duty in the Past, but Not Now	10.1	0.6	9.7	0.6	0.4	0.7	No
Total	100.0		100.0		0.0		

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 multiple comparison method at the $\alpha = 0.10$ level.

Table 9. Veteran Status Question Response Distribution – CAPI

Category	Test (%)	Standard Error (%)	Control (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significant
Never Served on Active Duty	92.7	0.4	92.1	0.4	0.6	0.6	No
Training Only	1.0	0.2	0.5	0.1	0.5	0.2	Yes
On Active Duty Now	0.7	0.1	0.7	0.2	0.0	0.2	No
Active Duty in the Past, but Not Now	5.6	0.4	6.7	0.4	-1.1	0.6	No
Total	100.0		100.0		0.0		

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 multiple comparison method at the $\alpha = 0.10$ level.

Tables 10, 11, and 12 show the gross difference rates by the control and test groups for mail, CATI, and CAPI. Mail response showed a statistically significant decrease in the gross difference rates for "Never Served" and "Training Only" for the test version compared to control. For all other categories, the differences in gross difference rates between the test and control version were not statistically significant.

Table 10. Veteran Status Question Gross Difference Rate (GDR)– Mail

Category	Test GDR	Standard Error	Control GDR	Standard Error	Test-Control (%)	Standard Error (%)	Significant
Never Served on Active Duty	0.7	0.1	1.3	0.1	-0.6	0.2	Yes
Training Only	0.9	0.1	1.6	0.1	-0.7	0.2	Yes
On Active Duty Now	0.2	0.0	0.3	0.1	-0.1	0.1	No
Active Duty in the Past, but Not Now	1.2	0.1	1.2	0.1	0.0	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 multiple comparison method at the $\alpha = 0.10$ level.

Table 11. Veteran Status Question Gross Difference Rate (GDR) – CATI

Category	Test GDR	Standard Error	Control GDR	Standard Error	Test-Control (%)	Standard Error (%)	Significant
Never Served on Active Duty	1.1	0.3	0.7	0.2	0.4	0.3	No
Training Only	1.2	0.3	1.1	0.3	0.1	0.4	No
On Active Duty Now	0.3	0.2	0.1	0.1	0.2	0.2	No
Active Duty in the Past, but Not Now	2.2	0.4	1.2	0.3	1.0	0.5	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 multiple comparison method at the $\alpha = 0.10$ level.

Table 12. Veteran Status Question Gross Difference Rate (GDR) – CAPI

Category	Test GDR	Standard Error	Control GDR	Standard Error	Test-Control (%)	Standard Error (%)	Significant
Never Served on Active Duty	0.5	0.1	1.1	0.3	-0.6	0.3	No
Training Only	0.6	0.2	0.7	0.2	-0.1	0.3	No
On Active Duty Now	0.5	0.2	0.2	0.1	0.3	0.2	No
Active Duty in the Past, but Not Now	1.2	0.2	0.9	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 multiple comparison method at the $\alpha = 0.10$ level.

Tables 13, 14, and 15 show the index of inconsistency values by the control and test groups for mail, CATI, and CAPI. Mail response showed a statistically significant decrease in the index of inconsistency values for "Never Served" and "Training Only" for the test version compared to control. CAPI response showed a statistically significant increase in respondents indicating "Training Only" on the test version. For all other categories, the differences in index of inconsistency values between the test and control version were not statistically significant.

Table 13. Veteran Status Question Index of Inconsistency Values – Mail

Category	Test Index Value	Standard Error	Control Index Value	Standard Error	Test-Control (%)	Standard Error (%)	Significant
Never Served on Active Duty	2.9	0.4	5.2	0.5	-2.3	0.6	Yes
Training Only	22.9	2.5	49.2	3.3	-26.3	3.9	Yes
On Active Duty Now	22.0	5.5	33.3	5.7	-11.3	9.2	No
Active Duty in the Past, but Not Now	6.4	0.6	5.7	0.6	0.7	0.9	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 multiple comparison method at the $\alpha = 0.10$ level.

