STUDY SERIES
*(Survey Methodology #2013-22)*

**2005 Census Internet Prototype Applications:**
**Usability and Accessibility Testing**
Elizabeth D. Murphy[1]
Susan Ciochetto[1]
Kathleen Ashenfelter[1]
Lawrence Malakhoff
David Coon
Nika Smith[2]


[1] Former Employees of the Census Bureau
[2] UserWorks, Inc.

Center for Survey Measurement
Research and Methodology Directorate
U.S. Census Bureau
Washington, D.C. 20233

**2005 Census Internet Prototype Applications:**
**Usability and Accessibility Testing**

**U. S. Census Bureau**
**Center for Survey Measurement**
**Human Factors and Usability Research Group**

**Elizabeth D. Murphy, Susan Ciochetto, Kathleen Ashenfelter (former Census**
**Bureau employees)**
**Lawrence Malakhoff (CSM), David Coon (DMD),**
**and Nika Smith (UserWorks, Inc.)**

**Originally Revised: June 30, 2005**

**Revised for CSM Report Series: November 5, 2013**

**Executive Summary**

Usability and accessibility testing of the Internet form for the 2005 National Census Test (NCT) took place in May, 2005. The purpose of usability testing was to identify any aspects of the user-interface design that might detract from the success and satisfaction of Internet respondents during the 2005 NCT. The purpose of the automated accessibility testing was to identify any violations of the Federal accessibility regulations (Section 508 of the Rehabilitation Act as amended in 1998). The purpose of testing for accessibility with actual participants was to identify aspects of the user-interface tagging that might be technically accessible but still awkward or difficult for users with disabilities in the 2005 NCT.

Methods. To the extent possible, testing methods were consistent with the methods described in the baselined test plan (Murphy, et al., 2005). For usability testing, SRD recruited 22 participants, who were evenly divided between the topic-based (i.e., all question topics –age, sex, race, etc.- for each person of the household before moving on) and the person-based (i.e., all questions asked for each person in a household before moving on) forms, by a process roughly resembling alternating. Most of the accessibility testing was accomplished by using automated tools, but we did recruit an internal blind user to evaluate both applications. We also recruited a Census employee with low-vision to evaluate the topic-based form.

Results. Exceeding the goal of 80 percent completion, all usability test participants completed their forms. Data-entry accuracy was high. Data-entry efficiency ranged from 70 percent for the person-based version to 90.5 percent overall, when adjusted for outliers. Navigational effectiveness was high. However, 82 percent of topic-based participants committed field-changing errors; and 70 percent of person-based participants had some difficulty moving between screens. Ninety-one percent of participants completed the form efficiently.

Upon comparing perceived completion time with actual completion time, we found that 55 percent of participants perceived their completion time to be less than their actual time; and 45 percent perceived their completion time to be longer than actual.

Subjective satisfaction was extremely high, reaching a mean of 7.9 on a 9-point scale, where 9 is highly positive and 1 is highly negative. The range of reported values was 3-9, and there was some scattered dissatisfaction with getting started, navigating back through the form, making changes, and accessing help.

Accessibility testing with two Census Bureau employees confirmed the Section 508 issues identified by the automated accessibility testing. Although both users with visual disabilities completed their forms, they often faced usability obstacles in the process.

Recommendations and their Disposition. The team discussed all findings and recommendations. Several usability recommendations were accepted and implemented.

Most were deferred for further research.  All but one of the Section 508 issues had been resolved by June 30, 2005; and the last one was nearing resolution.

**Table of Contents**

# List of Figures

**List of Tables**

# 1 Introduction

The 2005 National Census Test (NCT) is the latest in a series of tests that the U. S. Census Bureau has been conducting in preparation for the 2010 Census. The 2005 Census Internet form is one of the response modes to be tested in the 2005 NCT. In addition to conducting the usual software testing, the Census Bureau has tested the usability and accessibility of the Internet form's user interface in preparation for launching the site in August, 2005.

Usability is the extent to which a software user interface supports the accurate, efficient, and satisfying performance of the user's or respondent's tasks. Accessibility is a special case of usability: How readily usable is the user interface to persons with disabilities? The requirements for accessibility are specified by Federal regulation (29 U. S. Code 794d)[1], Section 508 of the Rehabilitation Act of 1973, as amended in 1998 and updated for electronic and information technology in 2002. Paragraphs (a) through (p) of Section 508 cover the accommodations that must be made so that people with visual, auditory, cognitive, and psychomotor disabilities are afforded access to information technology comparable to the access enjoyed by people without disabilities.

## 1.1 Purpose

The primary purpose of usability testing is to identify aspects of the user-interface design that lead to less than effective, efficient, and satisfying completion of the Internet form. We have evaluated the site against both qualitative and quantitative goals for usability and accessibility, as detailed in Appendix A. Based on test findings and trade-offs, several changes recommended and accepted by the 2005 National Census Test Internet Implementation team are being implemented in the software.

The purpose of accessibility testing is to evaluate the applications' compliance with the governing regulations as defined in Section 508 of the Rehabilitation Act. The Section 508 regulations essentially constitute the formal accessibility goals for the site. The team's intention is to correct any violations of Section 508 prior to launching the 2005 National Census Test Internet site. A further purpose of accessibility testing is to determine whether the user interface supports persons with disabilities in their efforts to provide their Census information.

Automated usability testing has identified violations of Section 508 (Coon, 2005; Malakhoff, 2005). Those findings are attached here as appendices for the purpose of documenting all the findings in one place. Testing with actual participants has evaluated whether the site is usable, as well as technically accessible, by persons with visual disabilities. This report presents and discusses those findings.

---

[1] The details of the Federal regulations can be examined at http://www.section508.gov

Identification of a particular usability or accessibility issue does not guarantee that it will be addressed through redesign. Issues identified through testing have been prioritized, and redesign is focusing on those issues that make it difficult for a respondent to complete the online form.

The purpose of the present document is to report on the methods for conducting the usability and participant-based accessibility testing of the 2005 Census Internet form as well as the results of the testing. This report closely follows the structure of the baselined test plan, which was developed by consensus of the Internet usability testing sub-team of the 2005 NCT Internet Implementation team.

## 1.2 Scope

Usability and accessibility testing with actual participants focused on the following respondent activities:
- Gaining access to the form
- Completing the form
- Responding to edit messages
- Reviewing entered data and making changes

Test participants used their own household information as the basis for completing the topic-based (i.e., all question topics -age, sex, race, etc.- for each person of the household before moving on) and person-based (i.e., all questions asked for each person in a household before moving on). A particular participant completed only the topic-based OR the person-based form. Participants were not asked to enter data for hypothetical households; but they were asked to make some hypothetical changes in their own data.

In this context, accessibility refers to compliance with Section 508 regulations. We recognize that complying with the Section 508 accessibility regulations does not guarantee a usable interface for people with disabilities (cf. Theofanos & Redish, 2003). Although automated accessibility testing focused on verifying compliance with Section 508, the testing also included some participants with visual disabilities who helped the usability-testing team evaluate the usability of the applications for people with similar disabilities. We did not evaluate usability for people with manual disabilities because of the tight and compressed testing schedules.

It was NOT within the scope of this testing to evaluate the relative efficiency, effectiveness, or user satisfaction with the topic-based approach versus the person-based approach. Doing so is one of the objectives of the 2005 National Census Test. Since usability and accessibility testing are non-experimental methods, it is not possible to address the topic-based versus person-based questions with the results of these tests. Although we present some results grouped by form versions (i.e., person-based and topic-based), we do not mean to imply a comparison based on these results because there was no random assignment of participants to form versions. Assignment was by convenience.

*1.3  Assumptions*

A major assumption was that the site would be available for usability testing on the days and times scheduled.  We assumed that no development or code updating would occur at the same time as a scheduled test session, unless there was a critical software error that would prevent further testing.  We assumed that the project manager would ensure conflict-free test sessions.

On the accessibility side, we assumed that the diagnostic software would be working and that we could purchase (or download) copies of the software commonly used by visually impaired participants to customize their user interfaces.

**2  Method**

This section details the non-experimental testing methods used in evaluating the 2005 Census Internet prototype applications for usability and accessibility.  Each kind of testing required test participants, facilities and equipment, materials, procedures, and logging of the videotapes.

*2.1  Methods of Usability Testing*

**2.1.1  Participants**

For usability testing, SRD recruited participants who met the following criteria:

- Related and non-related households for both the topic and person-based designs
- At least one year of self-reported experience with the Internet
- Households of size 1 for both applications
- Large households
- Households including an infant under the age of one

SRD recruited 11 participants for the evaluation of each of the two applications, for a total of 22 sighted and otherwise able participants.  When participants were being recruited, the recruiter informed them that they would be asked to provide authentic information about their households.  We did not recruit anyone who was uncomfortable with providing their own household data. Participants from outside the Census Bureau each received $35.00 to help with their expenses.

**2.1.2  Facilities and Equipment**

Usability testing took place in the Census Bureau's Usability Laboratory.  Located in Federal Office Building 4 at the Suitland Federal Center (Room 3223), the Usability Lab includes three testing rooms and three observation and recording consoles.  These

consoles are located in an observation room adjacent to the testing rooms. One-way glass allowed test administrators and observers to view the test participants. The test participant sat in one of the three testing rooms, facing a one-way glass and a wall-mounted camera, under a ceiling-mounted camera, and in front of a Liquid Crystal Display (LCD) monitor placed on a table at standard desktop height. Two microphones picked up sound in each testing room. The test administrator sat next to the participant throughout each session.

The lab has microphones and headsets for the test personnel as well as logging equipment and software, video cameras, scan converters, and digital recording equipment. The observation room has monitors at each console, which allowed observers a clear view of the participant's screen.

### 2.1.3 Materials

Except for the mailing packages, all testing materials are provided in the appendices.

### 2.1.3.1 Consent Form

Participants signed a standard form giving the Census Bureau permission to make a video recording and report on their comments and behavior (Appendix B). The consent form also informed participants that video clips from their tapes might be used in internal briefings to Census Bureau personnel without disclosing any of their personal information.

### 2.1.3.2 Questionnaire on Computer Use and Internet Experience

Prior to beginning the usability test, participants were given a short questionnaire to assess their experience with computers and the Internet (Appendix C). This questionnaire goes into more detail than the general recruiting screener, which asked how many years of experience they had had with the Internet.

### 2.1.3.3 Introduction to a Testing Session

A copy of the introduction that was given to test participants is attached to this report (Appendix D). A major purpose of the introduction was to assure test participants that they were assisting in evaluating the software, not having their own abilities tested. During the introduction, the participants had several opportunities to ask questions.

### 2.1.3.4 Mailing Package

The mailing package simulated the actual mailing package that households were to receive for the 2005 National Census Test. Test participants were asked to find the URL for the 2005 National Census Test, and they were observed as they looked for their Census Identification Number (CIN). David Coon and Myron Smith prepared 10 sets of the Control Form mailing packages. Five of the packages had a topic-based CIN, and

five of the packages had a person-based CIN. The paper questionnaires contained the test URL, not the production address. To allow reuse by more than one participant, the mailing packages were not sealed.

### 2.1.3.5 Satisfaction Questionnaire

A tailored version of the Questionnaire for User-Interaction Satisfaction (QUIS) appears in Appendix E. The QUIS items were adopted by consensus of the 2005 National Census Test Usability Sub-Team. The original QUIS was developed and validated at the University of Maryland (Norman, Shneiderman, Harper, and Slaughter, 1998). The Usability Lab uses the QUIS by permission of Dr. Kent Norman, University of Maryland.

### 2.1.3.6 Debriefing Questions

The kinds of questions to be asked during the retrospective portion of the test session appear in Appendix F. These questions were formulated, revised, and adopted by consensus of the 2005 National Census Test Usability Sub-Team (hereafter referred to as "the sub-team"). The test administrator was not limited to these questions but could ask others, depending on the specific situation for each test participant and the content of the think-aloud verbalizations.

### 2.1.4 Procedure

Prior to the start of a session, we determined the assignment of a participant to either the topic-based or person-based form. To control for as many extraneous situational variables as possible, we generally alternated the assignment of the topic-based and person-based applications to participants, starting with a coin toss to determine which pairing of participant and application to begin with. If we had done one week of testing with one application followed by one week of testing with the second application, situations that could not have been anticipated might have occurred in either week and might have affected just one group of test participants. Although usability testing is not controlled experimentation in a technical sense, we generally try to keep testing conditions about the same for all test participants.

Prior to conducting the actual test sessions, we held two dry runs with internal staff acting as test participants, one for the person-based prototype and one for the topic-based prototype.

Sequence of Activities. The test administrator greeted the test participant and gave a general introduction to the session. The test participant was asked to sign a consent form agreeing to have the session videotaped. In a short practice in thinking aloud, most of the test participants were asked to provide a running narrative while finding the local weather forecast on the Internet. After answering any questions from the participant, the test administrator handed the participant a Census mailing package containing the internal URL for the test site as well as a Census Identification Number (CIN).

The following sequence outlines the steps in the main test procedure:

1.  Participant successfully navigated the authentication process and went through the application, reporting his/her own data in a think-aloud manner.  Participants were instructed to stop before submitting their data.
    We asked participants to use true information about household members and, if necessary, assured them that the data would be encrypted, stored in a secure database, and protected by law (Title 13).
2.  When the participant reached the Review/Summary screen, the Test Administrator asked the participant how long he/she thought it took to provide the data.
3.  Participant took the QUIS.  Some talked through what they were thinking.
4.  Using QUIS ratings as a jumping-off point,  the Test Administrator queried the participant about any low values and segued into the debriefing questions, which are provided in Appendix F.
5.  The Test Administrator took the participant back to the Review/Submit screen and asked the participant whether he or she would like to make any changes in the answers after reviewing them.
    a.  If the participant identified something he or she wants to change, the Test Administrator asked how he or she would do that and to go ahead and make the change.
    b.  If the participant did not identify anything to change, the Test Administrator asked the participant what he/she would do when he/she reached this screen and found an error.  The Test Administrator asked the participant to make a change to his/her data using one of several pre-formulated changes.  Change scenarios appear in Appendix G.
6.  The Test Administrator asked the participant to complete the session, i.e., to submit the data.   If the participant could not figure out how to get back to the Review/Submit screen, the Test Administrator provided assistance.  Once on the Review/Submit screen, the participant was asked what he or she would do now to submit the data.
7.  The Test Administrator asked the participant a question about the "Thank You" screen, i.e.,   "What does this screen mean to you?"
8.  The Test Administrator asked if there are any points the participant would like to make that have not already been covered  [last question in the debriefing].
9.  The Test Administrator thanked the participant and had them sign the voucher indicating receipt of cash compensation ($35.00 for a non-Federal employee).

During test sessions, the Test Administrator's observations were guided by, but not limited to, the suggested areas for observation developed by the sub-team (Appendix H).

**2.1.5  Session Logging**

Although we had originally planned to use interns from the Usability Lab to log the session tapes, the two available interns had other commitments that prevented them from making rapid progress on the logging.  After discussion with 2005 NCT project management, SRD contracted with UserWorks, Inc. to perform the logging of the session

tapes using the Observer 5.0 logging software, developed by Noldus (2003).  The session logs document the occurrences of participant behaviors and permit calculation of time segments within each session.   Each session log was automatically synchronized with the corresponding videotape so that it was possible to locate particular incidents on the tapes by going to the logged time code.

Logging allowed the coding of participant behaviors, such as comments and instances of confusion, and supported the analysis of these behaviors for frequency of occurrence.  The set of behavior codes agreed upon by the sub-team appears in Appendix I.  Findings from behavior coding appear in the results section.  Highlight clips from the videotapes appear in the internal version of this report, which is authorized for examination by sworn Census Bureau employees.  To protect the security of Title 13 data, the version available online through the SRD Report Series does not contain the highlight tapes.

## 2.1.6  Browser Testing

David Coon (DMD) tested the applications with Internet Explorer (IE) 6.0.  Although he had intended to test with Firefox, he was not permitted to run Firefox on his computer for security reasons.

Using Internet Explorer, David Coon tested to ensure that the 2005 NCT Internet applications behaved in the desired way when going through the CIN/HUID entry page.  With IE he tested all paths of person-based questionnaires with each browser.  In IE, he tested different screen resolutions for both topic- and person-based applications.  He tested the following resolutions as expressed in picture elements (pixels):
- 600 x 800 pixels
- 1024 x 768 pixels
- 1152 x 864 pixels
- 1280 x 1024 pixels

David Coon reported that both applications seemed to work properly at all of these resolutions.

He also tested screen resolutions of 640 x 480 and 720 x 576.  Although horizontal scrolling was required at 640 x 480 and 720 x 576, the applications were designed to work without horizontal scrolling at a minimum of 600 x 800.  More than 95 percent of Web users have screen resolutions of at least 600 x 800 according to w3schools.com .

Browser testing followed sample paths of both the topic-based and person-based questionnaires with each browser.  The applications were tested with different font sizes, default fonts, and Javascript (enabled/disabled) for visual clarity, readability, and navigability.

## 2.2 Methods of Accessibility Testing

### 2.2.1 Visually Impaired Participants

In addition to the 22 sighted participants, we recruited two visually impaired participants to help us evaluate software accessibility and usability for persons with visual disabilities.[2] The blind participant was a regular, expert user of screen-reading software; the other visually impaired participant was a low-vision user who typically enlarges his font sizes and uses yellow text on a black background. The blind participant reported for a two-person household made up of unmarried partners. The low-vision participant reported for a four-person household including two adopted children. Both of these participants were Census employees and, therefore, did not receive any special compensation for their participation.

### 2.2.2 Facilities and Equipment

Automated accessibility testing took place in the Census Bureau's Accessibility Laboratory (Room 3222A-4). Testing tools available in the Accessibility Lab included the most popular screen reader, Job Access With Speech (JAWS 5.1, Freedom Scientific, 2004) and automated evaluation software (Insight/Infocus 4.2.2, SSB Technologies, 2004).

Accessibility testing with the blind participant took place in the Usability Lab and was videotaped as usual. Although we attempted to customize the lab's computer for the low-vision user, we were not able to duplicate the settings he normally uses in his office. He asked us to conduct the testing in his office so that he would be able to use his normal settings for font size and foreground/background colors. To accommodate his preference, we set up portable videotaping equipment in his office. We later copied the tape of this session to a tape that could be logged in the Usability Lab.

Automated testing assessed compliance with Section 508, and testing with participants assessed the usability of the two applications for people with visual disabilities. The automated testing was conducted separately by Larry Malakhoff and David Coon. Their reports are attached to this report for the sake of having all the findings in one place.

### 2.2.3 Materials

Materials for accessibility testing included the same content as those used in usability testing. The Test Administrator read the URL and the Census Identification Numbers to these participants. She also read aloud the consent form and the paper-and-pencil test

---

[2] We did not test with people who had disabilities in manual dexterity or cognition. SRD recommends recruiting people with a wide range of disabilities in future rounds of accessibility testing on Census Internet prototypes. The 2010 Census Internet form needs to be widely usable and accessible.

questionnaires to the visually impaired participants. She recorded their questionnaire responses on the paper forms.

## 2.2.4 Procedure

In testing the applications for accessibility, the test administrators looked for "showstoppers," i.e., anything that made it impossible for the participant to use either application. The test administrator's observations were guided, but not constrained by, the suggested areas of observation (Appendix H).

### 2.2.4.1 Testing with Software Tools

2.2.4.1.1 Diagnosis of Section-508 Compliance

The automated Insight/InFocus testing evaluated the prototype Internet applications for compliance with the Section 508 regulations (paragraphs (a) through (p)). For example, this testing identified missing label elements for radio buttons, checkboxes, and entry fields. The testing software looked for <noscript> tags wherever it found JavaScript. In addition to identifying any instances where color was used by itself to convey information, Insight/InFocus tested color combinations (foreground on background) for possible problems for colorblind/color deficient users. It looked for any use of embedded tables for page layout.

Some of the Insight/InFocus findings needed to be checked manually to verify the following conditions:
- HTML tables used in both applications contained proper markup and complied with Section 508.
- HTML markup for all images had appropriate ALT or Longdesc tags.
- Disabled navigation tabs contained appropriate ALT text indicating that the tabs were disabled.

As recommended by the Census Bureau's IT Standard 20 (2002), a manual verification was performed with Javascript turned on and off (enabled/disabled) to ensure that each application could be navigated both with and without Javascript enabled. See Appendix J for the findings and recommendations from tool-based accessibility testing.

2.2.4.1.2 Screen Reader

In testing the applications with a screen reader, Larry Malakhoff listened to each version using the principal screen-reading software used by blind people, JAWS, version 5.1 at the default settings. He tabbed to each field and listened for the label, then pressed <Enter> and entered some data. He verified that the forward and backward tab order matched the visual presentation. The interaction between the screen reader and form controls was evaluated for accuracy, and inconsistencies were noted.

2.2.4.1.3  Tabbing Order

The tabbing order (using TAB and SHIFT-TAB) was evaluated for logical forward and backward tabbing as described in the Census Bureau's IT Standard 20 (2002).  The logical order is left to right, then top to bottom on forward TAB and the same order in reverse for SHIFT-TAB).  Although tabbing order should have been under the control of the default browser since the contractor had not hard-coded tab stops, it was possible for the tabbing not to be in a logical order, depending on how the HTML was written.

**2.2.4.2  Testing with Visually Impaired Participants**

Procedures for the visually impaired test participants were similar to those used for sighted participants, but they were tailored for the needs of those with visual disabilities. As mentioned above, we covered some steps verbally.  For example, we asked the blind participant where, in the mailing package, she would expect to find the URL for the site and where she would expect to find her CIN; but the Test Administrator supplied this information to her.  We expected that any blind participant would use a screen reader, such as JAWS, but that other visually impaired participants would not be experienced in using JAWS.  If participants did not normally use a screen reader, they were not expected to use one during a test session.

**2.2.5  Session Logging**

Logging was done from the videotapes for all sessions with visually impaired participants.

*2.3  Observers*

The Census Bureau's Usability Lab invited members of the 2005 NCT Internet Implementation team to observe usability and accessibility testing.  To limit distractions to test participants, however, we attempted to restrict observers in the Usability Lab's observation room to no more than three at a time.  Observers were asked not to talk among themselves during a test session, since their voices could be heard in the testing rooms.

We coordinated the observation schedule with Sarah Brady (DSCMO).  Anyone who was unable to observe a "live" session of usability testing was invited to view the videotapes. For example, when a scheduled participant failed to appear, two observers from the Population Division viewed the videotape of another session.

Packets of screen shots were available for observers as well as test administrators.

*2.4  Analysis Techniques*

The analysis of usability data provided the basis for deciding whether the overall goals and the screen-specific goals were met.  Analysis of the accessibility results determined:

1) whether the applications were in compliance with Section 508; and 2) whether the applications were usable for respondents with the kinds of visual disabilities represented by participants in accessibility testing. We evaluated the user interface against the qualitative usability and accessibility goals by examining the frequency of the various coded behaviors, e.g., errors in navigating between screens and number of data-entry fields left blank.

Usability issues were identified by observation of the interactions between the test participants and the 2005 Census Internet prototype applications (topic-based and person-based). Instances of confusion were examined for possible design-based causes. Participant comments contributed to an understanding of user-system interaction, but were not taken as the final authority because users are sometimes misled by their expectations based on experience with other applications. Analysis attempted to uncover the thought process underlying participant comments and expectations. The analysis focused on participants' accuracy and efficiency.

For logged behaviors, we calculated descriptive statistics (e.g., frequencies, mean number of occurrences). For example, means were generated for participants' actual completion times, which were compared to their perceived completion times. Descriptive statistics were also generated for the QUIS ratings. Since the means tend (by definition) to dilute the low ratings, we examined the raw QUIS ratings for instances of participant dissatisfaction. The Insight/InFocus software provided descriptive statistics on the accessibility violations it identified. Aspects of the user interface associated with low means (i.e., below the mid-point of the scale) were identified as candidates for design changes.

## 3  Reporting

At the end of the testing period, we provided a short report to the 2005 Census Internet Usability Test Team. We revised this report with the team decisions on the various issues. We then developed the present final report covering usability and accessibility testing.

### 3.1  Short Report

Within three business days after the end of testing, we provided a short report listing prioritized usability and accessibility issues identified in the testing. The revised short report adhered to the following scheme for prioritizing the issues:
- The highest priority was reserved for any issue that could be considered a "showstopper," i.e., unless this issue is resolved, respondents will not be able to complete the form and submit their data.
- Moderate issues add to the difficulty of completing the form or the time spent doing so.

- Low-priority issues cause some minor annoyance or frustration to the respondent, but they do not increase completion time or difficulty significantly.

The original short report included recommendations for improving usability and accessibility. The revised short report documented the team's decisions on each issue. Unfortunately, this short report is no longer available.

### 3.2 Final Report

This final report includes methods, materials, and findings from usability and accessibility testing. Another version of the final report to the team includes video clips illustrating various usability issues. The final version, without the Title-13 protected video clips, will be posted on the Internet as part of the SRD Report Series.

## 4 Usability Results and Discussion

### 4.1 Outcomes for Usability Goals

Goals for usability were stated in the test plan. Here we report on the extent to which participants were able to meet the stated goals given the form design we provided to them.

#### 4.1.1 Completion of the form

**Goal**: 80 percent of test participants will be able to complete the form.

**Result**: Once test participants had gained access to the form, everyone was able to complete it, including the blind and low-vision users.

#### 4.1.2 Accurate and efficient data entry

**Goal**: Eighty percent of respondents should understand how data is entered and how it can be corrected. They should be able to easily identify the data needed and type entries and select options as needed. They should be able to correctly enter data the first time and/or deal with corrections of errors, inconsistencies, and blank entries following an edit message.

**Results for Accurate Data Entry**. The rate of data-entry errors was quite low. The following table presents the number of data-entry errors for both the person-based and the topic-based participants (N = 21):

**Table 1. Instances of Data-Entry Errors for Participants in the Person-based Group and the Topic-based Group during Usability Testing of Prototype Applications for the 2005 National Census Test (N = 21)**

| Form Version | Error in URL Entry | Error in completing field* | Left field blank** |
|---|---|---|---|
| **Person based** (n = 10) | 3 instances involving 3 Ps | 2 instances involving 1 participant (P) | 5 instances involving 1 P |
| **Topic based** (n = 11) | 4 instances involving 4 Ps | 3 instances involving 2 participants (Ps) | 4 instances involving 4 Ps |

*For example, user typed age in without noticing that the field was pre-filled
**User tried to continue to next page without realizing he/she had not responded fully.

Given the large number of opportunities for data-entry error, we consider data-entry accuracy to be quite high, probably exceeding the goal of 80 percent accuracy.

**Results for Efficiency of Data Entry.** Seventy percent for person-based and 73 percent for topic-based, ranging to 90.5 percent overall (when adjusted for outliers).

We are estimating the efficiency (speed) of data entry on the basis of the actual completion times recorded by the participants, as shown in Table 2. If we take the mean time in each group as a benchmark, participants in the person-based group who were at or below the mean time per person of three minutes and 45 seconds can be considered reasonably efficient. Seven of those ten participants met the data-entry-efficiency goal, while those who were above the per-person mean, especially the one who took ten minutes and 38 seconds, did not meet the goal.

Similarly, in the topic-based group, using five minutes and 20 seconds as a benchmark, eight of the 11 participants met the goal for date-entry efficiency, while three were above the per-person mean.

**Discussion**. As in the person-based group only one topic-based participant's per-person mean looks significantly above the overall per-person mean. So, out of the 21 participants for whom we have data[3], only two clearly exceeded the mean time spent per person in the household (Ps 10 and 15). This represents a 90.5 percent success rate for overall data-entry efficiency.

Participant 10 can be considered a true outlier because he spent a lot of time talking about unrelated topics while completing his form. He was not as familiar with the Internet as we expect actual respondents to be and was often disoriented, thinking he was in the form when he was really in the browser. Participant 15 can also be considered an outlier. She was often confused and did not seem to be adequately familiar with the Internet.

---

[3] Through the test administrator's oversight, Participant 3's session was not videotaped. However, we observed her to be extremely efficient at building up speed on the person-based form. This is attributable to the consistency of the design across persons, which allowed the participant to develop accurate expectations.

**Table 2. Times to Complete the Person-based and Topic-based Forms**

| Person-based participant # | Number of people in household | Actual length of time to complete form | Mean time to complete per person |
|---|---|---|---|
| 1 | 1 | 04:50 | 04:50 |
| 5 | 5 | 07:15 | 01:27 |
| 7 | 2 | 08:07 | 04:03 |
| 8 | 3 | 07:45 | 02:35 |
| 13 | 2 | 06:27 | 03:13 |
| 14 | 3 | 08:14 | 02:44 |
| 15 | 2 | 21:16 | 10:38 |
| 16 | 6 | 13:40 | 02:16 |
| 20 | 4 | 11:32 | 02:53 |
| 22 | 6 | 16:39 | 02:46 |
| **Overall mean time per member of household** | | | **03:45** |
| **Topic-based Participant #** | **Number of people in household** | **Actual length of time to complete form** | **Mean time to complete per person** |
| 2 | 3 | 22:42 | 07:34 |
| 4 | 5 | 19:55 | 03:59 |
| 6 | 2 | 07:12 | 03:36 |
| 9 | 5 | 08:07 | 01:37 |
| 10 | 2 | 34:55 | 17:28 |
| 11 | 4 | 17:22 | 04:20 |
| 12 | 1 | 05:31 | 05:31 |
| 17 | 6 | 16:02 | 02:40 |
| 18 | 2 | 12:49 | 06:24 |
| 19 | 4 | 09:24 | 02:21 |
| 21 | 4 | 12:47 | 03:11 |
| **Overall mean time per member of household** | | | **05:20** |

### 4.1.3  Effective and efficient navigation

**Goal**:  Eighty percent of respondents to the person-based form should be able to quickly gain an understanding of how the survey is organized into a household page, person pages, and a review/submit page.  Eighty percent of respondents to the topic-based form should be able to quickly gain an understanding of how the survey is organized into pages by topic and a review/submit page.  They should also understand how the navigation works using buttons and tabs.  They should be able to move through the form in a direct, forward manner initially and later, if needed, by jumping directly to other pages.

**Results for Effective Navigation**.  Given the large number of opportunities for navigational errors and difficulties, we consider the rate of navigational effectiveness to be reasonably high for both the person-based and topic-based applications.

The following table provides an overview of navigational errors and difficulties experienced by participants:

**Table 3.  Navigational Ineffectiveness in Usability Testing of Prototype Applications of a Census Internet Form for the 2005 National Census Test  (N = 21)**

| Form Version | Error in field focus | Error in changing fields | Difficulty in navigating between screens |
|---|---|---|---|
| **Person based** (n = 10) | 1 instance | 7 instances involving 5 participants (Ps) | 8 instances involving 7 Ps |
| **Topic based** (n = 11) | 1 instance | 13 instances involving 9 participants | 6 instances involving 3 Ps |

**Discussion**.  Some error is simply random and unaccountable.  Some portion of the error may have been attributable to a few participants' relative inexperience with the Internet.  Although some participants rated themselves as high on Internet experience, their behavior sometimes indicated that their self ratings may have been inflated. This was the case for participants 10 and 15, as noted above.

Between the two of them, participants 10 and 15 were involved in five of the field-changing errors and one of the between-screen navigation problems represented in Table 3.  Although this suggests that lack of experience with the Internet may have contributed to many of the recorded errors and navigational difficulties, the involvement of nine of 11 participants in topic-based field-changing errors suggests that Internet inexperience was not the whole story.  When 82 percent of topic-based users commit field-changing errors and 70 percent of person-based users have difficulty moving between screens, it would be wise to look for ways to reduce or eliminate these kinds of errors and difficulties.

**Results for Navigational Efficiency**.  The completion times reported in Table 2 suggest relatively efficient navigation for all participants except 10 and 15.  By observation, Participant 3 was extremely efficient in moving through the person-based form and in making corrections.  If we include Participant 3 in the calculation, 91 percent of the participants can be described as having completed the form efficiently (20/22).

The data on perceived completion time versus actual completion time are mixed.  As shown in the next table, 45 percent (9) of 20 participants estimated that their completion time was longer than it really was, while 55 percent (11) of 20 participants estimated that their completion time was shorter than it actually was.  Participant 4's estimate is essentially identical to the actual time.  Thus, a majority reported an efficient experience; but we should not ignore the large minority who reported a relatively less efficient experience with the prototype Internet form.

**Table 4.  Participants' Perception of Completion Time Compared to Actual Completion Time (N = 21), in Usability Testing of the Prototype Applications for the 2005 National Census Test:  An Indicator of Navigational Efficiency**

| Person-based P# | Perceived Length of Time | Actual Length of Time | Difference (perceived -actual) |
|---|---|---|---|
| 1 | 10:00.0 | 04:50.6 | +05:09.4 |
| 5 | 10:00.0 | 07:15.0 | +02:45.0 |
| 7 | 10:00.0 | 08:07.1 | +01:52.9 |
| 8 | 03:00.0 | 07:45.4 | -04:45.4 |
| 13 | 10:00.0 | 06:27.5 | +03:32.5 |
| 14 | 07:00.0 | 08:14.3 | -01:14.3 |
| 15 | 20:00.0 | 21:16.9 | -01:16.9 |
| 16 | 20:00.0 | 13:40.4 | +06:19.6 |
| 20 | 10:00.0 | 11:32.8 | -01:32.8 |
| 22 | 15:00.0 | 16:39.9 | -01:39.9 |
| **Topic-based P#** | **Perceived Length of Time** | **Actual Length of Time** | **Difference (perceived - actual)** |
| 2 | 20:00.0 | 22:42.7 | -02:42.7 |
| 4 | 20:00.0 ➔ | 19:55.3 | +00:04.7 |
| 6 | 03:00.0 | 07:12.6 | -04:12.6 |
| 9 | 04:00.0 | 08:07.5 | -04:07.5 |
| 10 | 10:00.0 | 34:55.9 | -24:55.9 |
| 11 | 30:00.0 | 17:22.4 | +12:37.6 |
| 12 | 07:00.0 | 05:31.1 | +01:28.9 |
| 17 | 15:00.0 | 16:02.4 | -01:02.4 |
| 18 | 15:00.0 | 12:49.4 | +02:10.6 |
| 19 | 07:00.0 | 09:24.6 | -02:24.6 |
| 21 | 15:00.0 | 12:47.8 | +02:12.2 |

Note: Participant 4's estimate was judged equal to the actual.
The mean for time estimates greater than the actual time was 4:05 minutes (range = 1:28 to 12:37).[4]  The mean for time estimates lower than the actual time was 4:35 minutes (range = 1:02 to 25:55).

### 4.1.4  Subjective satisfaction

**Goal**:  Eighty percent of respondents should be satisfied with the process of filling out the survey; they should not be frustrated or confused.  They should also be satisfied with the interaction and the time required to fill out the survey.  When the respondent submits

---

[4] Participant 4's overestimate of 0:04 seconds was omitted from the analysis because it was judged not to be significantly different from the actual completion time.

the form, he or she should be confident that it was duly submitted and received by the Census. Finally, the respondent should come away with a good impression of the Census Bureau.

**Results for Subjective Satisfaction**. On the whole, subjective satisfaction was extraordinarily high, as measured by the QUIS.

As shown in Table 4, very few questions received ratings below the midpoint of the nine-point scale, where 1 was an extremely negative score and 9 was an extremely positive score. There was some dissatisfaction with getting started, navigating back through the questionnaire, making changes, and accessing help. We touch on each of these issues to some extent in the following discussion based on observations made during usability testing, including accessibility testing with participants.

**Table 5. Summary of Results for Subjective Satisfaction Based on Participant Ratings on the Questionnaire for User Interface Satisfaction (QUIS) as Modified for this Usability Test***

| QUIS Item | Mean | Range |
|---|---|---|
| 1. overall reaction | 7.59 | 6-9 |
| 2. getting started | 7.23 | 3-9 |
| 3. screen layout | 8.27 | 5-9 |
| 4. amount of info on screen | 7.82 | 6-9 |
| 5. arrangement of screen | 8.05 | 5-9 |
| 6. order of questions | 8.32 | 6-9 |
| 7. ease of going back | 7.65 | 4-9 (2 Ps = N/A) |
| 8. ease of making changes | 7.80 | 4-9 (2 Ps = N/A) |
| 9. use of terminology | 8.14 | 5-9 |
| 10. adequacy of instructions | 7.91 | 5-9 |
| 11. accessing help | 7.69 | 4-9 (9 Ps = N/A) |
| 12. ease of entering data | 8.45 | 7-9 |
| 13. progression through q'aire | 8.36 | 7-9 |
| 14. length of time to complete | 7.64 | 5-9 |
| Overall Mean Rating | 7.92 | |

Ps = Participants
*The exact content of the QUIS items is provided in Appendix E.

Our usual finding is that participants report relatively high levels of satisfaction if they are able to finish the test tasks, which all participants did in this case. The high satisfaction ratings do not reflect the amount of difficulty that we observed most participants having as they tried to get started. Similar disconnects between participant performance and satisfaction have been documented in the usability literature (e.g., Andre and Wickens, 1995).

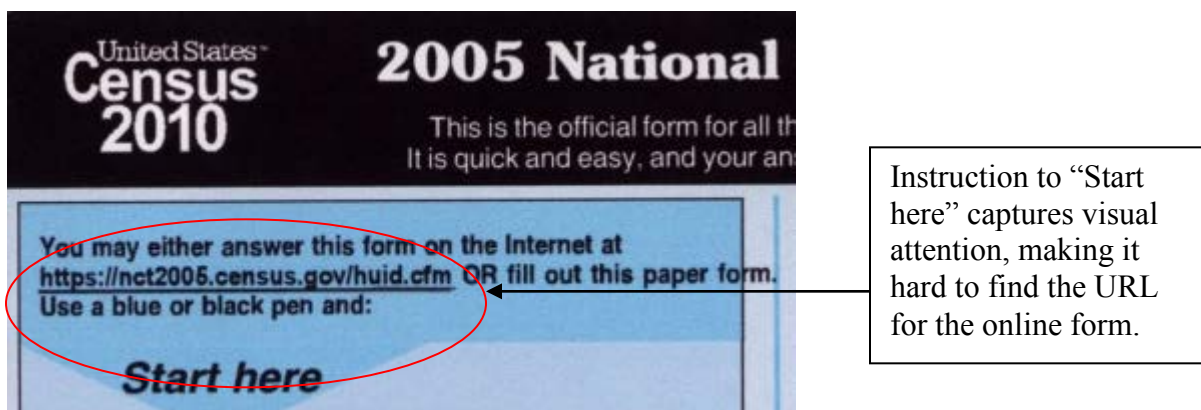## 4.2 Observations and Recommendations from Usability Testing

In the following sections we identify the various design features or items on which we are reporting. All items are facets or aspects of the prototype applications that we evaluated. Each item is followed by a discussion based on observations from testing and recommendations for its resolution. In some cases, we identify issues that cannot be resolved prior to launch but that can be kept in mind for future resolution.

We have prioritized our recommendations according to the following scheme:
- High = an issue that can be considered a "showstopper." That is, unless this issue is resolved, respondents will not be able to complete the form and submit the data.
- Moderate = an issue that is mildly confusing but the respondent can figure it out; it may add time to the overall difficulty of completing the form or the time spent doing so.
- Low = an issue that is minimally annoying to the respondent but one that does not increase completion time or difficulty significantly.

The documentation of each issue includes the team's decision about resolving the issue. In many cases, the team decided to recommend further research during the scheduled time for Round-2 usability testing.

## Item 1: Internet Address (URL)



Instruction to "Start here" captures visual attention, making it hard to find the URL for the online form.

**Figure 1: Location of URL for the 2005 Census Internet site on the front page of the paper**

As shown in Figure 1, the Uniform Resource Location (URL) for the online form is provided in the upper left corner on the first page of paper form. It is given in black text on a medium blue background.

*1a. Placement of URL*: One participant never did find the URL. He opened the paper form and immediately went to the area on the form that said "Start Here." He didn't consider anything above that. This may be because of the poor contrast between the

color of the type and the background color or because of the size of the font. A low-vision participant read "nct" as "net."

> Recommendation: The URL should be more obvious on the paper form. Since "Start Here" is meant to get the respondent started on filling out the form, the URL really shouldn't come after that instruction. However, it needs to be more prominent in the area above the instruction. Perhaps the font or background color could be changed, or the contrast could be enhanced. People tend to overlook text in a shaded area, thinking that it is not important.

>> Priority: High for 2010
>> Team Decision: Provide usability data to content R&D team for future research

The following table presents the time each participant spent trying to gain access to the form, i.e., the total time spent reviewing the paper form, locating the URL, typing in the URL, and locating the Census ID number:

**Table 6. Time Spent by Participants Attempting to Access the Prototype Internet Form for the 2005 National Census Test during Usability Testing (N = 21)**

| Participant# | Access Time (mins) |
|---|---|
| 1 | 02:00 |
| 2 | 07:15 |
| 4 | 03:05 |
| 5 | 01:03 |
| 6 | 04:11 |
| 7 | 04:53 |
| 8 | 03:59 |
| 9 | 01:59 |
| 10 | 10:12 |
| 11 | 03:02 |
| 12 | 03:12 |
| 13 | 01:51 |
| 14 | 02:37 |
| 15 | 06:34 |
| 16 | 06:56 |
| 17 | 03:30 |
| 18 | 03:19 |
| 19 | 04:10 |
| 20 | 02:27 |
| 21 | 03:09 |
| 22 | 02:25 |
| **Mean** | **03:54** |
| **Range** | **1:03 to 10:12** |

A concern is that potential respondents will give up if they cannot find their Census ID Numbers quickly and get into the form easily. Although resolving this issue is outside the scope of designing the Internet form, attention needs to be paid to the design of the respondent's end-to-end experience. If the respondent experience is cumbersome and confusing at the beginning, the take-up rate is likely to be affected. People will feel that their time is being wasted, and they will not have a good feeling about what they are getting themselves into. The process of accessing the form needs to be smooth, quick, and relatively simple, not awkward, time consuming, and complicated.