Table 14. Veteran Status Question Index of Inconsistency Values – CATI

Category	Test Index Value	Standard Error	Control Index Value	Standard Error	Test-Control (%)	Standard Error (%)	Significant
Never Served on Active Duty	5.1	1.3	3.3	1.1	1.8	1.6	No
Training Only	47.0	8.6	56.7	13.0	-9.6	14.8	No
On Active Duty Now	38.4	24.6	18.5	17.9	19.9	29.7	No
Active Duty in the Past, but Not Now	11.3	2.0	6.2	1.5	5.1	2.5	No

Source: U.S. Census Bureau, 2010 American Community Survey Content Test, September to December 2010
multiple comparison method at the $\alpha = 0.10$ level.

Table 15. Veteran Status Question Index of Inconsistency Values – CAPI

Category	Test Index Value	Standard Error	Control Index Value	Standard Error	Test-Control (%)	Standard Error (%)	Significant
Never Served on Active Duty	3.0	1.0	7.2	1.9	-4.2	2.1	No
Training Only	29.8	7.9	48.7	14.5	-18.9	16.4	No
On Active Duty Now	28.6	8.2	15.6	9.0	12.9	12.7	No
Active Duty in the Past, but Not Now	9.6	2.1	7.5	2.1	2.1	2.9	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 multiple comparison method at the $\alpha = 0.10$ level.

Table 16 shows the item missing data rate for the veteran status question by the control and test groups for mail, CATI, and CAPI. CATI response showed a statistically significant increase in the item missing data rate on the test version compared to the control. For the other two modes, the differences in item missing data rates between the test version and the control version were not statistically significant.

Table 16. Veteran Status Item Missing Data Rates – Data Collection Mode

	Test (%)	Standard Error (%)	Control (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significant
Mail	7.7	0.2	7.8	0.2	-0.2	0.3	No
CATI	3.7	0.4	2.7	0.4	1.0	0.6	Yes
CAPI	3.9	0.5	3.6	0.4	0.2	0.6	No

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

* Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a one-sided test

The L-fold index value for mail response was significantly lower on the test version (6.37) than on the control (8.83). For the other two modes, the differences in L-fold index values between the test version and the control version were not statistically significant.

5.10 Reliability of the Data by Mail Response Stratum

For each mail response stratum, do the changes to the veteran status question affect the item missing data rate, response distribution, or reliability of the data?

Tables 17 and 18 show the response distribution for each veteran status category by the control and test groups for high and low mail response strata. The low response stratum showed a statistically significant increase in the percent indicating "Training Only" on the test version compared to control. For all other categories, the differences between the test and control version were not statistically significant.

Table 17. Veteran Status Question Response Distribution – High Mail Response Stratum

Category	Test (%)	Standard Error (%)	Control (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significant
Never Served on Active Duty	88.4	0.3	87.9	0.3	0.6	0.4	No
Training Only	1.7	0.1	1.5	0.1	0.2	0.2	No
On Active Duty Now	0.5	0.1	0.6	0.1	-0.2	0.1	No
Active Duty in the Past, but Not Now	9.4	0.3	10.0	0.3	-0.6	0.4	No
Total	100.0		100.0		0.0		

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 multiple comparison method at the $\alpha = 0.10$ level.

Table 18. Veteran Status Question Response Distribution – Low Mail Response Stratum

Category	Test (%)	Standard Error (%)	Control (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significant
Never Served on Active Duty	91.7	0.2	91.7	0.2	0.0	0.3	No
Training Only	1.2	0.1	0.9	0.1	0.3	0.1	Yes
On Active Duty Now	0.5	0.1	0.5	0.1	0.0	0.1	No
Active Duty in the Past, but Not Now	6.6	0.2	6.9	0.2	-0.3	0.3	No
Total	100.0		100.0				

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 multiple comparison method at the $\alpha = 0.10$ level.

Tables 19 and 20 show the gross difference rates by the control and test groups for high and low mail response strata. The high response stratum showed a statistically significant decrease in the gross difference rates for "Never Served" and "Training Only" on the test version compared to control. The low response stratum showed a statistically significant decrease in the gross difference rates for "Never Served" and "Active Duty in the Past" on the test version compared to control. For all other categories, the differences in gross difference rates between the test and control version were not statistically significant.

Table 19. Veteran Status Question Gross Difference Rate (GDR) – High Mail Response Stratum

Category	Test GDR	Standard Error	Control GDR	Standard Error	Test-Control (%)	Standard Error (%)	Significant
Never Served on Active Duty	0.6	0.1	1.2	0.2	-0.5	0.2	Yes
Training Only	0.8	0.1	1.3	0.1	-0.5	0.2	Yes
On Active Duty Now	0.3	0.1	0.3	0.1	0.0	0.1	No
Active Duty in the Past, but Not Now	1.3	0.2	1.1	0.2	0.1	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 multiple comparison method at the $\alpha = 0.10$ level.