*1b. Complexity of URL*:
Participants missed the "s" in https:// and got an error screen. When one participant reviewed the address, she started with the nct…and when she found no error with that, she went back to the beginning of the URL and reviewed the entire address.

Several participants commented on the length and complexity of the URL. One started typing at the nct2005, completely omitting the https://. He got a screen that suggested alternatives. He clicked on one of them. When that didn't seem to be what he wanted, he typed in the entire URL in the search area of the browser. This gave him another error message. Finally, he said he would type in the URL on the address line. This was successful.

> Recommendation: If possible, the complexity of the URL should be reduced.

>> Priority: Moderate
>> Team Decision: The URL used for usability brought the participant directly past the firewall and to the secure servers. This added some complexity to the URL. For the 2005 NCT, the respondent will first go to an unsecured site – www.nct2005.census.gov -- and from there be directed to the first page behind the firewall. Therefore, this issue is resolved.

*1c. Accessing the site*: One participant saw the URL for the Census Bureau home page in the lower right corner of the letter and thought this was the address referred to in the letter. Some thought that they could get to the site off the Census Bureau's home page (www.census.gov).

> Recommendation: For 2010, it should be possible for respondents to get to the site from the Census home page.

>> Priority: High for 2010
>> Team Decision: The URL will most likely be a link from the Census home page for the 2010 census. Since the 2005 NCT will only impact a sample of the population, having the URL on the site may entice those who are not in sample to try to get into the site. Therefore, the URL will not be on the Census home page for the 2005 NCT.

*1d.  Placing the URL on the letter*:  A couple of participants thought the URL should be on the letter instead of just referring to it in the letter.

> Recommendation:   If the URL is placed on the letter, the letter should also inform the respondent that they would need the Census ID Number from the form so that the respondent does not discard the form and keep only the letter.  See issue in HUID section.

> Priority:   High for 2010
> Team Decision:  The URL and a note about the need for the respondent to keep the form will be part of the letter that is included in the 2005 NCT mailing package.

## Item 2:  HUID screen
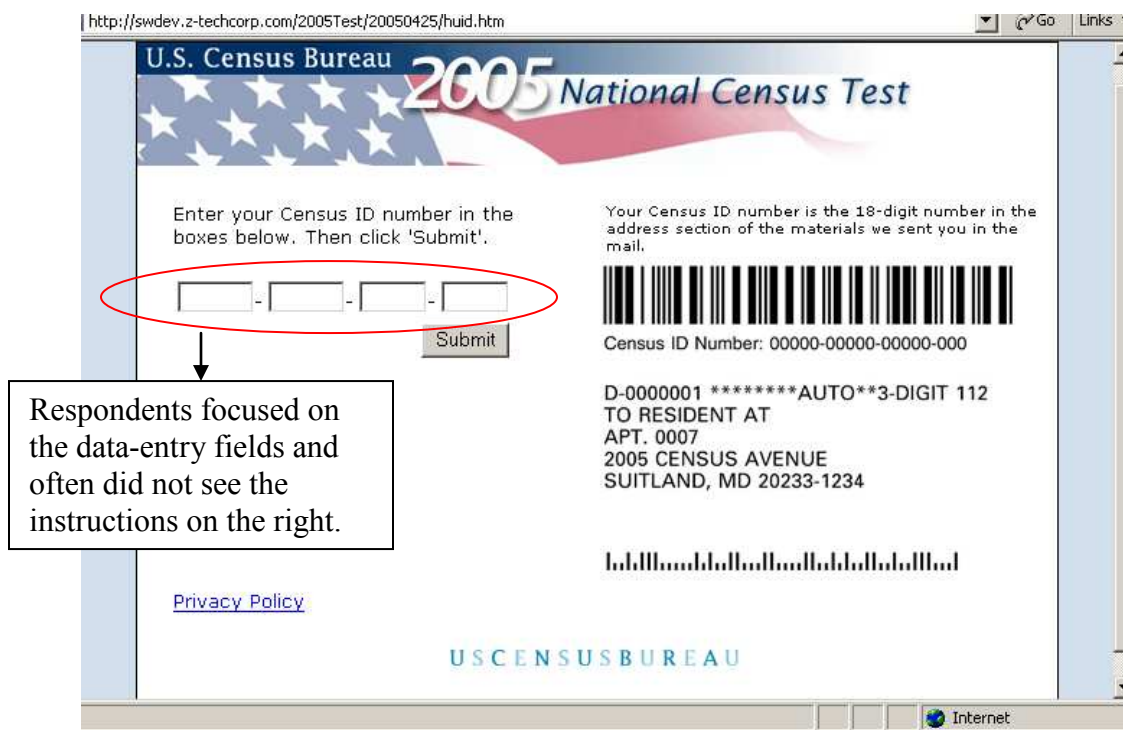


**Figure 2.  Housing Unit Identification (HUID) screen:  Respondent's task is to find their Census ID Number and enter it into the data-entry fields at the left of the screen; graphic on the right is intended to help respondents find their Census ID Number on the paper form mailed to them.**

*2a.  Finding the Census ID Number*.  It was obvious to all participants that the first thing they needed to do to continue was to enter a Census ID. However, many struggled with what it was and where to get it.  A couple thought it was their Social Security Number (SSN), but they noticed that there were four boxes for numbers and SSN only needs three.  Some people started leafing through the materials, picked up the form and glanced

over it.  Many never did see the instruction on the HUID screen about where to find the number.

As shown in Figure 3, the Census ID number was located on the back of the paper form, below a bar code.  When the instruction on the screen was pointed out to them, they said the font was too small, it looked like fine print that they didn't need, or the right side of the screen looked like a graphic so they ignored it.  One participant said she doesn't read things on the right side.  People do not expect to find critical information on the right side of the screen.  Finding the Census ID Number actually requires the respondent to turn the paper form over and upside down.
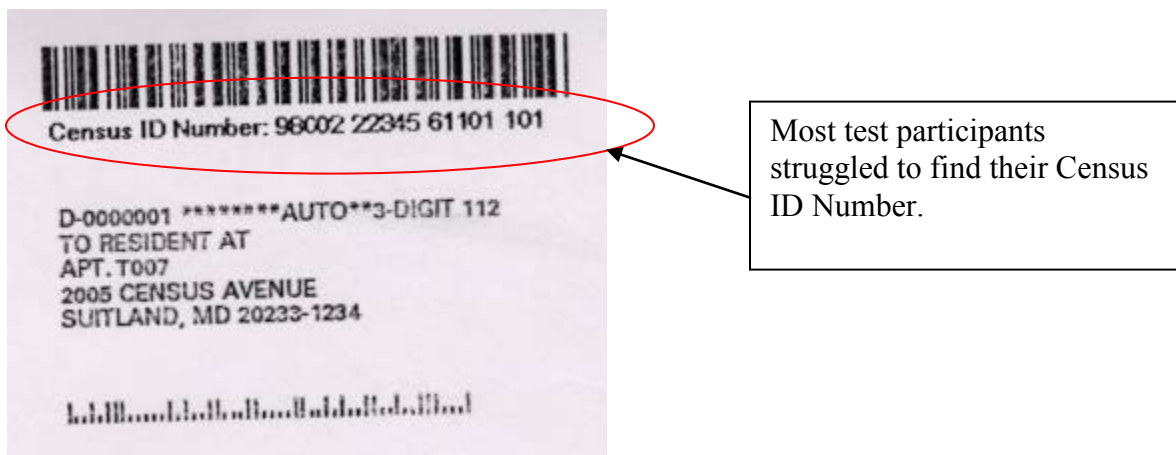


**Figure 3.  Sample Census ID Number, located on the back of the paper form underneath a bar code, above the mailing address (Image scanned from paper form).**

Recommendations:

1) Put a note on the letter in the mailing package informing the respondent that the paper form contains the ID number that they will need when filling out the form on the Internet – especially if the URL is placed on the letter.  This would increase the likelihood that the respondent doesn't throw away the form before completing the census.

> Priority = high for 2010
> Team Decision: This will be done for the 2005 test.

2) Slightly redesign the look of the screen.  Move the instruction so that it is underneath the data-entry field and put it larger type.  Bold the words Census ID number.  Line this text up with the graphic so that a red arrow can point to the ID number.   If possible, highlight the word Census ID Number in the graphic.  Shrink the address portion of the graphic so it is not as overpowering.

Priority = high for 2005 because some potential respondents will quit when they have trouble finding the HUID.

Team Decision: The team agreed to change the HUID screen as follows: Move the instruction so that it is underneath the data-entry field, and put it in the same size type as the instruction to enter the Census ID number. Move the graphic down so that a red arrow can point from the instruction to the graphic. We noted that the look of this might differ with various browser/OS combinations.

The next figure illustrates the revised version of this screen, which will be used in the production version of the site.
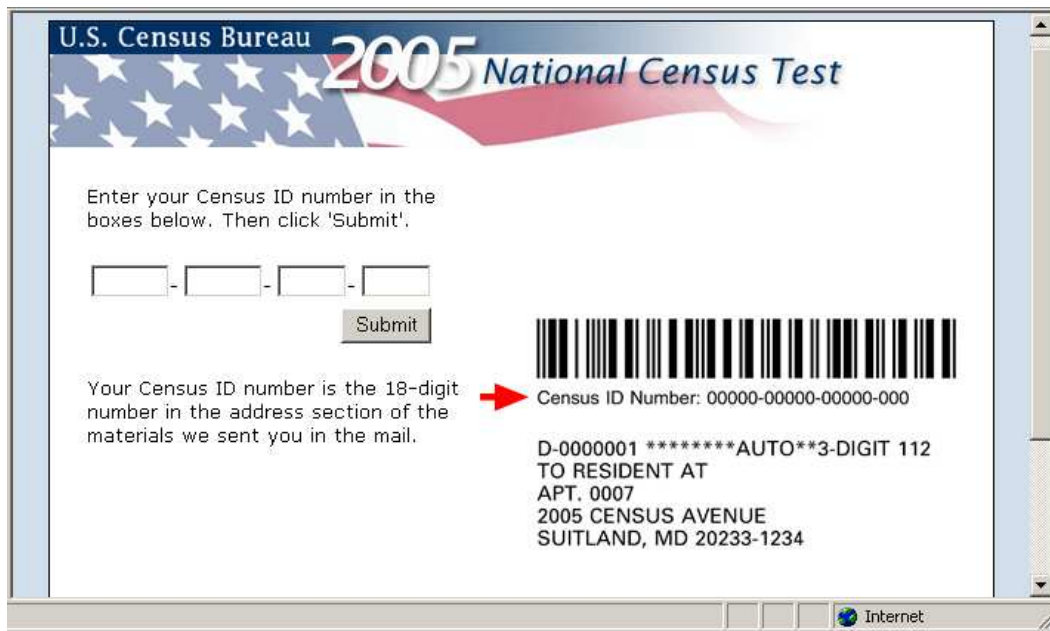


**Figure 4. HUID screen as revised after team discussion of usability recommendations**

An alternative solution was suggested by a participant: Reverse the order of the instructions and the data-entry fields, i.e., since we read from left to right, tell people where to find their HUID before asking them to enter it.

Priority: Moderate for 2005

Team Decision: We can experiment with putting the graphic on the left versus right in future usability tests.

*2b. Clarity of the Census ID on the Paper Form.* Once they found the number, all but one participant thought it was easy to enter. This one participant said the print was fuzzy and she couldn't tell if one of the digits was a 6 or 8. She said she hoped the print would be clearer on the actual census form.

Recommendation: None. The quality of the type on the form used for the test is not necessarily representative of the final form.

*2c. Autotabbing.* The application automatically tabs between groups of numbers. Those participants who looked at the screen when entering the number noticed this and continued without interruption. Those who had their heads down when entering the number used the tab key between the groups (indicating they saw the grouping on the form and remembered seeing the grouping on the screen). This, however, meant that they skipped a block (i.e., the autotab took them to the second block and when they pressed tab the cursor went to the third block). Eventually they saw this and figured out what had happened. During the debriefing, all but one participant said they liked the autotabbing feature.

Recommendation: None

### Item 3: Before You Start Screen (B4UStart)

There were two versions of this screen, one for the topic-based version of the application and one for the person-based version. The following figures show both the topic-based version of this screen (Figure 5) and the sample screen section of the person-based version (Figure 6). The versions were the same in all respects except the sample screen, which included sub-tabs in the person-based version.

In order to inspect the sample screen from the application, the respondent had to scroll down through the second page of either screen. Many respondents elected to start without reading all the information. If test participants are willing to skip this content in a laboratory setting, how much more likely are actual respondents to ignore it?

*3a. Reading the Before You Start Screen:* A few people just clicked on the Start button; some scanned over what was displayed above the fold (about the top of the graphic); fewer actually scrolled down to look at everything; even fewer actually read everything. One participant said it was "too much."

Recommendation: None

*3b. Before you get started paragraph*: One participant was expecting to get tips of the day from the site.

Recommendation: Change the first sentence on the screen to read, "Before you get started, review the sample page below." This is more in line with what we want the respondent to do. This will remove the expectation that we are providing tips to the respondent.

Priority: Moderate
Team Decision: This will not be changed at this time. Since another round of testing will not be conducted before the application is fielded, we will

not know the impact of the change. This can be explored in future usability testing.

*3c. Start button at top of page*:  Several participants remarked that there was important information on this screen and that we shouldn't allow respondents to skip over this page (i.e., we should remove the Start button at the top of the page).

> Recommendation:  We do not recommend removing the button at the top of the page.  Although we agree that it may help respondents if they read the page, behavior of web respondents suggests that they may not read it even if we make them scroll down to the bottom of the page to click on the Start button.
>
> > Team Decision:  There was some interest in removing the Start button at the top of the page, but the team decided to keep the button.  This can be explored in future usability testing.

*3d. Instructions for use of Next and Previous Buttons:* One participant scrolled down to see the Next and Previous buttons after reading about them.  This page obviously didn't meet her expectations.  We believe that the placement of the instruction for the Next and Previous buttons at the top of the screen may make respondents complacent about reading the rest of the screen.  This instruction seems obvious to the respondents, which may make them think that the rest of the page and the site itself are also intuitive.

> Recommendation: The instructions for the Next and the Previous buttons should be shown the same way that other navigation methods are shown -- as call outs from the graphic.
>
> > Priority:  Moderate
> > Team Decision:  This will not be changed at this time. Since another round of testing will not be conducted before the application is fielded, we will not know the impact of the change. This can be explored in future usability testing.

*3e. Instructions not to use Forward and Back Buttons*:  Several participants commented when they read that their browser back and forward button shouldn't be used.  A couple of participants who did not read the screen actually used the back button.  Some were able to press the refresh button and return to the application, others did not know how to handle the "Page Expired" message.

> Recommendation:  The instruction not to use the browser forward and back buttons should be above the Start button.  Moving this may make respondents less complacent about reading the rest of the page because it is giving them information that is specific to the application.
>
> > Priority: Moderate

Team Decision:  There was a suggestion to move the instruction so that it is the first line on the screen.  However, the team felt that because of the structure of the instruction (i.e., a warning triangle and the instruction NOT to use the browser buttons), placing the instruction as the very first thing on the page might be interpreted by the respondent as implying they had incorrectly used the browser button.  The instruction will be moved so that it is after the paragraph, "Before you get started…" and before the sentence "If you prefer to skip this screen…" The format of the instruction will remain the same.  It will be centered on the line.
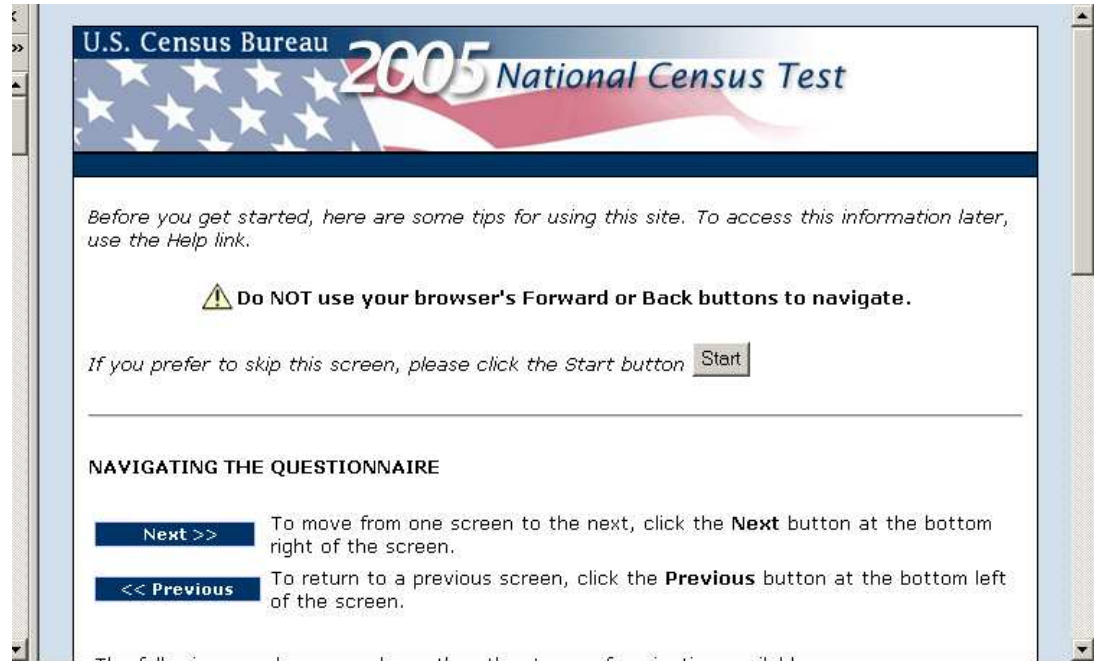


**Figure 5: Instructions not to use the browser navigation buttons**

The following sample screen shows the other types of navigation available.
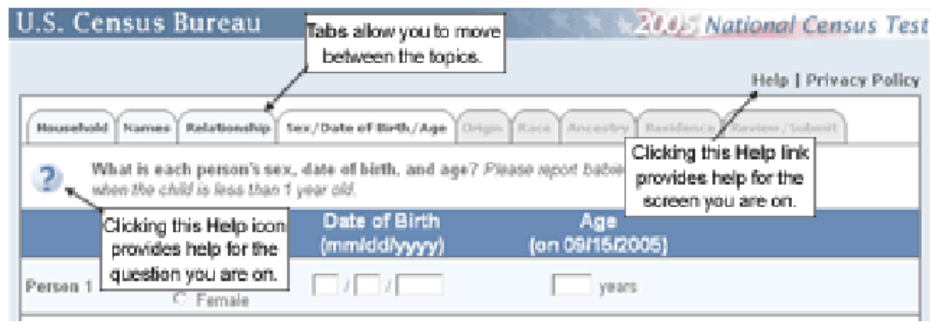


- **White Tabs** indicate the topic you are currently on.
- **Blue Tabs** indicate a previously visited screen or question, which is still accessible to you.
- **Gray Tabs** indicate a topic that has not yet been visited, and therefore, is not yet accessible to you.

**SUPPORTED BROWSERS**

This site works best with one of the following browsers:

- Internet Explorer 5.5
- Internet Explorer 6.0
- Firefox 1.0

**Figure 6. Topic-based version of the Before-You-Start Screen**

The following sample screen shows the other types of navigation available.
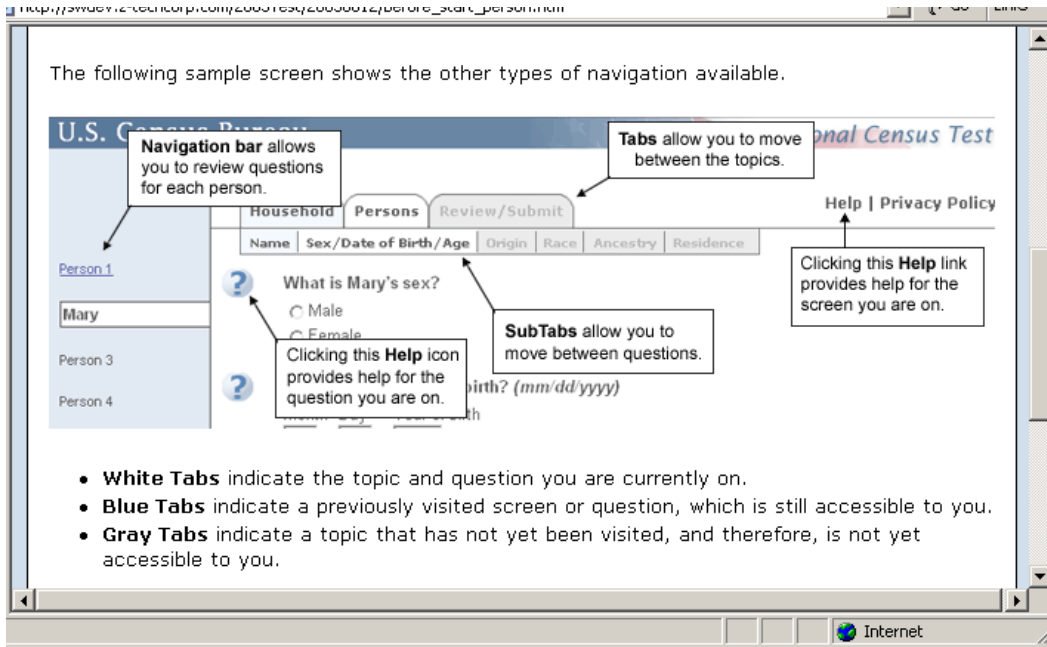
**Figure 7. Sample screen from the person-based version of the Before-You-Start Screen**

*3f. Graphic*: Some participants were confused by the graphic of the sample screen – one participant tried to answer the question about Mary (the sample person on the graphic). Another was confused about the Help link information because the B4UStart screen does not have a Help link.

>   Recommendation:  The graphic should be moved up and display a shot of an entire screen (perhaps some of the white space in the middle of the screen can be removed so that the graphic fits on the screen).  We think this will make it clearer that it is a screen shot.

>   Change the title above the graphic to read "Sample Screen" or "Sample Page." Delete the sentence, "The following sample screen shows the type of navigation available."  Web users are looking for key words; most will not read entire sentences or paragraphs.

>>   Priority: Moderate
>>   Team Decision:  The graphic will not be changed at this time. Since another round of testing will not be conducted before the application is fielded, we will not know the impact of the change. This can be explored in future usability testing.

*3g. Tooltip in Graphic*. One participant was confused by the tooltip that came up over the graphic saying "Sample Person-based Screen."  This text had no meaning to her.

Recommendation: Remove the tooltip that comes up over the graphic.  If this must appear for accessibility requirements, change the wording to read, "Sample picture of a screen."

Priority: Moderate
Team Decision: The wording of this alt tag was discovered in the accessibility testing and will be changed.

*3h.  Explanation of tabs*: From the reaction of the participants when reading the explanation of the tabs, we think it is too wordy.  Those who read the explanation did not seem to follow it or remember it later.  Forgetting this kind of information is entirely normal when many demands are being made on limited memory functions (e.g., Baddeley, 1990; Reason, 1990).

Recommendation:  Revise the wording of the instruction for the tabs to be more in line with a web application – that is, bulleted points instead of full sentences.  For example,
    White tab- current topic and question
    Blue tab – accessible -- already visited question
    Grey tab – not accessible -- not yet visited question

Priority: Low
Team Decision:  This will not be changed at this time. Since another round of testing will not be conducted before the application is fielded, we will not know the impact of the change. This can be explored in future usability testing.

*3i.  Supported browsers:*  One participant, who uses Netscape as his browser at home, got to the bottom of the screen and said it didn't do him much good to know the supported browsers at this point.  He said he could either quit the application and open it in IE or just continue and know that if something didn't work well, it was probably because of the browser he was using.  He said he would not quit and start again in IE.

Recommendation: None

***Item 4: Household Screen (Popcount – Tenure)***

**Before you answer the first question, count the people living at your house, apartment, or mobile home using our guidelines.**

☐  We want to count people where they usually live and sleep.
☐  For people with more than one place to live, this is the place where they sleep most of the time.

<span style="color:blue">EXCLUDE these people</span>                                              <span style="color:blue">INCLUDE these people</span>
*We will count them at the other place.*

**Figure 8. "Exclude" and "Include" links preceding the question on number of people in the household**

*4a. Popcount. Include/Exclude links*:  As illustrated in Figure 7, the EXCLUDE and INCLUDE links were confusing to some participants.  They do not follow logically from the text displayed before them.  That is, they seem to appear "out of the blue" for no apparent reason. One dry-run participant, who grew up with computers, thought that clicking on those instructions would cause the computer to take an action of excluding or including people.  Likewise, others did not recognize those as help links, but we think that was because they weren't expecting a help link to be positioned there. If they did think these links were help, they weren't sure how the help would function. Those who had read the residency instructions on the paper form recognized the "exclude" and "include" as referring to those lists of people.  Others knew right away that they were help links.

We think there are at least two issues here.  First, participants were told to count the people using "our guidelines," and then they were given a couple of guidelines (the two bulleted statements).  So, now they are ready to give the count.  By telling them to include or exclude someone, the text confused them because they hadn't told us anything yet.  How could they exclude someone they hadn't reported yet?  So, what they intuitively knew to be links (most reported this in the debriefing), they dismissed because they were expecting to put in a value for the number of people in their household.

The second issue is the title of the links.  It suggests an action on the part of the respondent or the computer to include/exclude **these** people. It is unclear to which people "these" is referring, since no person has been mentioned yet.

> Recommendation: Merge the INCLUDE/EXCLUDE links and title it "More Guidelines (Help)." This gives the respondent cues that these are a continuation of the guidelines started in the two bullet points. It reinforces the inference that the two bullet points ARE guidelines and clarifies what the respondent will get by clicking on the link.  Including the word Help in parenthesis verifies that it is

information and that respondents should expect the link to behave as a help link usually behaves.

> Priority: Moderate
> Team Decision: While there was support for this change, the team did not think that stakeholders outside the team would be able to react and approve this change in the time frame required (end of day). Note that team members did try to get approval but it wasn't possible in the short time frame.

*4b. Wording of Popcount instructions*: In line with the discussion of the include/exclude links presented above, we recommend changes to the wording of the Popcount instructions to help tie the instructions to the question.

> Recommendation: Instead of a period, use a colon at end of the sentence "Before you answer the first question…" to signify there is more to come. A colon following "guidelines" would signify that what follows are the guidelines. Otherwise, the respondent is likely to be uncertain of where to look for the guidelines.

> > Priority: Moderate
> > Team Decision: Team members from POP and DSSD were comfortable with making this revision. However, the Internet staff in DSCMO did not think it was worth the risk and resources to make this change if the other changes for this instruction (change the include/exclude link) were not made.

Again, the two bullets are too wordy for Internet readers. Revise the first bullet to read, "Count people where they **usually** live and sleep." Bold the word usually so that readers who skim will get the main thrust of the bullet. The second bullet should read, "For people with more than one place to live, count them whey they sleep **most of the time**". Again, bold the words most of the time. Also, using the word "count" continues the theme of these being guidelines for counting the people in the household.

> Priority: Moderate
> Team Decision: Team members from POP agreed that the words "usually" and "most of the time" could be put in bold font. They did not, however, want to change the wording of the bullets because this would deviate from what was printed on the paper form.

**Figure 9. From top to bottom: Questions on popcount, undercount, and tenure**

*4c. Undercount question*. A couple of participants were confused by this question (See the middle question in Figure 9). They thought of their situation and said that if there were additional people, they would have counted them in the previous question.

There was, however, one participant with an interesting living arrangement. He lived with unrelated people in a home owned by a church that rented out rooms. There was a person who had not lived at the house for quite a while but kept up his rent so that he could use the address as his permanent mailing address. The participant did not count this person in the Popcount question because the person did not usually sleep there. He did report "yes" in the undercount question and marked the "nonrelatives" checkbox. He said the "people staying here temporarily" box did not apply because this person didn't stay there.

> Recommendation: None

*4d. Tenure question*. When they got to the tenure question (last question in Figure 9), about half the participants read, "In this house..." rather than "Is this house..." This misreading of "in" for "Is" has happened repeatedly in previous usability tests.

One participant who owned the house but had housemates who paid him rent had to think about how to answer this question because two of the response options were appropriate and he could only mark one. He chose the "owned" response because he was filling out the form.

Several participants read through at least the first two response options if they owned their home.  This was probably because the word "owned" for both categories was in their field of vision.  They did remark, however, that the response options were wordy.

Recommendation:  None

### Item 5:  Household/P1 Name Screens (P1 Name – Same last name)

*5a.  P1 Name:*  Several participants got hung up on this question, as shown in Figure 10, but none of the hangups resulted in inaccurate data. Most had to read it several times before getting the gist.  One participant saw the word "renter," said the question wasn't relevant, and skipped the question.

When the next question asked for Person 1's telephone number, this participant wondered where she was going to be asked to enter Person 1's name.  She then went back to this question.  Another participant was confused because she said they had rental property.  While she didn't think we wanted to collect data for the rental's address, that was the only "renter" she could identify with.  The participant who rented in a household in which the owner also lived, entered the name of the owner as Person 1.  Another participant, after reading the question several times finally said, "So, you want the name of any adult living here."  The question structure may be the problem.  The word "this" (this is Person 1) refers grammatically to the adult referred to in the second sentence, but many participants didn't read the second sentence because it didn't apply to them.

Recommendation:  The question for collecting the name of Person 1 is wordy and cognitively difficult.  It should be revised for reading on the Web.  In addition, the term Person 1 needs to be mentioned closer to the beginning of the question, not the end.

**Figure 10. Questions on Person 1's name and telephone number; same last name for everyone in the household. Text is identical in both the person-based and topic-based versions of the form.**

> Priority: High for 2010
> Team Decision: These results will be passed along to the content R&D group.

*5b. Telephone Number question.* Autotabbing on the phone number was acceptable for most but not all participants. Some were habitual tabbers.

> Recommendation: None

When asked to change the telephone number, many participants put the cursor over the number and tried to type over it. This approach did not work. They figured out that they needed to delete the entry and then type in the new entry.

> Recommendation: If possible, allow respondents to type over numbers already entered.

> > Priority: Moderate
> > Team Decision: It is possible to type over an entry if the respondent toggles the insert key on the keyboard. A change could be made to the application that would automatically allow this. However, this change would be time-consuming and somewhat risky to make. Therefore, we

will not make this change for the 2005 application but note it for future Internet applications.

A couple of participants questioned why we asked the phone number of Person 1. They said that everyone in the household uses the same phone number. In the case of the renter living in the same home as the owner, she said that if she has put herself down as Person 1, then she could enter a phone number. She just left it blank. Another person in a household of unrelated people wanted a "don't know" box.

> Recommendation: None

*5c. Same Last Name question*. Most participants did not question why we were asking this question (Figure 10). One participant suggested it had to do with sociological research.

One participant whose son lived with her had a hyphenated last name (maiden name-married name). She said "yes" to this question, but a problem with the application didn't prefill last name on the Name screen. When she got to the Name screen, she only entered her married name as her son's last name.

> Recommendation: Fix application so that if same last name is marked "yes", the last name of Person 1 is prefilled for others in the household. (Invoking the edits on this page may have something to do with this problem.)

>> Priority: Moderate
>> Team Decision: This is a known defect that has been fixed in a later version of the application.

In another instance, the participant said everyone in the household had the same last name. However, on the Name screen, he added JR after his son's last name. When he went back to the previous screen and came forward to the Name screen, the JR had been deleted.

> Recommendation: Once a respondent has made a change to the last name on the Names screen, don't overwrite it with the last name of Person 1.

>> Priority: Moderate
>> Team Decision: This is a known defect that has been fixed in a later version of the application.

Add a field for suffixes to last name. Use this suffix when displaying names in the left navigation bar of the person-based application. This would help distinguish people with the same first name on the left navigation bar.

> Priority: Moderate

Team Decision: This would add a field to the output record unless post-processing is done to append this to the last name that was entered. This change would be time consuming and somewhat risky to make.
Therefore, we will not make this change for the 2005 application but note it for future internet applications

## *Item 6: Name/Relationship*

*6a. Name.* On the person-based application, the question asks for Person X's name (e.g., Figure 11). However, when the participant navigated back to this screen, Person X was replaced with the name that was entered. For example, the question would read, "What is John's name?"

Recommendation: Hard code the words "Person X's" in the question so that it isn't replaced with the name that the respondent enters.

Priority: Low
Team Decision: This is a major change to the application and carries a high risk. Since it will happen infrequently and should not cause major confusion to the respondent, we will not make this change to the application.



**Figure 11. Person-based approach to collecting Person 2's name and relationship to Person 1**

Many participants did not enter a middle initial.  Either they didn't know it or just tabbed over it.

> Recommendation:  None

  *6b.  Relationship*.  Most participants in the person-based version (Figure 11) claimed to have seen the "Related" and "Unrelated" headings, but many did not see them in the topic-based version (Figure 12).  We think there is something about the dark blue bar that makes people overlook the embedded white headings.

> Recommendation: If we must have the horizontal blue bars that separate the questions from the response options, take the names of the headings out of the blue bars and put them in black above the columns of options.  (The horizontal blue bars serve no functional purpose and may have the effect of causing people to miss the question.)
>
>> Priority: Moderate
>> Team Decision:  This will not be changed at this time. Since another round of testing will not be conducted before the application is fielded, we will not know the impact of the change. This can be explored in future usability testing.



**Figure 12.  Topic-based approach to collecting data on relationship to Person 1.**

Even if participants didn't see the headings, they understood that there was a logical grouping to the lists.  One participant in an unrelated household started down the first

column and before finishing it, abandoned it for the second column.  For others in the household, he immediately went to the second column.

      Recommendation:  None

One participant suggested putting a vertical line between the columns on the topic-based application.  While we don't necessarily recommend this, we think it shows that this participant made a distinction between the two groupings.

      Recommendation:  None

### Item 7:  Sex/Date of Birth (DOB)/Age

*7a.  Sex.*  One participant missed filling in the sex of Person 3 in the topic-based application (Figure 13).  She entered her sex, date of birth and then when she clicked on age, the software calculated it for her.  She commented about being  pleased to see this.  Then she entered the sex and date of birth for the second person and was anxious to click on the age box to have it calculate age.  For person 3, she went straight to date of birth and clicked on the age box.  As she was hitting the Next button, she realized she had missed sex for person 3.  Although she didn't read the edit message that came up, she knew what the problem was and said she was pleased that the computer was checking her work.

      Recommendation:  None

**Figure 13. Topic-based approach to collecting sex and date of birth, with autocalculation of age if date of birth is entered according to the given format (mm/dd/yyyy). In the example shown, date of birth does not follow the specified format.**

7b. Date of Birth/Age. Some participants expected DOB to autotab and some did not. Those who did not had to make corrections. Some expected autotab from the end of date of birth to age.

 Recommendation: None

One participant put in a single digit month and did not get an edit message.

 Recommendation: Fix the application so that the edit message comes up when only one digit is entered for month.

  Priority: Moderate

  Team Decision: The application is correct as is. No edit message should come up when only a single digit month is entered.

Everyone seemed surprised, but pleased that age was calculated and filled in (Figure 14). One person-based participant built up speed to the extent that she stopped checking to see if the ages were correct. When she went to put in her fifth daughter's year of birth, however, she realized she had an error in the fourth daughter's year. She used the names in the left bar to go back to the fourth daughter's data and make the correction.

**Figure 14. Person-based approach to collecting sex and date of birth, with autocalculation of age if date of birth is entered according to the prescribed format (mm/dd/yyyy).**

Another participant entered the wrong year of birth for second son. When she saw the age calculation, she recognized that she had the wrong age and went back and changed year of birth.

We think that calculating age is good because it relieves respondents of the cognitive burden of figuring out age especially for a date that is in the future. When the household member was under one year old, the computer calculated the age as 0. One participant expected this because he had read the instruction; another did not.

In a couple of instances, the participant didn't verify the age, assuming the computer would get it right.

> Recommendation: None

An edit message incorrectly came up when a person was under 1 year old.

> Recommendation: Fix the application so that an edit message does not come up when 0 is entered for age.

>> Priority: High
>> Team Decision: This is a known defect that has been fixed in a later version of the application.

Some participants saw the mm/dd/yyyy (Figure 14). One participant entered 2 digits for month and day, saw that it filled the data-entry space, then entered 2 digits for year. She noted that it didn't fill the space, saw the "yyyy" and changed year to four digits.

One participant said she much preferred entering numbers instead of having a drop-down box. She thought it was easier to make a mistake with a drop-down box.

One participant in an unrelated household wanted a "don't know" box because he didn't know the birth dates of others in the household. However, when he guessed and saw that he could go on without entering something, he thought this was just like entering a "don't know."

    Recommendation: None

### Item 8: Origin

*8a. Origin*. Participants had no trouble answering this question in either the topic-based version (Figure 15) or the person-based version (Figure 16). The use of the word "of," while grammatically correct, was not conversational in tone. It caused a couple of participants to have to read the question twice.

    Recommendation. None.



**Figure 15. Topic-based approach to collecting Hispanic, Latino or Spanish origin for everyone in the household.**

**Figure 16. Person-based approach to collecting information on Hispanic, Latino, or Spanish origin for each person in the household.**

### *Item 9: Race*

*9a. Race.* For the topic-based version of the question (Figure 17), it seemed that people were not reading the question at all. They looked immediately at the grid of names by response categories and thought they knew what they were being asked. For the person-based version (Figure 18), it appeared that participants read the question for Person 1 and maybe for Person 2; after that, they had built up an expectation for the race question to occur at this point. We think the blue bar separating the question from the responses in the topic-based application is a factor in this behavior.

**Figure 17.  Topic-based approach to collecting racial information for all persons in the household.**



**Figure 18.  Person-based approach to collecting racial information for each person in the household.**

Recommendation:  For the future, consider using a lighter blue with dark text for the headers.   The dark blue acts like a frame separating the question from the response options.  A lighter blue might lessen this effect.  Conduct research on two or three combinations of foreground and background in Round 2.

Priority:  Moderate for 2010
Team Decision:  This can be explored in future usability testing.

A couple of participants questioned the statement about Hispanic origins not being races. One dry-run participant said that people might be insulted by that statement.

> Recommendation:  Soften the statement about Hispanic origins by saying, "For this census test, Hispanic origin is not considered a race."

> > Priority: Moderate
> > Team Decision: This is the statement that is on the paper form, so it won't be changed in this application.  This recommendation will be passed along to the content R&D team.

During the debriefing, one participant was curious about what the example link for other race would give him.  He said he expected to see a list versus just being told to use it if none of the other categories applied. Figure 19 shows the help given in both the topic- and person-based applications:
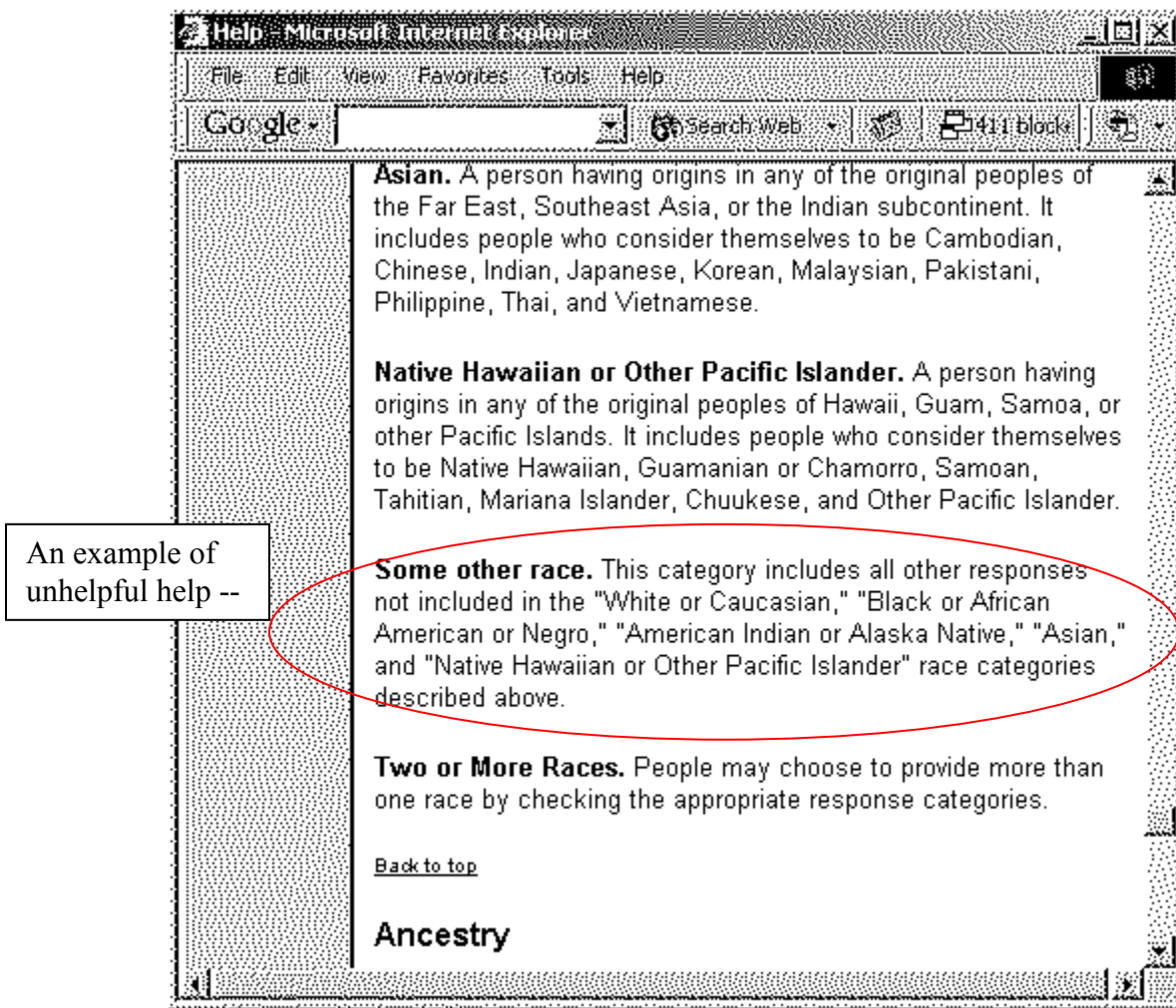
An example of
unhelpful help --

**Figure 19. Help provided when the respondent clicks on Examples for "Some other race"**

Recommendation: Revise the help text for other race to give examples.