Table 20. Veteran Status Question Gross Difference Rate (GDR)– Low Mail Response Stratum

Category	Test GDR	Standard Error	Control GDR	Standard Error	Test-Control (%)	Standard Error (%)	Significant
Never Served on Active Duty	0.7	0.1	1.0	0.1	-0.4	0.2	Yes
Training Only	0.9	0.1	1.1	0.1	-0.1	0.2	No
On Active Duty Now	0.4	0.1	0.2	0.1	0.2	0.1	No
Active Duty in the Past, but Not Now	1.3	0.1	0.9	0.1	0.4	0.2	Yes

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 multiple comparison method at the $\alpha = 0.10$ level.

Tables 21 and 22 show the index of inconsistency values by the control and test groups for high and low mail response strata. Overall, mail response shows a statistically significant decrease in the index of inconsistency values for "Never Served" and "Training Only" for the test version compared to control, which is reflected in each mail response stratum. For all other categories, the differences in index of inconsistency values between the test and control version were not statistically significant.

Table 21. Veteran Status Question Index of Inconsistency Values – High Mail Response Stratum

Category	Test Index Value	Standard Error	Control Index Value	Standard Error	Test-Control (%)	Standard Error (%)	Significant
Never Served on Active Duty	3.0	0.4	5.3	0.7	-2.3	0.8	Yes
Training Only	23.6	3.0	47.0	4.3	-23.4	5.0	Yes
On Active Duty Now	25.2	7.4	24.3	7.2	0.9	10.9	No
Active Duty in the Past, but Not Now	7.3	0.9	6.0	0.9	1.3	1.2	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 multiple comparison method at the $\alpha = 0.10$ level.

Table 22. Veteran Status Question Index of Inconsistency Values – Low Mail Response Stratum

Category	Test Index Value	Standard Error	Control Index Value	Standard Error	Test-Control (%)	Standard Error (%)	Significant
Never Served on Active Duty	3.7	0.5	6.5	0.8	-2.8	1.0	Yes
Training Only	38.1	3.8	62.2	5.0	-24.0	6.6	Yes
On Active Duty Now	29.4	6.1	20.3	5.9	9.1	8.1	No
Active Duty in the Past, but Not Now	9.0	1.0	6.6	0.9	2.3	1.3	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 multiple comparison method at the $\alpha = 0.10$ level.

Table 23 shows the item missing data rate for the veteran status question by the control and test groups for high and low mail response strata. The differences in item missing data rates between the test version and the control version were not statistically significant.

Table 23. Veteran Status Item Missing Data Rates – Mail Response Stratum

	Test (%)	Standard Error (%)	Control (%)	Standard Error (%)	Test-Control (%)	Standard Error (%)	Significant
High Mail Response Stratum	5.7	0.3	5.6	0.3	0.1	0.4	No
Low Mail Response Stratum	6.8	0.2	6.9	0.3	-0.1	0.3	No

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

* Statistical significance of differences is determined at the $\alpha = 0.10$ significance level using a one-sided test

The differences in L-fold index values between the test version and the control version for high and low mail response strata were not statistically significant.

5.11 Behavior Coding

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

Overall, the test series of questions performed significantly better on interviewer behavior than the control series by 19 percentage points. In particular, the control version of the initial veteran status question (see VET1C in Appendix C) had an especially low rate of interviewer standard behavior – 15% compared to 59% and 64% for the control version of the Reserves question (see MILC) and the control version of the active duty question (see ACTIVEC), respectively. The four items in the test version ranged from 52% to 82% standard behavior for interviewers. On the respondent side, the control series performed significantly better than the test series, though the magnitude of the difference was rather small – four percentage points. Among the four test items the test version of the initial veteran status question (see ACTIVET in Appendix C) had the lowest rate of standard behavior, at 53%, while the other three test items (see RESERVES, TRAINING, and ACTIVET) ranged from 62% to 73% standard behavior.

See Tables 24, 25 and 26 in Appendix A.

6. SUMMARY

The estimate of veterans from the test version is comparable to the estimate from the control version. There are fewer inconsistencies between the veteran status question and the period of service question. There were lower gross difference rates and index of inconsistency values for the “Never Served” and “Training Only” response categories, and no statistical differences in the estimate of veterans, the estimate of active duty respondents or the item missing data rates between the control and the test.

In the supplemental questions, there was a statistically significant decrease in the item missing data rates for service-connected disability, and the percent of multiple answers. Additionally, the test series of questions performed significantly better for both interviewer and respondent behavior than the control series.

Further, the VA recommends the adoption of the test version of the veteran status questions based on design, word count, and statistical performance of the test version. While the VA prefers simpler questions that collect only necessary information, they were also pleased to see no statistically significant difference in the gross count of veterans, and fewer cases of missing data or multiple responses to a single response question.

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 Ruiter, Jennifer Tancreto, Anthony G. Tersine JR., and Mary France Zelenak.

References

Clark S. and Raglin D. (2006). “Evaluation Report Covering Military Service Status,” 2006 American Community Survey Content Test Report P.5.a, 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 24. Veteran Status Question Behavior Coding Results: "Standard" Question-Asking (exact reading/slight change or correct verification) by Mode

Category	Test % Standard	Standard Error	Control % Standard	Standard Error	Test-Control (%)	Standard Error (%)	Significant
Overall	55.0	0.0	35.0	0.0	19.0	0.0	Yes
CAPI	45.0	0.0	31.0	0.0	14.0	0.0	Yes
CATI	65.0	0.0	39.0	0.0	25.0	0.0	Yes

Table 25. Veteran Status Question Behavior Coding Results: "Standard" Question-Asking (exact reading/slight change or correct verification) by Language Spoken

Category	Test % Standard	Standard Error	Control % Standard	Standard Error	Test-Control (%)	Standard Error (%)	Significant
Overall	55.0	0.0	35.0	0.0	19.0	0.0	Yes
English	70.0	0.0	43.0	0.0	27.0	0.0	Yes
Spanish	45.0	0.0	31.0	0.0	15.0	0.0	Yes

Table 26. Veteran Status Question Behavior Coding Results: "Standard" Question-Asking (exact reading/slight change or correct verification) by Question

Category	Test % Standard	Standard Error	Control % Standard	Standard Error	Test-Control (%)	Standard Error (%)	Significant
Overall	55.0	0.0	35.0	0.0	19.0	0.0	Yes
ACTIVEC			64.0	0.1			
ACTIVET	82.0	0.1					
MILC			59.0	0.2			
RESERVES	56.0	0.0					
TRAINING	70.0	0.1					
VET1C			15.0	0.0			
VET1T	52.0	0.0					

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

Figure B-1. Control Version of the Veteran Status Question

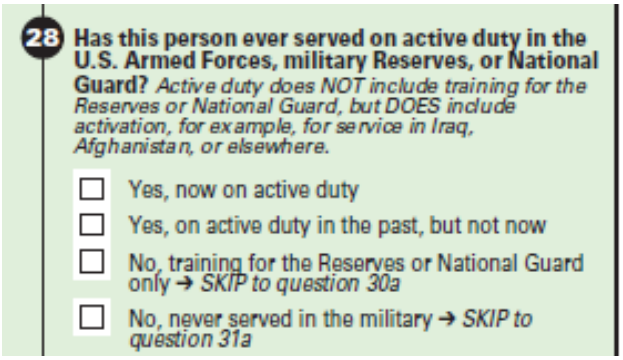
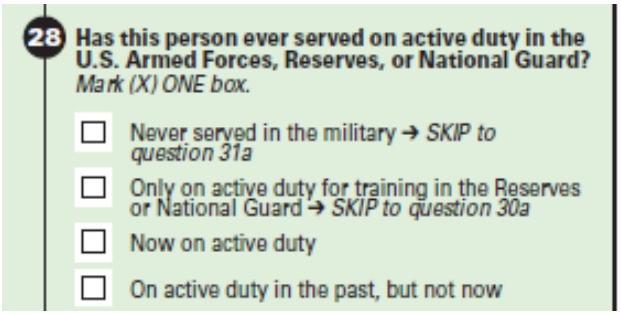


Figure B-2. Test Version of the Veteran Status Question



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

Control Questions:

? [F1]

[(Has <Name> / Have you)] ever served on active duty in the U.S. Armed Forces, military Reserves, or National Guard? Do not include training for the Reserves or National Guard but do include activation, for example, for service in Iraq, Afghanistan, or elsewhere.