Priority: Low for 2005; Moderate for 2010
Team Decision: The help text will not be revised. This recommendation will be passed along to the content R&D team.

None of the participants clicked on any of the examples when initially completing the form.

No one complained about "African Am.," being abbreviated. However, when they entered it into the ancestry field, they wrote it out as "African American."

Several participants were surprised that the term "Negro" is still being used. Many wouldn't read the word Negro aloud. When one participant saw "Black, African Am,

54

Negro" on the review/submit screen, she wondered where that term came from.  It made her go back to the question to see if it was part of the category she checked.

Many participants didn't necessarily recognize that they could enter more than one race. One participant who had participated in an earlier study actually went back to the race screen and said she had recommended that people be able to enter more than one race. When she saw this was possible, she said she was pleased.  Several people who typed in African American and American Indian on the Ancestry screen only checked the African American box on the race screen.   Another participant who explained on the Ancestry screen that his wife was both African American and white only marked the African American box on the race screen.  It seems as though participants were entering whichever race they identified with.

One participant remarked that she was surprised that "Native American" wasn't on the form.

> Recommendation.  None

### *Item 10:  Ancestry*

Many participants reported that they did not understand the difference between the ancestry question (Figure 20) and the race and origin questions.  They saw the ancestry question as redundant.  For example, one participant, who had said in the origin question that her five daughters were Hispanic (on their father's side) and listed herself as being of Polish ancestry, checked "Don't Know" for all her daughters on the ancestry question. When probed about this, she said she did not know anything further than Hispanic.  A discussion with the daughters' paternal grandfather had indicated some Mexican descent along with others, so she did not have a clear enough idea of what to put down other than Hispanic.  She felt that she had given the Census Bureau enough information.  Another participant reasoned that ancestry was just getting more details than race.  Another questioned why this screen said African American and the previous one (race) said Black, African American.

**Figure 20. Topic-based version of the ancestry question and response area. Question text was identical in the person-based version of the Internet form for the 2005 National Census Test.**

Another participant, who indicated that she was African American on the race question, initially checked the "Don't Know" box for herself. When answering the question for her son (topic-based), she read Italian and then skipped over to Jamaican. She said her son's father was Jamaican and wondered how far back she should go to report. She used the question mark to get help. She saw that she should go back several generations, so she changed her ancestry to be African American and American Indian. She entered three ancestries for her son. As everyone did, she said that three was the maximum she could enter.

> Recommendation: None

No one understood that more than three ancestries could be entered. The three text fields send a very strong signal that three is the limit. However, test participants understood that they were not required to enter something in all three spaces.

> Recommendation: Communicate to the respondent that more than three ancestry categories can be entered.

> > Priority: Moderate
> > Team Decision: No change will be made to this item.

One participant entered African American for himself. He readily checked the "Don't Know box" for others when he didn't know. He said during the debriefing that he could

guess the ancestry for one of the housemates, but didn't know for sure, so he wasn't comfortable entering anything.

> Recommendation: None

Participants generally glanced over the examples. However, some didn't see them as examples but rather as a list that they should choose from. The African-American participants in our tests generally found this question easy (although redundant) because they saw an example that was appropriate for them. This statement is supported by the fact that a couple of participants asked why they had to type in their entries rather than choosing from a list. Others wanted more "examples."

> Recommendation: Participants seem to be looking for themselves in the list of examples. Since there is space on the form, more examples should be provided -- perhaps Swedish, Norwegian, or some Scandinavian country? European countries are not well represented (e.g., England, Ireland), nor are South American countries.

>> Priority: Moderate
>> Team Decision: The examples will not be changed.

A participant commented that ending the examples with "etc." is a slap in the face to all those who are not mentioned. It puts everyone else in the same generic bucket.

> Recommendation: Do not use "etc." to end the examples.

>> Priority: Moderate
>> Team Decision: The examples will not be changed.

Participants made various decisions about what entries to include. One father entered only his own ancestry for his son, omitting the boy's Caribbean ancestry on his mother's side. Others went back to one grandparent on each side of the family. There were various strategies for answering this question.

One participant tried to enter her tribe (Nanitoke Lenne Lenape), but it wouldn't all fit. She was the only one to specify a tribe.

One participant was unsure of how to mark her husband who was Jewish. She read the help. Eventually she entered Germany and Austria.

A couple of topic-based participants wanted a way to enter the same for everyone. A couple of participants cut and pasted from the example list.

> Recommendation: None

*Item 11:  Other Residence*

*11a.  Other Residence.*  In general, participants seemed to understand the question and be able to answer it (Figure 21).   Many read the response sub-options for the "Yes" to help them determine their answer.  For example, one participant initially entered "no" for each of his children but then reconsidered after skimming the sub-option.  He said that they sometimes stayed with their grandmother or aunt and changed his answers to "yes."  During the debriefing, he went to help for this question and saw that they should have been marked "no."



**Figure 21.  Person-based version of the overcount (Other Residence) question.  Except for the content of the fill, the question text and response options were the same in the topic-based version.**

One participant said her son sometimes stays with his father.  She went to help to see if that should count.  Help didn't seem to directly answer her question, but from the text she got the sense that he should not be counted as living away.

One participant said he sometimes rents out his house for the summer while he goes off and stays in the national parks.  He wondered if he should mark "to stay at seasonal residence."

Recommendation: Scenarios from real life such as those we encountered might be useful in help.

Priority: Low for 2005; Moderate for 2010

Team Decision: The help file will not be changed.

In interviews with two people from the same unrelated household, participants checked different sub-options for one of the housemates. They both explained that the housemate lives at <u>this</u> address to be close to work. Her "permanent residence" is with her family. She stays at this address Mon-Thurs. Both marked that this housemate lived elsewhere – one marked "other reason," and the other marked "to be closer to work."

Recommendation: None

### Item 12: Review/Submit

*12a. Getting back to Review/Submit.* One participant went back and changed an answer as instructed during the test. However, she hesitated to click on the Review/Submit tab (Figure 22) to go back to that screen. She assumed it would submit her data, but she wanted to review her data again. She suggested making "Submit" a separate tab.

Recommendation: None

**Figure 22. Review/Submit screen: Content and layout are identical in both person-
and topic-based versions, except that the person-based version has a left navigation
bar listing the names of household members. This example shows the change in
labeling of "Residence" to "Other Residence," based on usability findings.**

*12b. Instructions.* Although most participants wanted to review their data before
submitting, the "Submit Census Information" button and the amount of information that
fits on the screen act as barriers between the review instruction and the information to be
reviewed. Some did not understand what was meant by "the underlined answer" or "the

[Blank]" because they had not yet seen an underlined answer or a blank (Figure 22). Also, the instruction on how to send information was not where it was needed. The participants were thinking about reviewing their data and were instead given an instruction about submitting it.

> Recommendations:  Revise the screen layout so that it is more in line with how the respondent uses it.  First, split the instructions for changing a response and completing an answer left blank into two different instructions to reduce the complexity.  These instructions can be side-by-side.  Keep three bullets, with only the instruction on the first bullet being different.  The first bullet for changing a response should say, " Click on the underlined answer."  The first bullet for completing a blank answer should say, "Click on the underlined [Blank]."

>> Priority: Moderate
>> Team Decision:  This change will not be made. This can be explored in future usability testing.

> Next, move the instruction for sending information to the bottom of the screen, after all of the data since this is where the respondent will use this information. This will bring the data further up on the screen and display some "above the fold."

>> Priority: Moderate
>> Team Decision:  This change will not be made. This can be explored in future usability testing.

> The Submit button can remain at the top of the screen as well as the bottom, but should not be in the line of vision between the instructions and the data.   It should be moved  over to the right side of the screen.

>> Priority: Moderate
>> Team Decision:  This change will not be made. This can be explored in future usability testing.

In general, the layout of the data on the page was understandable to the participants. However, the item names "Origin" and "Residence" gave them pause.  Many did not remember that Origin referred to Hispanic, Latino, or Spanish origin.  They thought instead that it meant that they did not put in any Origin when, in fact they had (perhaps they were remembering entering an answer for Ancestry).  Participants associated the word "residence" with their current residence, and they thought it was supposed to be their current address.  Thus, they were surprised by a "no" for residence.

> Recommendation:  Change the title of the Origin question to Hispanic Origin or fill the response as "No, not Hispanic, Latino, or Spanish" or "Yes, Hispanic, Latino, or Spanish."

Priority:  Moderate
Team Decision: If approval for this can be obtained by the end of the day, the change will be made.  Otherwise, it will remain as is.  (NOTE: approval was not obtained.)

Recommendation:  Change the title of the Residence question to Other Residence.

Priority:  Moderate
Team Decision:  DSSD is comfortable with making this change (as illustrated in Figure 22).

A couple of participants asked why there was a question mark after their answer to the Ownership question.  They reasoned that the computer was asking them if it was correct instead of just displaying what they had entered.

Recommendation:  Remove the question mark after the Ownership response.

Priority: Moderate
Team Decision:  This is a known defect that has been fixed in a later version of the application.

## Item 13:  Thank You Screen

*13a.  Thank you.*  Participants understood this to mean that they were finished.  Many said they would stop reading after they saw the words "Thank you" (Figure 23).  Some commented that it all looked "boring" and was too much to read.  One wanted color in the seal.

One participant suggested having the email address available on the paper form.

Recommendation: None
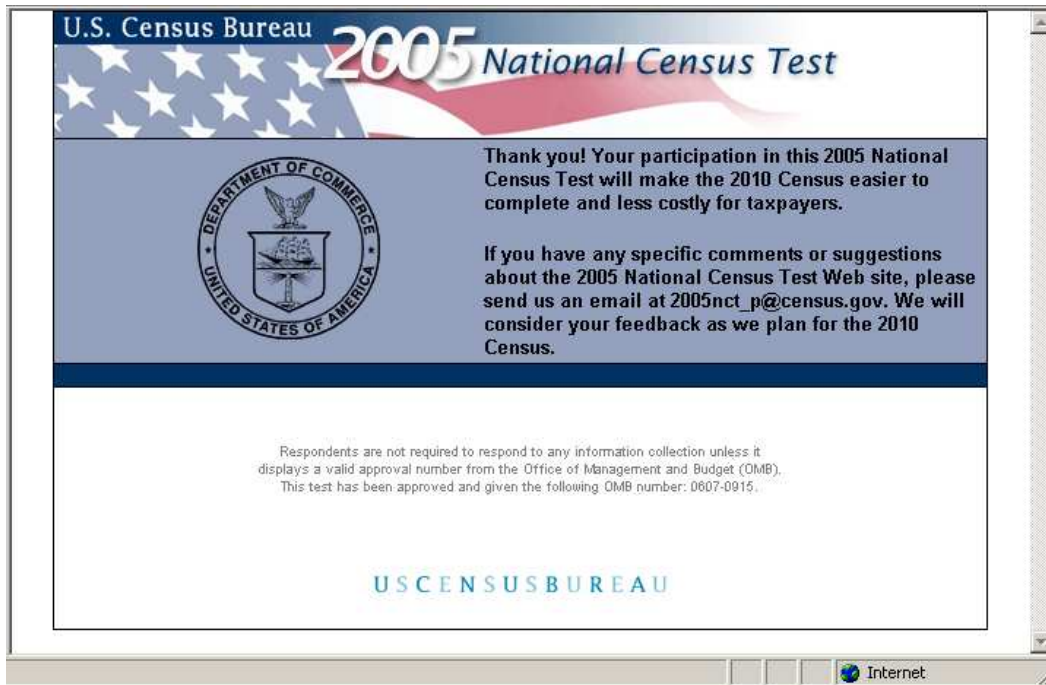
**Figure 23.  Thank-you screen:  2005 National Census Test**

*Item 14:  Help text*

*14a.  Help text wording*:  In general, help is too wordy.  People do not want to read sentences and paragraphs of the kind illustrated in Figure 24.  They are looking for key words.

> Recommendation:  Revise the help text, using appropriate writing style for the Web. For guidance on writing for the Web, see Redish (2012).
>
> > Priority:  Moderate
> > Team Decision:  The help file will not be changed.  This can be explored in future usability testing.
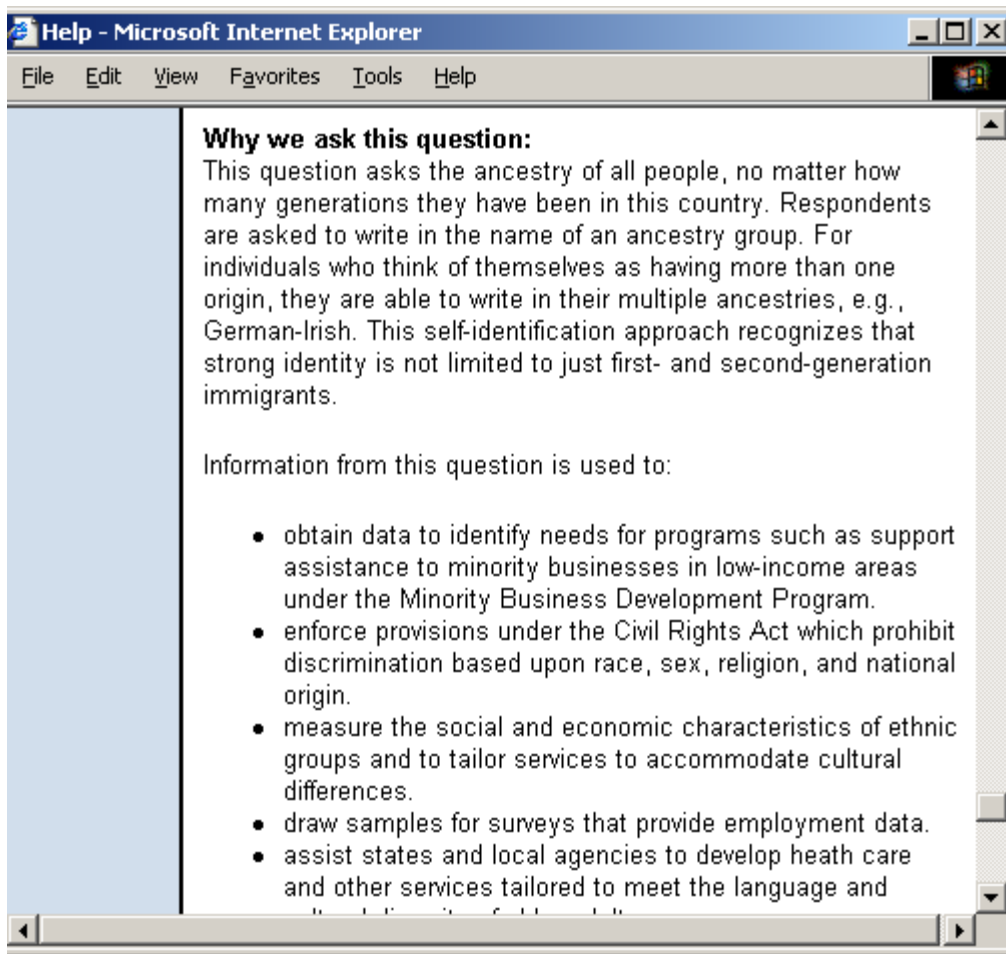
**Figure 24. Example of pop-up help text taken from the explanation given for the question on ancestry.**

*14b. Question Mark Icons*: Those who read the Before-you-start tips understand that the question marks are help icons. For the most part, others do not understand that these icons can be clicked for help with a question. They thought a question mark was just there to indicate the next question. Examples of the question-mark icon are provided in several previous figures (e.g., Figure 20, Figure 21).

> Recommendation: None

*14c. Links*: The Exclude/Include links are problematic. People don't know what they are trying to communicate, and some are afraid to click on them without knowing where the links will take them. See earlier discussion about these links and the illustration in Figure 7.

> Recommendation: See earlier recommendation (Item 4).

*14d. "Other race" help*: Participant was looking for a list of acceptable options, not just a statement saying to use "other" if it is not any of the other races listed (cf. Figure 19).

Recommendation: Enhance the content of the help to answer questions that are frequently asked.

*14e. Ancestry*: The help refers to countries and uses nouns, whereas the examples use adjectives. This caused one participant to enter "Germany" instead of "German" for ancestry. Also, the help does not indicate that more than one entry can be entered on each line.

Recommendation: Include an example in the ancestry question that illustrates that more than one ancestry can be entered on a line. For example: Irish, Polish.

Priority: Moderate

Team Decision: See earlier discussion.

One participant said she was reluctant to go to help because she wondered how she would get back.

Recommendation: None

## Item 15: Edits

*15.1 Generally Positive Reaction to Edits*. In general, participants who received edit messages (Figure 25) when completing the form reacted positively to them. They remarked that the red arrows and highlighting of the affected areas in red led them immediately to where the change(s) had to be made. For example, one participant didn't notice right away that age was calculated. She typed it in. This caused a 3-digit number to be entered. When the participant hit the Next button, the edit message came up. She knew there was a problem and, because of the highlighting, immediately saw what it was. She made the change and went on. She did this for the second person also. The age calculated for the second person was different from the value she would have entered. This caused her to pause. She realized the age was calculated as of September and that she would have had her birthday but the second person would not have. By the third person, she let the software calculate age.

Recommendation: None
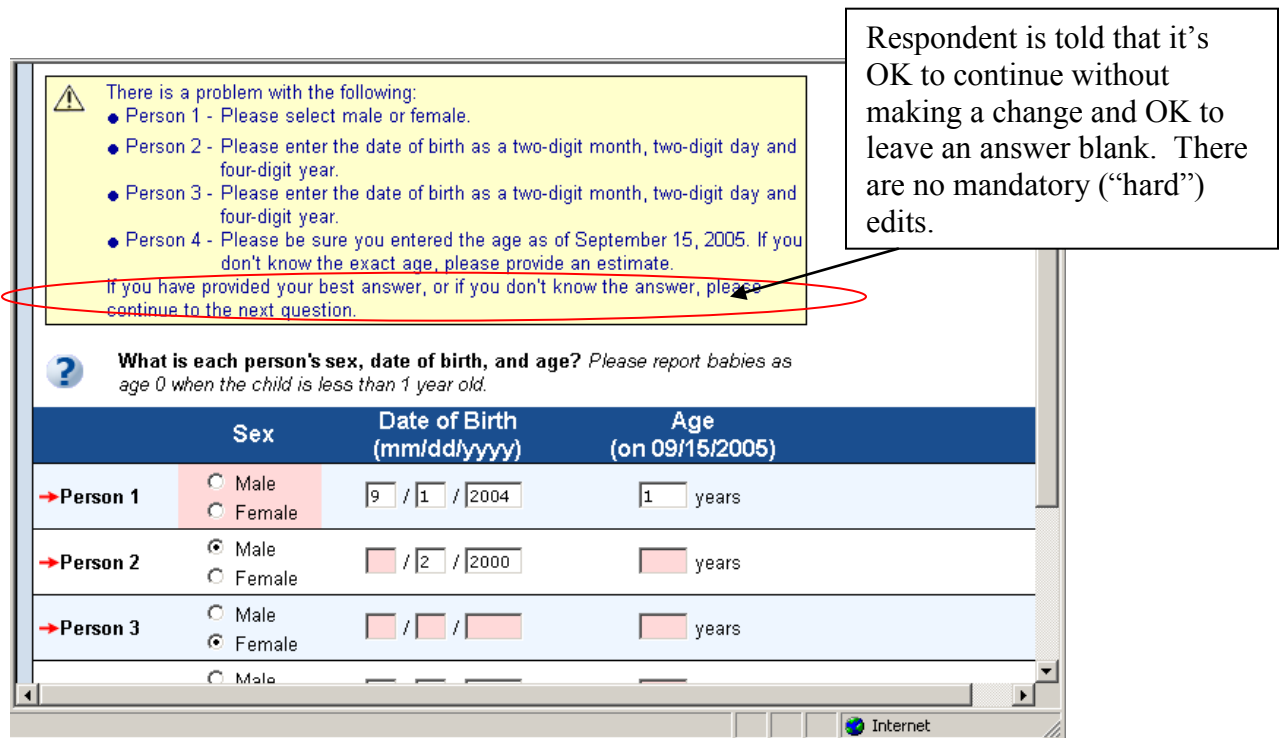
**Figure 25. Examples of edit messages in the topic-based application: Having left some answers blank, the respondent receives several instructions upon attempting to go on to the next page.**

15.2 Using Previous button. In the topic-based application, one participant tried to go to the previous screen from the Name screen and got an edit message. However, since this wasn't what she was expecting, she never saw the message. Instead, she remarked that maybe the application would not let her go back. She tried again and it worked so she said she must have done something wrong the first time she tried.

> Recommendation: Pressing the Previous button should not display the edit messages for the current screen.

> > Priority: Moderate
> > Team Decision: This is how the application was designed to work. It would be a major change to make at this point. We should consider this for future Internet applications.

One participant didn't know the birth dates of his housemates. He was uncomfortable guessing for the first person, but he guessed an age, got the edit message and was relieved that he could go on without having to enter something. He was comfortable giving an age or leaving it blank for everyone else.

The part of the message that says the respondent can go on without entering something was important to several participants. It seemed to soften the edit message and reassure the participant that he/she could continue. This part of the message is illustrated within the red oval drawn on Figure 26. Previous usability testing had shown that respondents

assumed they could not go on without resolving edit messages, when there was no such encouraging text.  This text was added to assure respondents that it was OK to continue without making a change and OK to leave something blank.  It seemed to work well.

Recommendation:  Research different approaches to edit messages

Priority:  Moderate for 2010
Team Decision.  This can be done, time permitting, during the second round of usability testing.

One participant used an apostrophe to approximate an acute accent over a letter in his daughter's name.  He received an edit message on this.  He was annoyed and said it was "insulting [his] intelligence" because he had purposely entered what he had entered.  He thought it would be better to ask the user to verify what they had entered instead of saying there is a problem with what they entered.  The apostrophe was removed without informing him that a change would be made.

Recommendation:   Allow additional special characters (e.g., apostrophe, umlaut) in first and last names.

Priority:  Moderate; if the respondent keeps getting edits on this, it will raise the annoyance level; the correct spelling of their name is important to them.
Team Decision:  No change will be made to the 2005 application. This, however, most likely will be done for 2010

*Item 16:  Navigation*

*16a.  Use of tabs*.  The tabbed design for the topic-based version of the Internet form is shown in Figure 26; and the tab/sub-tab design for the person-based version is shown in Figure 27.  Some participants recognized right away how the tabs worked; others never caught on to it.  At least one participant described the tabs as a kind of progress indicator, but she repeatedly mentioned that we should have a time indicator.  She likes to know how much she has left to do. (She also commented that the Internet form should look exactly like the paper form.)

No participant navigated backwards using the tabs.  However, when trying to return to the Review/Submit screen several participants (particularly with the person-based instrument) first hit the Next button a couple of times and realized there had to be a quicker way.  Then they remembered reading there was a tab they could hit to return to the Review/Submit screen and quickly found it.

Recommendation:  None
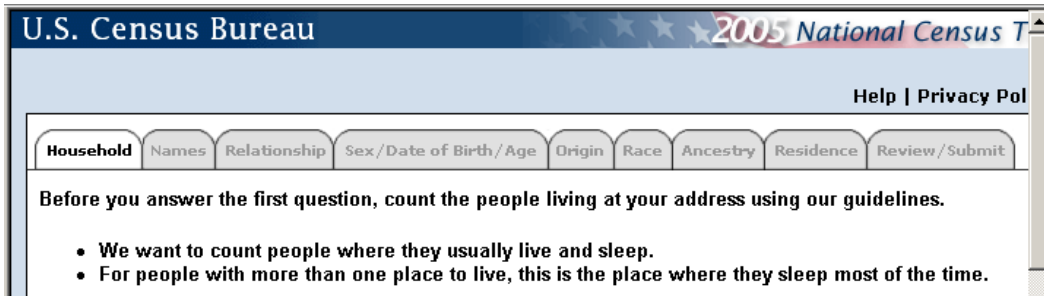
**Figure 26.  Tabs for navigation within the topic-based version of the Internet form for the 2005 National Census Test**



**Figure 27.  Tabs and sub-tabs for navigation within the person-based version of the Internet form for the 2005 National Census Test.**

*16b. Left  Navigation*.  The person-based version featured a left-navigation bar, as shown in Figure 27, for example.  One participant who had to go back and change her daughter's date of birth, clicked on the name to go back.  Aside from this participant, others either did not notice the bar or did not need to use it, so they did not question how it worked.  Before they are clicked, the names are in pale blue on a pale blue background (low contrast) and are easy to miss.  They do not draw the user's attention, especially if the household has only one or two people.

> Recommendation:  Increase the foreground/background contrast of the names in the left navigation bar. If possible, use a larger font for the names.
>
> Priority:  Moderate
> Team Decision: This will be explored in future rounds of usability testing.

16c.  *Enter key*.  Several participants hit the Enter key after typing in an answer.  This invoked the edits for the screen because some items on that screen were left blank.  After seeing the edits, they figured out that it was because they hit Enter and "learned" to use the Tab key to navigate.

Recommendation: Disable the Enter key so that it does not invoke the edits when hit.

Priority: High
Team Decision: This is a known defect that has been fixed in a later version of the application.

*16d. Review/Submit*. Participants understood that the answers on the review/submit screen were links and that they could click on them to return to the question (Figure 21). However, a couple of participants thought that when they hit the Next button after making a change, they would be returned directly to the review/submit page.

Recommendation: None

## Item 17: UI Design

One participant remarked that it seemed odd to be answering questions about himself in the third person (a one-person household). For example, asking, "what is John's age" instead of "what is your age."

Recommendation: In future iterations, if the user enters "1" for Popcount, do not fill the following questions with Person 1's name. Instead, use "your."

Priority: Low for 2005, Moderate for 2010
Team Decision: We should consider this for future Internet applications.

One topic-based participant who had several people living in the household said that he would like to see "floating headers" so that he would not have to scroll up to re-read the question. He described "floating headers" as the locking of the rows and column on an Excel spreadsheet so that the titles are always visible.

Recommendation: Consider implementing floating headers in future prototypes.

Priority: Moderate for 2010
Team Decision: Look into the feasibility of doing this and research during the second round of usability.

*Screen design*. The dark blue bar has unintended effects on user performance. Examples of the blue bar appear in several previous figures (e.g., 12, 13, 14). The blue bar seems to frame the response options and draw the user's eye initially to what is below the bar. We think that if people are able to figure out the question from the response options, they are not reading the question. If they do not understand what they are being asked from the options (e.g., in ancestry), they go ahead and skim the question. We realize that this hypothesis is counter to Tourangeau's (2000) model of survey response, but it is based on actual observation, not on what we think respondents should do.

Recommendation:  Our participants seemed to be fairly typical web users.  That is, they skimmed rather than read.  They first read the response categories and, if they thought they knew the question from the responses, they didn't read the question.  The dark blue horizontal bars may also be keeping people from reading the questions. We should keep these points in mind for future web development.

Priority:  Moderate for 2010
Team Decision:  This will be explored in future rounds of usability testing.


## 5.  Findings and Recommendations from Accessibility Testing with In-House Participants

 Accessibility testing took place with two Census employees during the same time period as usability testing.  One participant could not see the screen and was an expert user of a screen reader (JAWS).  She first tested the topic-based version and came back the next day to evaluate the person-based version.   The second participant described his condition as "severe tunnel vision."  We refer to him as "the low-vision user."  He customizes his screen but does not use JAWS.  We attempted to customize a monitor in the Usability Lab for his use but were not able to achieve the setting he is used to.  We therefore tested the topic-based version in his office with his usual settings:  a black background with yellow fonts and large text.  These settings dramatically changed the look of the Internet form.

The blind user tested both the topic-based and person-based versions of the questionnaire with the JAWS screen reader for two household members.  The low-vision user completed the topic-based version of the form for four household members.  The URL was inaccessible for the blind user and difficult to read for the low-vision user.  For both participants, the ID number on the printed form itself was inaccessible; someone needed to read it to the participants.  The low-vision user did not see the instructions on how to find the Census ID number.  He said he has a "rule" not to look at things on the right since they are usually advertisements.

In general, the difficulties these participants encountered matched the findings previously documented in the accessibility reports for the topic-based and person-based forms.  The blind participant commented that the person-based version was easier to use than the topic-based version since it did not have the labeling problems she encountered in the topic-based form.

Both versions of the form had the same accessibility issues with the skip links, ancestry, and review/submit screens.  Perhaps because she listened to everything (sometimes more than once), the blind user never triggered an edit message.  This is in contrast to the sighted users who read as little as possible.

*Item 18:  Skip links*

The skip link to bypass the top horizontal navigation bar is not programmed properly. The target of the skip link occurs before the first question, but the skip link itself is not present.

Recommendation:  The skip link should occur just prior to the "household" tab.

Priority: High
Resolution:  Corrected

*Item 19:  Hispanic Origin and Race (checkbox tables)*

The Hispanic and Race checkbox tables displayed similar traits.  Lack of labeling in the topic-based version made answering these questions very inefficient. The blind participant said that to answer either question, she had to remember the response options when tabbing to the checkboxes.  She needed to back up to the table header to hear the response options because no checkboxes were labeled.  When checking a box, JAWS would read the name of the next person or "table end... "checkbox checked".  The participant commented, "This is about as badly labeled as you can get."

The blind user remarked that she was able to figure out the Hispanic origin response because it was just had "yes" and "no" as responses.  She noted, however, that if she belonged to the third or fourth racial group, it would be very difficult to answer without back-tabbing and manually counting tab stops.

Recommendation:  Improve labeling of these checkbox tables.

Priority: High
Resolution:  Labeling has been improved.

*Item 20:  Ancestry*

When the blind participant encountered the ancestry data-entry boxes, she was not sure what to do.  She thought she was limited to three entries, one per box.  She said it would help if the fields had labels.

Recommendation:  Label the data-entry fields.  Indicate that more than one ancestry can be entered into each field.

Priority:  Moderate
Resolution:  The data-entry fields have been labeled.

*Item 21:  Review/Submit screen*

Some responses did not seem to make sense, such as "Origin:  no" and "Residence:  no"

 Recommendation:  Revise the labels for these elements as suggested previously in this report.

  Priority:  Moderate
  Resolution:  This has been done for Residence but not for Origin.


*Item 22:  Navigation bar*

The navigation bar in both versions consists of white, gray, and light blue tabs.  The low-vision user's customized settings did not apply to the display of these or any other images. If anything, they became virtually non-functional, isolated from the rest of the screen.  He never noticed the tabs because his software could not increase the font size.  When directed to them, he could not distinguish any difference between the tabs.   As a result, this participant was unable to quickly navigate the form, except by pressing the "previous" and "next" buttons.

Recommendation:  Research alternative methods to provide quick navigation for blind and low-vision users.

 Priority: Moderate for future versions
 Resolution:  Research is to be conducted during the second round of usability testing, time permitting.


*Item 23:  Graphics*

In both the person-based and topic-based versions, graphics were included as part of the "Before you start" screen and in help.  The low-vision user was not able to read any of the content of an image or the text call-outs because his software could not access this information.

 Recommendation:   Make image content and text call outs accessible.

  Priority: High
  Resolution:  This was resolved for the topic-based version.  Image content is now described below the image; so it was not necessary to provide text for the call-outs in the topic-based form.  However, the text below the image did not adequately describe all call-outs in the person-based version.  Use of LONGDESC was recommended and has been implemented successfully.

A second round of usability testing, focusing on alternative designs, occurred in October 2005 (Murphy, Ashenfelter, Ciochetto ,Agarwal, & Malakhoff, 2013).

## 6. References

Andre, A. D., & Wickens, C. D. (1995). When users want what's *not* best for them. *Ergonomics in Design*, 10-13.

Baddeley, A. (1990). *Human memory: Theory and practice*. Boston: Allyn and Bacon.

Freedom Scientific. (2004). *JAWS® 5.1*. St. Petersburg, FL: Freedom Scientific Blind/Low Vision Group (http://www.freedomscientific.com).

Murphy, B., Malakhoff, L., Coon, D., Brady, S., Charbonneau, D., Ciochetto, S., Fratino, S., Lins, J., Smith, M., Reiser, C., & Ross, A. (March 24, 2005). *2005 Census Internet Prototype Applications: Plan for Usability and Accessibility Testing*. Washington, DC: U. S. Census Bureau.

Murphy, E.D., Ashenfelter, K. T., Ciochetto, S.,Agarwal, A., Malakhoff, L.A. (2013). Usability Testing of Alternative Design Features for the 2005 National Census Test (NCT) Internet Form: Methods, Results, and Recommendations from Round 2 of Usability Testing. *Research and Methodology Directorate, Center for Survey Measurement Study Series (Survey Methodology #2013-11)*. U.S. Census Bureau. Available online at <http://www.census.gov/srd/papers/pdf/ssm2013-11.pdf>.

Noldus. (2003). *The Observer$^R$* (Reference Manual, Version 5.0). Wageningen, The Netherlands: Noldus Information Technology b.v.

Norman, K. L., Shneiderman, B., Harper, B., and Slaughter, L. (1998). *Questionnaire for User-Interface Satisfaction*. College Park, MD: University of Maryland, Human-Computer Interaction Laboratory.

Reason, J. (1990). *Human error*. New York: Cambridge University Press.

Refnes Data. (2005). W3schools: Browser statistics. Retrieved from http://www.w3schools.com/browsers/browsers_stats.asp on March 2, 2005.

Redish, J. (2012). *Letting Go of the Words* (2$^{nd}$ ed.). Waltham,MA: Elsevier/Morgan Kaufmann.

SSB Technologies. (2004). *InSight 4.0 /InFocus 4.0*. San Francisco, CA: SSB Technologies, Inc.

Theofanos, M. F., & Redish, J. (2003).  Guidelines for accessible – *and usable* – Web sites: Observing users who work with screenreaders.  *interactions*, *10* (6), 38-51.

Tourangeau, R., Rips, L. J., and Rasinkski, M. (2000). *The Psychology of Survey Response*. New York: Cambridge University Press.

U. S. Access Board (Architectural and Transportation Barriers Compliance Board).  (2000).  *Electronic and information technology accessibility standards*.  Washington, DC:  Federal Register (Dec. 21).  Retrieved January 15, 2005, from http://www.access.board.gov/secton508/508standards.htm

U. S. Census Bureau.  (2002).   IT Standard 20.0.0:  *Design requirements and guidelines for Web-based user interfaces*.  Washington, DC: Systems Support Division, Software and Standards Management Branch.

**Appendix A – Usability and Accessibility Goals**

*A.1  Overall Goals*

The team will consider the user interface successful if 80 percent of the participants are successful in achieving the overall goals.

1. Successful completion of the Form:  Respondents should be able to successfully complete and submit the form.  This includes successfully finding the URL for the site on the questionnaire.

2. Efficient and accurate data entry: Respondents should understand how data is entered and how it can be corrected.  Respondents should be able to easily identify the data needed and type entries and select options as needed.  Respondents should be able to correctly enter data the first time and/or deal with corrections of errors, inconsistencies, and blank entries following an edit message.

3. Efficient and effective navigation: Respondents should be able to quickly gain an understanding of how the survey is organized into a household page, person pages, and a review/submit page (Person-based) or pages by topic and a review/submit page (Topic-based).  They should also understand how the navigation works using buttons and tabs.  Respondents should be able to move through the form in a direct, forward manner initially and later, if needed, by jumping directly to previous pages.

4. Subjective satisfaction with the process of filling out the survey: Respondents should not feel frustrated or confused while filling out the survey but should be satisfied with the process, the interaction, and the time required to fill out the survey. When the respondent submits the form, he or she should be confident that it was duly submitted and received by the Census.  Finally, the respondent should come away with a good impression of the Census Bureau.

*A.2  Screen-Specific Goals*
1. Comprehension of the purpose of the screen.

2. Efficient and accurate entry of data.  Probe for why user doesn't provide data for a question, if not obvious during think aloud.  This includes observing entry of invalid data.

3. Successful exit and navigation using the "Next" button or other appropriate link.

4. Accessible for users with disabilities (i.e., compliant with Section 508).

**Appendix B – Consent Form**


The Census Bureau routinely tests products used for collecting information about the U.S. population in order to keep the country informed about changes and trends.

You have volunteered to take part in a study to evaluate a set of Web pages that may be used for collecting information about households during a planned National Census Test. This study will help the Census Bureau identify improvements that need to be made to the Web pages before they are used in the National Census Test.

In order to have a complete record of your comments, your interview session will be videotaped. We plan to use the tapes mainly to help us remember the details of your session. Staff involved in this design research will have access to the tapes. The tapes may also be used for training others to conduct this type of research and in presentations to Census Bureau staff. Quotations and video clips from your session may be used in the presentation and in our report. None of your personal information will be used in presentations, and you will not be identified by name. Your information is protected by Federal law.

This study is being conducted under the authority of Title 13 United States Code. The OMB control number for this study is 0607-0725. Respondents are not required to respond to any information collection unless it displays a valid approval number such as this one.

I have volunteered to participate in this Census Bureau product design study, and I give permission to be videotaped and for my tapes to be used for the purposes stated above.


_____          _____
  Researcher's Signature                      Participant's Signature



_____          _____
      Printed Name                               Printed Name



_____          _____
          Date                                       Date

**Appendix C – Questionnaire on Computer Use and Internet Experience**

1. Do you use a computer at home or at work or both?
   (Circle all that apply.)

   Home
   Work

2. If you have a computer at home,

   a. What kind of modem do you use at home?
      Dial up
      Cable
      Other _____
      Don't know _____

   b. Which browser do you typically use at home?  Please indicate the version if you can recall it.
      Firefox
      Internet Explorer
      Netscape
      Other _____
      Don't know _____

   c. What operating system does your browser run in?
      MAC OS
      Windows 95
      Windows 2000
      Windows XP
      Other _____
      Don't know _____

3. On average, about how many hours do you spend on the Internet per day?
   0      1-3     4-6     7+

4. Have you ever filled out a survey on the Internet?
      a.  If yes, about how many surveys do you think you have filled out on the Internet? _____

       b. If yes, have you filled out a survey on the Internet in the last two months?

            _____ Yes            _____ No

5. Please rate your overall experience with the following:

|           | no experience |   |   |   |   | very experienced |   |   |   |
|-----------|---|---|---|---|---|---|---|---|---|
| Computers | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Internet  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

**Appendix D – General Introduction**


Thanks for coming by today.  My name is _____, and I'll be working with you to evaluate a prototype of the online form for use in the next Census, in 2010.

First, I'd like to ask you to read and sign this consent form.  It explains the purpose of the testing and informs you that we would like to videotape the session, with your permission.  Only those of us connected with the project will review the tape.  We will use it mainly as a memory aid.  We may also use clips from the tape to illustrate key points about the design of the Web pages.

[Hand consent form; give time to read and sign; sign own name and date.]

Next, I'd like you to complete this questionnaire on your experience with computers and the Internet.  [give questionnaire and start the tape while the participant is working]

[Accept completed questionnaire.]  Thank you.

I want to emphasize that this is not a test of your skills or abilities.  You are helping us test a user-interface design. It is the Internet version of the Census short form. Your feedback is valuable, and we appreciate your help.

The Web pages that you will be evaluating represent one way of presenting questions from the Census short form on the Web.  While you are going through the form, I'm going to ask you to think aloud.  That is, I'd like you to read out loud anything that you are reading, explain why you are doing something, and talk to me about anything that raises a question in your mind. Tell me about your thought process and your method of navigating. If you expect something to happen, tell me what you expect.

As a practice in thinking aloud, I'm going to ask you to find the local weather forecast on the Web.  Use whatever method you would usually use, but keep talking to me about your strategy for finding the forecast.

[time for practice in thinking aloud]

OK, during the session, I will remind you to talk to me if you get quiet.  For example, I'll ask you to tell me what you're thinking.

Do you have any questions?

When you get to the last page of the Web form, please do not press the submit button to send your data.  We will be going back through the form and talking about your experience.  If you submit your data, the form will be cleared, but it would be better if it

still had the data you entered.  I will ask you to submit your data when we finish our walkthrough of your completed form.

OK, now, we'll have you get started with the first screen.

--------------------------------------------------------------------------

[Probe only to have them think aloud – no specific probing – keep it general]

When they get to the Review/Submit screen, remind them not to submit and ask, "How long do you think that took?"