❍ 1. Yes
❍ 2. No

Ever Served ☐

VET1C

MILB

When (were you/ was <Name>) on active duty?

❍ 1. Now on active duty
❍ 2. On active duty during the last 12 months, but not now
❍ 3. On active duty in the past, but not during the last 12 months

Active Duty

MILC

(Has <Name>/ Have you) ever been in the U.S. military Reserves or the National Guard?

❍ 1. Yes
❍ 2. No

Reserves Nat. Guard

Test Questions:

? [F1]

[Has <Name>/Have you] ever served on active duty in the U.S. Armed Forces, Reserves, or National Guard?

❍ 1. Yes
❍ 2. No

Ever Served ☐

VET1T

? [F1]

[Has <Name>/Have you] ever been in the Reserves or National Guard?

❍ 1. Yes
❍ 2. No

Reserves or Guard ☐

RESERVES

? [F1]

[Was <Name>/Were you] on active duty ONLY FOR TRAINING in the Reserves or National Guard?

❍ 1. Yes
❍ 2. No

Training Reserves or Guard ☐

TRAINING

? [F1]

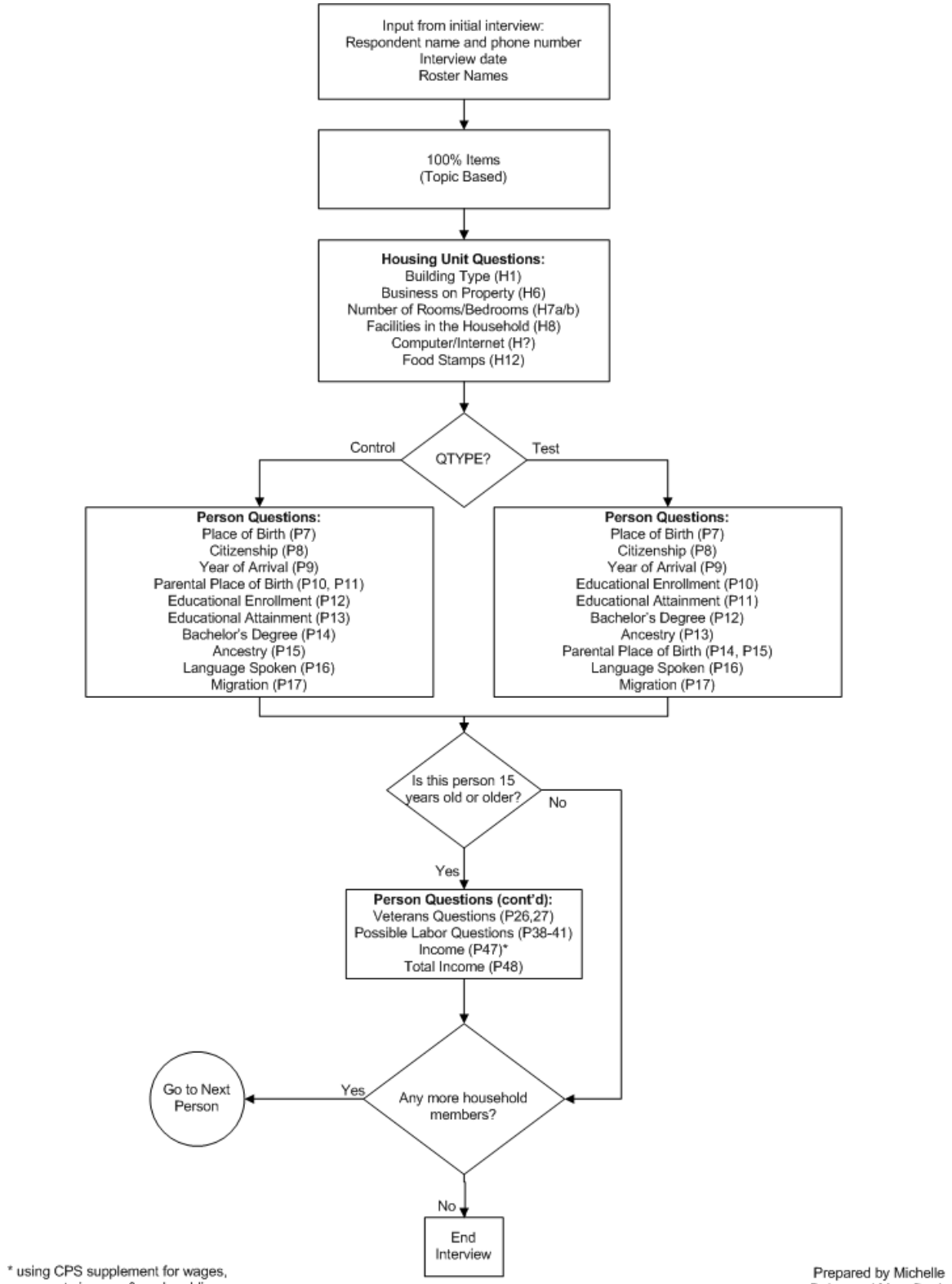
[Is <Name>/Are you] currently on active duty?