During the retrospective portion, cover the debriefing questions.

After the retrospective debriefing, ask if they want to comment on anything we did not talk about.

Finally, have them complete an invoice and give them an envelope.
[Invoice requires name, address, and signature.]

**Appendix E – Questionnaire for User Interaction Satisfaction (QUIS)**

Please circle the numbers that most appropriately reflect your impressions about using this electronic questionnaire.

| | | | | |
|---|---|---|---|---|
| 1. | Overall reaction to the electronic questionnaire: | terrible | | wonderful |
| | | 1  2   3   4   5   6   7 | | 8   9   not applicable |

1.  Overall reaction to the electronic questionnaire:  terrible  ... wonderful
1 2 3 4 5 6 7 8 9 not applicable

2.  Getting started:  difficult  easy
1 2 3 4 5 6 7 8 9 not applicable

3.  Screen layouts:  confusing  clear
1 2 3 4 5 6 7 8 9 not applicable

4.  Information displayed on the screens:  inadequate  adequate
1 2 3 4 5 6 7 8 9 not applicable

5.  Arrangement of information on the screens:  illogical  logical
1 2 3 4 5 6 7 8 9 not applicable

6.  Order of presenting questions:  confusing  clear
1 2 3 4 5 6 7 8 9 not applicable

7.  Going back to previous questions:  impossible  easy
1 2 3 4 5 6 7 8 9 not applicable

8.  Making changes to answers:  difficult  easy
1 2 3 4 5 6 7 8 9 not applicable

9.  Use of terminology throughout the questionnaire:  inconsistent  consistent
1 2 3 4 5 6 7 8 9 not applicable

10.  Adequacy of instructions:  inadequate  adequate
1 2 3 4 5 6 7 8 9 not applicable

11.  Accessing help:  difficult  easy
1 2 3 4 5 6 7 8 9 not applicable

12.  Overall experience of entering information:  difficult  easy
1 2 3 4 5 6 7 8 9 not applicable

13.  Progression through the questionnaire:  slow  fast
1 2 3 4 5 6 7 8 9 not applicable

14.  Length of time to complete the form:  seemed very long  seemed very short
1 2 3 4 5 6 7 8 9 not applicable

**Appendix F – Debriefing Questions**

The debriefing questions will key off the results from the Questionnaire for User-Interface Satisfaction (QUIS). Those items that are low will be asked about during the debriefing. If any of the events occur that are in listed under our observation areas then debriefing questions will be asked about them. Any probes based on observation of the participant will also be part of the debriefing. The areas we want to probe about are listed below:

1. How easy was it to find the URL? What steps did you take to find the website?
2. Allow for a general question of all users that asks if they noticed the tabs, color indicators, etc. How we discuss these results in the usability testing report is important. If a user does not notice the tabs, etc. and they are able to navigate through the form using the "Next" and "Previous" buttons without any difficulties then the fact that they did not notice the alternative navigational tools is not an issue. The probing questions will need to differ between the topic- and person-based designs.
3. Observe users' reactions and ability to understand the multiple ways of accessing help (via examples, question mark icon, include/exclude links or help link) and probe during retrospective. We will probe with a general question about if the user needed help where would they go and we will probe those users who have problems and need help if they noticed the help links.
4. We will probe during the retrospective with a general question about if the user thought it was okay if someone marked more than one race.
5. What does the "Thank You" screen mean to you?

**Appendix G – Change Scenarios**

Once the participant has reviewed his or her data on the Review/Summary Screen, we want to have him or her make at least one change in the entered data to assess the extent to which the design supports the respondent's expectations. We suggest the following scheme for having the participant make changes in his or her data:

- If participant makes changes to data based on his/her own review, not necessary to go through any scenarios.
- If participant verbalized uncertainty as to who should/should not be counted while providing the data, have the participant change the number of people in the household.
- If, while providing the data, the participant verbalized uncertainty as to whether anyone in the household is Hispanic, Latino, or Spanish, ask the participant to change origin for that person.
- If the participant verbalized uncertainty about anyone's race/ancestry while providing the data, ask the participant to change the person's race/ancestry.
- In a person-based interview with a large household (several children), have the participant change a child's date of birth.
- In a household with unrelated individuals where the participant doesn't know the date of birth for someone in the household but guesses at his/her age, have the participant change the age of that person.
- If none of the above scenarios apply, have the participant change the telephone number for person 1 (or enter one if it was left blank).

**Appendix H – Observation Guide for Test Administrators**

Note: This guide will be reviewed and updated, if necessary, prior to testing in order to adjust for any changes that have occurred in survey content since the plan's last revision.

The following are key observation areas that support the overall goals and the screen specific goals for the usability testing. These observation areas are called out for test administrators to pay special attention to. These observation areas could lead to probes during the retrospective.

Observation/Probes pertaining to general application features:

➢ Observe users' reactions to navigation, color, tabs (do they understand what they are for) and then probe during retrospective. This includes observing if users provide data topic by topic in the person-based design or person by person in the topic-based design. For example on sex/DOB/age screen do users go down column (topic) or across row (person)?

➢ Observe users' reactions to the edit messages and their ability to handle them and then probe during retrospective. Specifically in the topic-based design, observe if users realize that they need to scroll down to where errors occurred in order to correct them.

➢ Observe users' reactions and ability to understand the multiple ways of accessing help (via examples, question mark icon, or help link).

Observations pertaining to design of screens:

➢ Do we have sufficient contrast between the text and the actual button on the "Next" and "Previous" buttons? *(Topic & Person)*

➢ Before You Start Screen-
  • Observe users' behavior with the screen. Observe if they are reading it, etc. *(Topic & Person)*

➢ Household Screen-
  • Observe users' interactions with the Include/Exclude instruction links in the residence rules. If a user has difficulty determining whom to include in the POP count, probe during the retrospective about using the links to get information. *(Topic & Person)*
  • Observe users' reactions to the same last name checkbox. Do they notice the question and understand what it means. Observe if users have problems with the name for Person 1. *(Topic-based specific)*
  • Observe users' interactions with the POP Count question. Do they understand that the "people" label means "Number of people" and that they actually need to fill it in? *(Topic & Person)*

➢ Observe users' reactions for one person households when tabs disappear. Do they notice or does it confuse them? *(Topic-based specific)*

➢ Name Screen- Observe users' reactions to the pre-populated last name and Person 1's name. Observe if users include themselves in the list of names. *(Topic-based specific)*

➢ Sex/DOB/Age Screen-
- Observe users' reactions to the age calculation. *(Topic & Person)*
- Do users see all items on this screen? *(Topic-based specific)*
- Observe users' reactions to the mm/dd/yyyy format for DOB. *(Topic & Person)*

➢ Origin Screen-
- Observe users' reactions to the grid format of question. *(Topic-based specific)*

➢ Race Screen-
- Observe users' reactions to the grid format of question. Are they able to use the grid correctly? *(Topic-based specific)*
- Observe users' reactions to the examples links. Are users overwhelmed by the number of examples on the screen? *(Topic & Person)*
- Are respondents understanding that they can report more than one race; do they readily report more than one race? We will probe during the retrospective with a general question about if the user thought it was okay if someone marked more than one race. *(Topic-based specific)*

➢ Review/Submit Screen-
- Do users understand the navigation options? *(Topic & Person)*
- Do users successfully make edits? Test this by giving them edits to make using their own data. *(Topic & Person)*

Observations pertaining to content of screens:

➢ Household Screen-
- Do users understand or are they confused by the sentence "We will count them at the other place" below the exclude link? Do they know what "other place" means? *(Topic & Person)*
- Is the "start with" instruction confusing in the question for Person 1's name? *(Topic-based specific)*

➢ Sex/DOB/Age Screen-
- Observe users' reactions to the instruction for babies. Do people with babies under the age of 1, report correctly? *(Topic & Person)*

- Observe any age and date of birth discrepancies (the chance of this occurring is rare since age is calculated). *(Topic & Person)*

➢ Race Screen-
- Are respondents finding the ordering of the race response categories confusing?  Are there specific comments about the race response categories? *(Topic &Person)*
- Do users understand that they can report more than one race; do they readily report more than one race? *(Topic &Person)*

➢ Ancestry Screen-
- Observe users' reactions to the three boxes for write-ins.  Are they confused by the three separate boxes?  Do they understand that they do not need to put something in each box?  Do respondents understand they can report more than one ancestry or tribe? *(Topic & Person)*
- Are respondents providing detailed i.e., White Mountain Apache, or general i.e., Apache, responses to the question on ancestry or tribe? *(Topic &Person)*

**Appendix I – Behavior Codes for Logging**


**Table 7   Behavior Codes for Logging Test Sessions**

| Behavior Code | What code means… |
|---|---|
| Answer change (A) | Participant changed their answer (including calculated age).  They could have changed the answer because of an edit or another reason. |
| Error due to overlooked information (E) | There was something on the screen that could have helped the participant provide an answer but they don't notice and enter inaccurate data (final data entered).  Logger has to know that the data are inaccurate. |
| Field cursor/tabbing problem (F) | Participant has a problem either tabbing or moving between fields on a screen. |
| Confusion (C) | Participant is confused about the screen.  Logger should add comments about the cause of the confusion.  Is it the design of the screen, content of the question, confusion about the help contents-contents were not what they were looking for? |
| Screen navigation problem (S) | Participant has problems navigating between screens.  This includes using the browser back button. |
| Need assistance from administrator to continue (N) | Note that administrators assist as a last resort.  Thus, other behavior codes may be logged for this situation. |
| Left blank-Design  (D) | Participant left an item blank because of a screen-design issue, such as they didn't see item or they didn't understand the item because of the design. |
| Left blank- Other  (O) | Participant left an item blank because they don't know the answer or they refuse.  Regardless of the design, participant would have left the item blank. |
| Technical/System issue (T) | A server error occurs, etc. |
| Negative response to edit message (R) | Participant is served an edit message and they don't understand the edit message, don't see it, etc. |
| Comment  (M) | Participant or test administrator has a comment.  (Logger will distinguish between participant and administrator.) |
| Use of links (e.g., help) (L) | Participant clicks on a hyperlink.  (Logger will indicate the specific link selected.)  Clicking on a hyperlink indicates the participant understood it as a link; but not clicking on a hyperlink does not mean it was not understood to be a link. |

**Appendix J.  Findings and Recommendations from Tool-based Accessibility Testing**

Three rounds of tool-based accessibility testing were conducted using JAWS and/or InSight/InFocus.  At the time of delivering the draft version of this final report, all but one of the accessibility (Section 508) violations had been corrected, and the last one was near resolution.

**Final Findings, 5/2/05**

*J.1  Person-Based Internet Application*

by David Coon, Decennial Management Division

References in this document to "paragraph rules" apply to paragraph letters of § 1194.22, Section 508 (29 United States Code 794d)

Global Issues that apply to all screens:

Larry's report on the topic-based version also speaks to some of these global issues in detail.

**Skip Links:**

**Priority:** HIGH, required by Paragraph (o) rule.

There is an attempt to provide a skip link, but the link has no target.  The application should have a skip link for the vertical navigation bar that starts immediately after the Privacy Policy Link and ends immediately before the tabs bar.
The Navigation bar skip link should be read by JAWS as "Press Enter to skip the left navigation bar."

The Left navigation bar has no title block to indicate that a navigation bar is present. It could be confusing to respondents when JAWS reads the links without any context.

The first screen that JAWS detects the Navigation bar is on the Origin screen.  It should be able to be detected on earlier screens.

There should be a skip link for the tabs bar that starts to the left of the tabs bar and ends after the tabs bar.  There is currently a placeholder for the skip link at the end of the tabs bar, but it has no target.
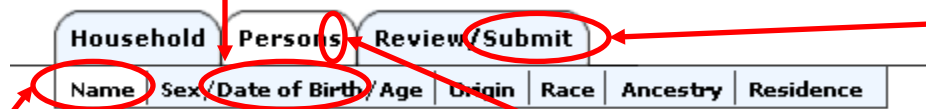
The Navigation Tabs have typos in the alt tags.

The ALT tag for PERSON should say "Persons" (missing "s")
The ALT tag for the NAME tab says "Names" (should say "Name").
The ALT tag for the Review/Submit tag misspells "Subumit"
The ALT tag for "dateofbirth" should contain spaces between the words "date of birth"

The visited navigation tabs must have ALT text indicating the state of the tabs, i.e. that they are "visited" tabs.

The unavailable navigation tabs must have ALT tags indicating the state of the tabs, i.e. that they are "unavailable" tabs.

Z-Tech verbally agreed to provide ALT text for these tabs when we met to discuss their Section 508 checklist.

**Form Controls**

A strict interpretation of the paragraph (n) rule would mandate that all form controls (radio buttons, check boxes, and text entry fields) must have an explicit label.

This document calls out a priority for certain fields that are not intuitive with JAWS. Technically, strict compliance with Section 508 would mandate that all form controls have an explicit label.

Testing without JavaScript

The paragraph (l) rule states that if scripts (e.g. JavaScript) are required, the system still must be useable if scripting is disabled or unavailable.

I tested both the topic and person based forms without JavaScript disabled, and everything seemed usable. The Yes/No radio checkbox in the coverage question doesn't trigger without JavaScript, but a server-side edit makes the change, so it has the "equally functional alternative" required by the paragraph (l) rule.

Without JavaScript, the help still pops up in a separate window (though the window is full sized, unlike the small JavaScript window).

The age autocalculation does not work, without JavaScript, and has no equally functional alternative, but we can live with this, per the access board response cited in the age screen discussion.

**Cascading Style Sheets**

Per the paragraph (d) rule, the application must be readable without requiring style sheets.

**Finding:** JAWS had some limited difficulty reading the screens properly when style sheets were disabled.  Some of the form controls could not pick up the associated labels when stylesheets are disabled.

**Priority:** Low. The intent of this rule is for color blind or low vision individuals to use their own style sheet.  The application works ok with a number of standard "user defined" style sheets.  It only behaves strangely if style sheets are disabled or not supported, and it still is useable with disabled stylesheets, but the interface is awkward.  If the missing explicit labels are appended to form control, then this problem should be remedied.



This image may require a LONGDESC tag.  It communicates lots of information that is not in the ALT tag.
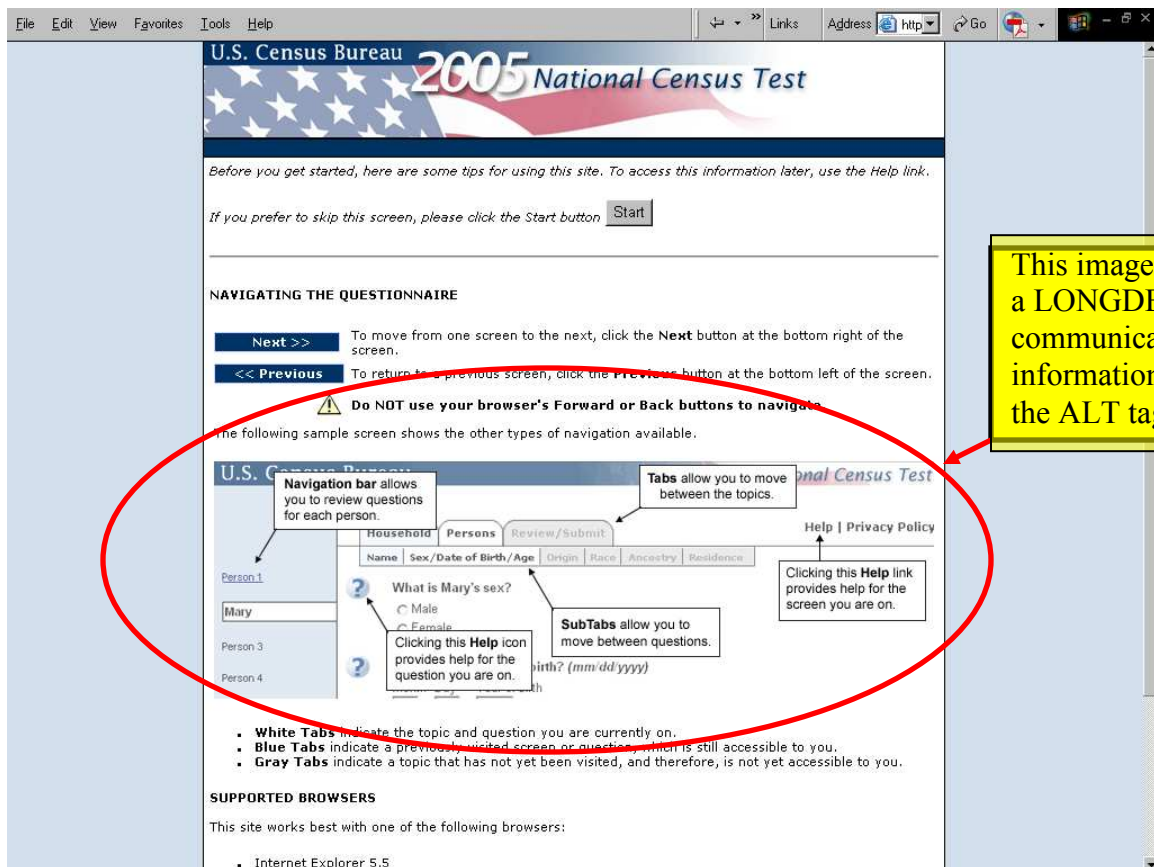
Finding :    **1. The image on this screen communicates lots of information to the user that is otherwise not communicated by the ALT tag.**

**2. The Start button requires an explicit label, per the Paragraph (n) rule.**

<u>Priority:</u> Medium to High

<u>Recommendation:</u>

1. A LONGDESC tag could be included that describes the callouts for the vertical navigation bar and horizontal tabs in plain text. This image is less problematic on the Topic Based version because that version does not have subtabs or a vertical navigation bar.

Household Screen
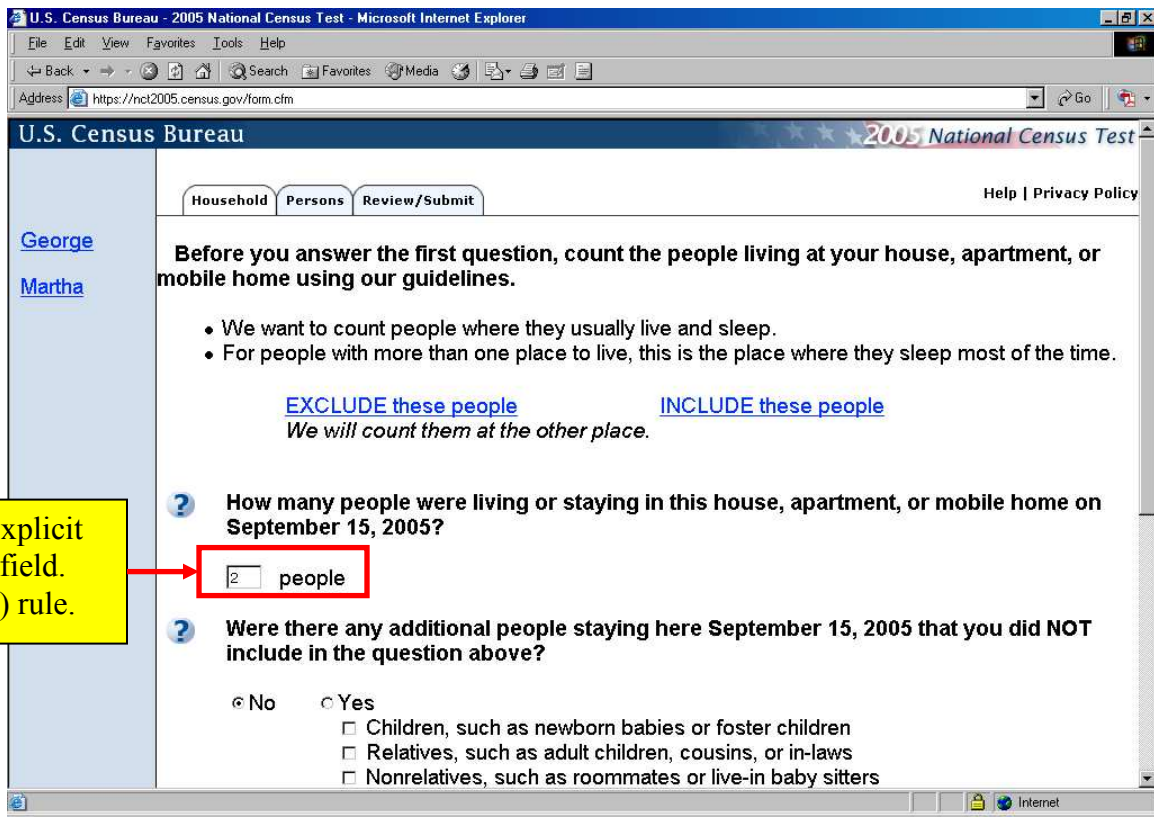


<u>Finding</u> : There is no explicit label on the Number of People Field paragraph (n) rule.