❍ 1. Yes
❍ 2. No

Active Duty Now ☐

ACTIVET

Appendix D: Flow of the Content Follow-Up



Appendix E: Information Page

Test Design

Treatments	Two question versions with different wording (see page 4).
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. *CATI and CAPI interviews will be recorded using Computer-Assisted Recorded Interviewing (CARI) technology.*
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 veteran status question result in a similar or higher estimate[1] of veterans?	Compare the estimate of veterans between the control and test versions.
2	Do the changes to the veteran status question decrease the number of respondents that classified themselves as "Yes, on active duty now" but marked a period of service that was not "September 2001 or later"?	Calculate the number of respondents in the control version that classified themselves as "Yes, on active duty now" but marked a period of service that was not "September 2001 or later" and compare with the number of respondents that did the same in the test to evaluate response inconsistencies related to the estimate of veterans. This comparison is based on the belief that respondents are more likely to mark the veteran status question incorrectly than the period of service question. Inconsistencies between these two variables in the control version may be an indicator that respondents are marking the first "yes" category of the veteran status question in error.

[1] We believe that respondents, who are veterans, are currently mistakenly marking the first answer choice "Yes, now on active duty" because of the word 'yes' instead of marking the second answer choice "Yes, on active duty in the past, but not now." The new version of the question does not include the words 'yes' or 'no' in any of the answer choices.

No.	Research Questions	Evaluation Measures
3	Do the changes to the veteran status question lower the item missing data rate?	Compare the item missing data rates between the control and test versions.
4	Do the changes to the veteran status question result in a lower percentage of people who mark "now on active duty," and consequently, a higher percentage of people who mark "on active duty in the past, but not now"?	Compare the response distributions between the control and test versions.
5	Do the changes to the veteran status 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.
6	Do the changes to the veteran status question affect the item missing data rate for service-connected disability?	Compare the item missing data rates for service-connected disability between the control and test versions of veteran status.
7	Do the changes to veteran status question reduce the occurrence of multiple answers?	Compare the number of multiple answers to the question between the control and test versions.
8	For each mode of data collection, do the changes to the veteran status question affect the item missing data rate, response distribution, or reliability of the data?	For each mode (mail, CATI, CAPI), compare the item missing data rates, response distributions, and reliability measures between the control and the test versions. *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).*
9	For each mail response stratum, do the changes to the veteran status question 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.
10	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
1, 2	The estimate of veterans from the test version is comparable to or higher than the estimate from the control version. Additionally, there are fewer inconsistencies between the veteran status question and the period of service question, thus requiring less editing.
3-5	The item missing data rates and reliability measures, along with seeing an increase in the number of respondents who mark that they are veterans and seeing a decrease in the number of respondents currently on active duty when comparing the control and test versions, will be considered together when determining which question version performs better.

Supplemental Information

Research Question(s)	Criteria
6-10	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
Q.28 Has this person ever served on active duty in the U.S. Armed Forces, military Reserves, or National Guard? Active duty does NOT include training for the Reserves or National Guard, but DOES include activation, for example, for service in Iraq, Afghanistan, or elsewhere. ☐ Yes, now on active duty ☐ Yes, on active duty in the past, but not now ☐ No, training for the Reserves or National Guard only → SKIP to question 30a ☐ No, never served in the military → SKIP to question 31a	**Q.28 Has this person ever served on active duty in the U.S. Armed Forces, Reserves, or National Guard?** Mark (X) ONE box. ☐ Never served in the military → SKIP to question 31a ☐ Only on active duty for training in the Reserves or National Guard → SKIP to question 30a ☐ Now on active duty ☐ On active duty in the past, but not now