<u>Priority</u>: Medium, This is still usable but the interface is awkward without the explicit label.

<u>Priority:</u> Medium to High

**Recommendation:**

1. A LONGDESC tag could be included that describes the callouts
for the vertical navigation bar and horizontal tabs in plain text. This image is less
problematic on the topic-based version because that version does not have subtabs or a
vertical navigation bar.



**Finding:** The text fields must have an explicit identifier associated with the text field.
JAWS cannot distinguish which field is for FIRST NAME versus MI versus LAST
NAME.

**Priority:** High, required by Paragraph (n) rule.

**Finding:** The text fields for phone number lack an explicit identifier and technically
should have an explicit identifier associated with the text field.

**Priority:** Low.  See callout in picture above.

Sex/DOB/Age Screen:

**What is George's date of birth?**

Month  Day  Year of birth

`9` / `9` / `1950`

**What is George's age?**

Age (on 09/15/2005)

`50` years

> JavaScript autocalculation of age is acceptable, per the Access Board ruling cited below.

No problems identified  (except for global issues noted above).

Although JavaScript is used to perform an auto-calculation of age, In response to a comment on the notice of proposed rulemaking  The Architectural and Transportation Barriers Compliance Board (known as the Access Board) stated the following:

> If "a web page uses a script only to fill the contents of an HTML form with basic default values, the web page will likely comply with this requirement, as the text inserted into the form by the script may be readable by a screen reader ."

In keeping with this ruling by the access board, it is our recommendation that this does not present a problem, even though there is no equally functional alternative, since users can still access the form—they just don't have the added benefit of the pre-filled age.

Hispanic Origin Screen:

No problems identified (except for global issues noted above).
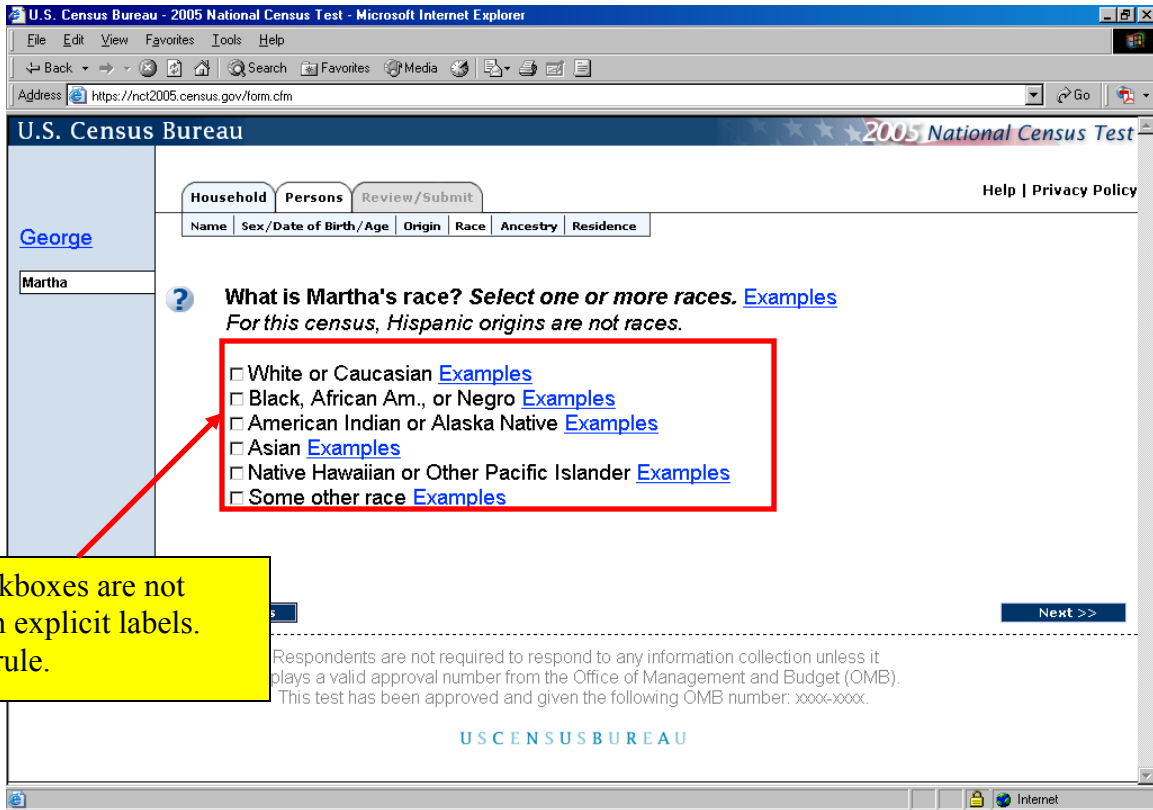
Race Screen:

**Finding:** The interface is awkward because the labels for each checkbox appear in a separate table cell of a nested table. While less than desirable, the current design will still work. Under the current design, it is difficult to tell which checkbox is associated with each racial group.

**Priority:** Low. Ideally, there should be an explicit label for each checkbox, per the paragraph (n) rule.

There is some kind of embedded help link in addition to the Question mark icon Link.
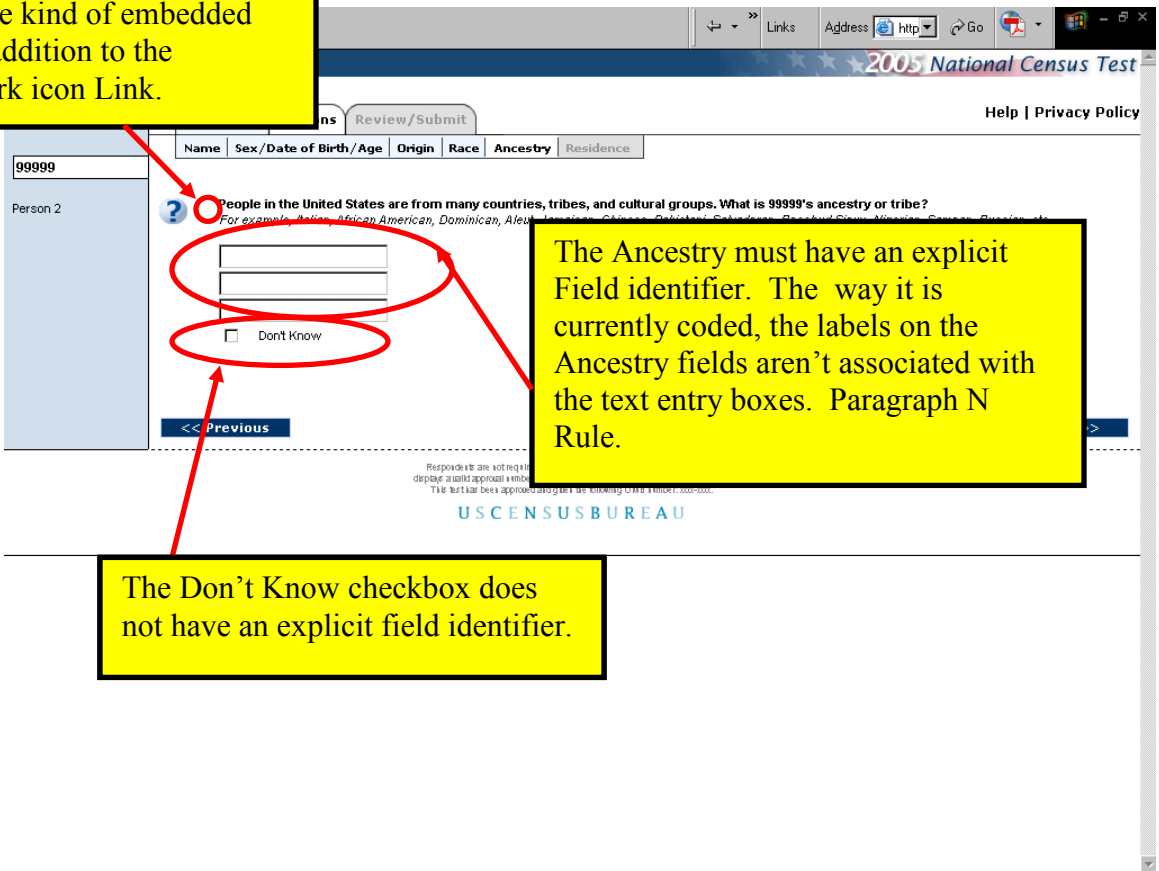
The Ancestry must have an explicit Field identifier. The way it is currently coded, the labels on the Ancestry fields aren't associated with the text entry boxes. Paragraph N Rule.

The Don't Know checkbox does not have an explicit field identifier.

**Finding:** There is an extra help link discovered by JAWS that is embedded within the page that is read between the main question and the italicized instruction. This link is in addition to the question mark help icon. This link did not appear on any other person-based screens.

**Priority:** Medium, while confusing to respondents, this should not be considered a show stopper.

**Finding:** The text fields for Ancestry must have an explicit identifier associated with the text field. The purpose of the fields is unknown by JAWS since they are so far away from the question text.

**Priority:** High, required by Paragraph N rule.

**Finding:** The "don't know" checkbox does not have an explicit label that can be detected by JAWS. JAWS reads the label associated with this box only when the box is checked.

**Priority:** Medium to High. The interface can still be used with JAWS, but is awkward.

**Coverage Screen**

**Finding:** There are no explicit labels associated with the coverage question as required by paragraph (n).

**Priority:** Low. This question seems to vocalize correctly in JAWS even without the explicit labels.



**Person 2 Name Screen:**
**Finding:** The text fields must have an explicit identifier associated with the text field. JAWS cannot distinguish which field is for FIRST NAME versus MI versus LAST NAME. They

**Priority:** High, required by Paragraph (n) rule.

Relationship works OK. The interface is a bit awkward, but it is usable with JAWS.

There is no ALT tag associated with this image.
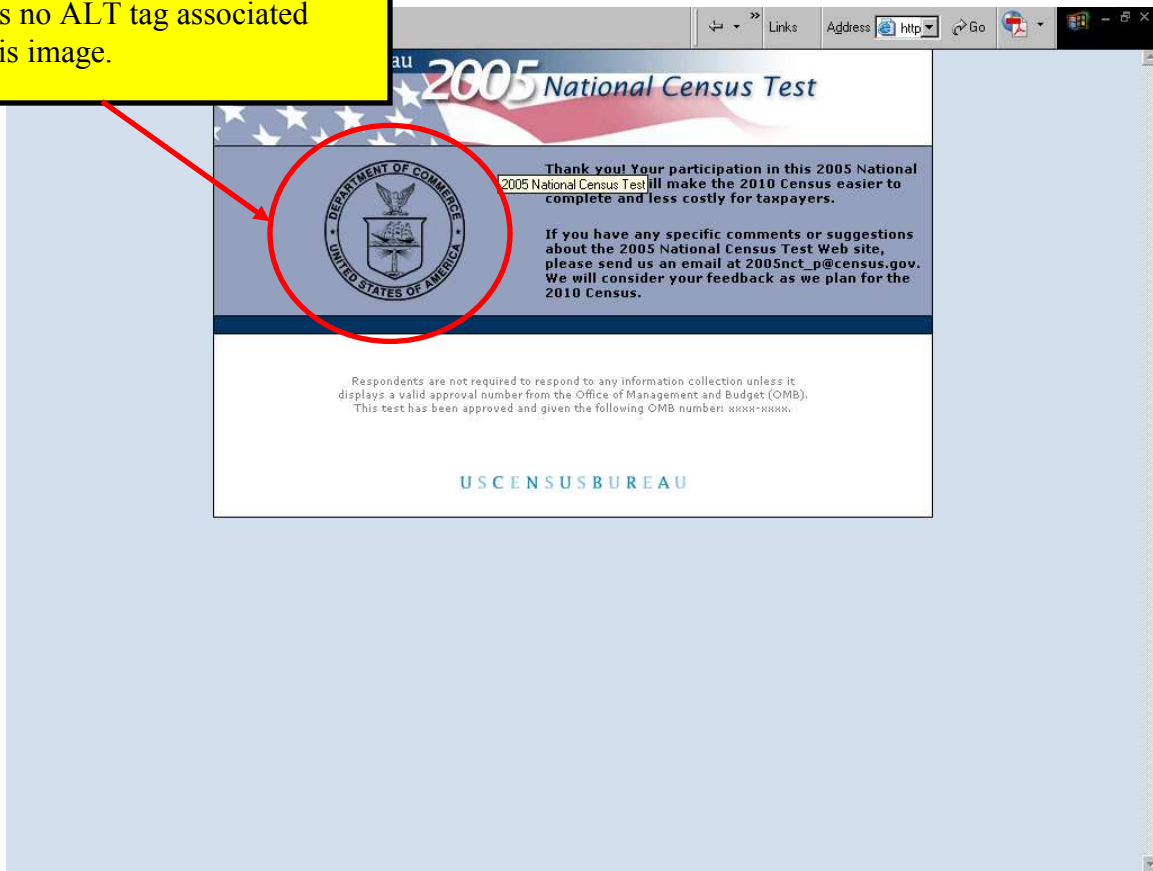
**Thank You Screen**

**Finding:** There is no ALT tag on the Department of Commerce Seal (welcome5.gif).

**Priority:** High

**Recommendation:** Add an ALT tag to this image that says "Department of Commerce Seal."

There are no ALT tags associated with these images.

Your session will time out in 5 minutes. If you need more time click, "continue" to proceed.

Continue >>

USCENSUSBUREAU



There are no ALT tags associated with these images.

Your session has timed out. To start over, you must close this window, reopen your browser, and reenter the URL for the 2005 National Census Test.

USCENSUSBUREAU

**Timeout screens**

**Finding:** There are no ALT tags on either timeout screen (see above).

**Priority:** High (paragraph a rule).

**Recommendation:** Add descriptive ALT tags to these screens for all images.

## J.2 An Accessibility Evaluation of the Topic Based version of the 2005 National Census Test Internet Form

Figure 1. The opening screen

Figure 2.  Information about tab color was not accessible.
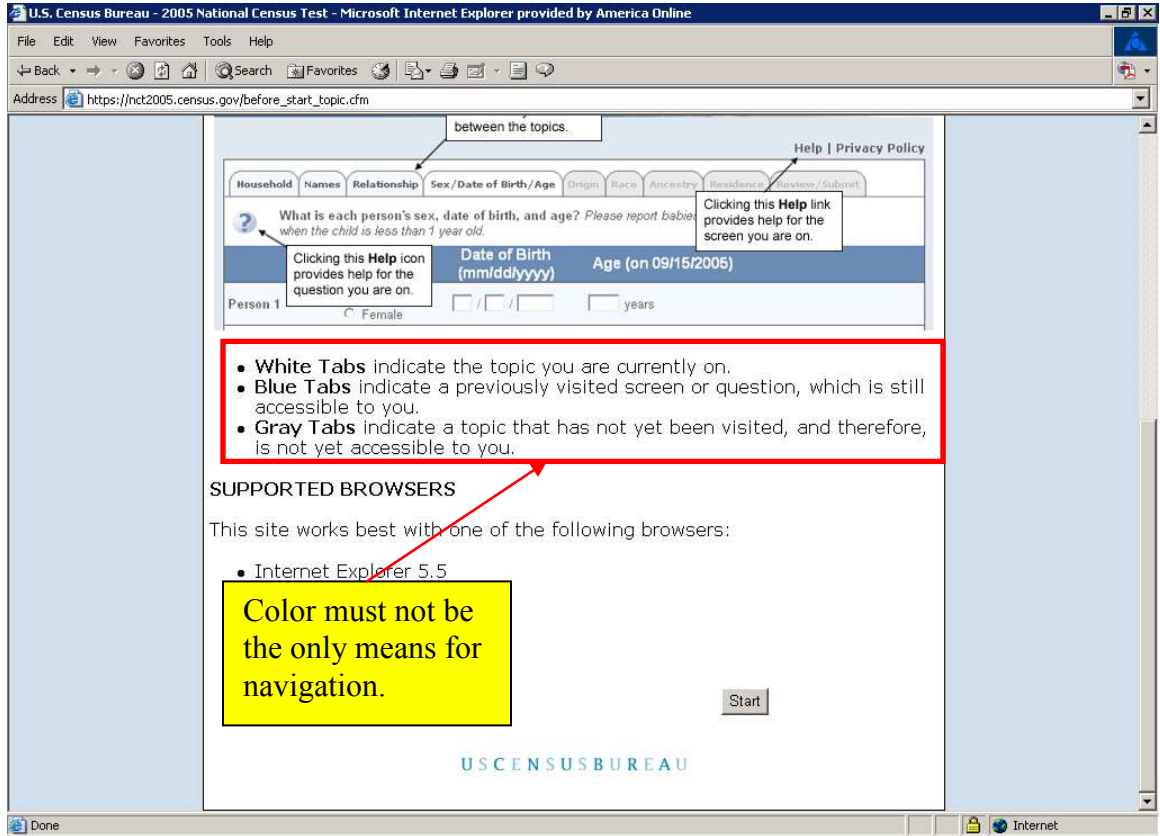


**Finding 1**:  **Blind individuals and persons with vision deficits will not be able to distinguish between white, blue and gray tabs.**
Priority:  High

**Details**:  When tabbing through the interface, there was no information provided about tab links, other than the text.

Recommendation:    The screen reader user should be able to hear "visited link" for tabs used earlier, "unavailable" for unvisited tabs, and only the tab text for the current tab. This language should be included on this page and the help screen.

Figure 3. **(Global)** Text displayed did not match the spoken text.
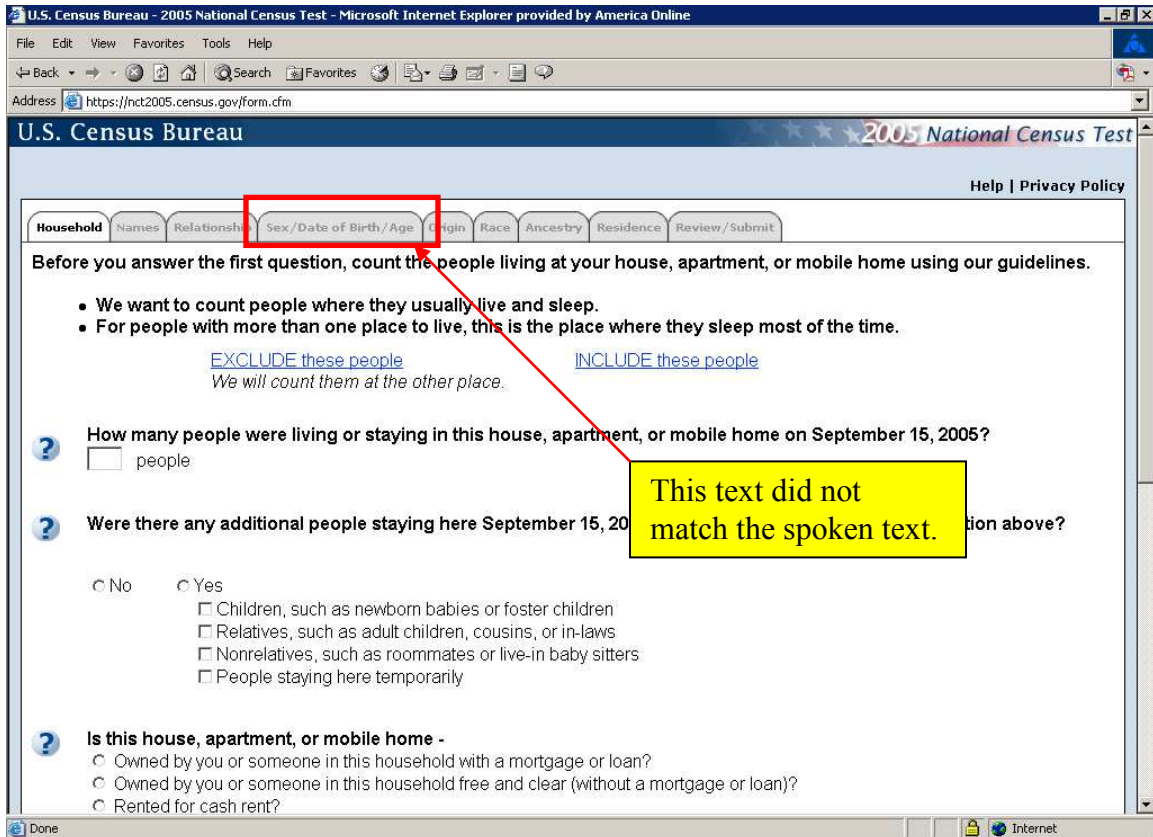


**Finding 2**:  **Tab text for "Sex/ Date of Birth/Age" did not match the spoken text and was not understandable.**
Priority:  High

**Details**:  The screen reader attempted to pronounce "Date of Birth" as one word.  When the source code was examined, the entry was keyed in as "dateofbirth", without any spaces between the words.

Recommendation:  Put a space between each word so it is read properly.

Figure 4.  **(Global)** There was no skip link to bypass the tab link navigation.
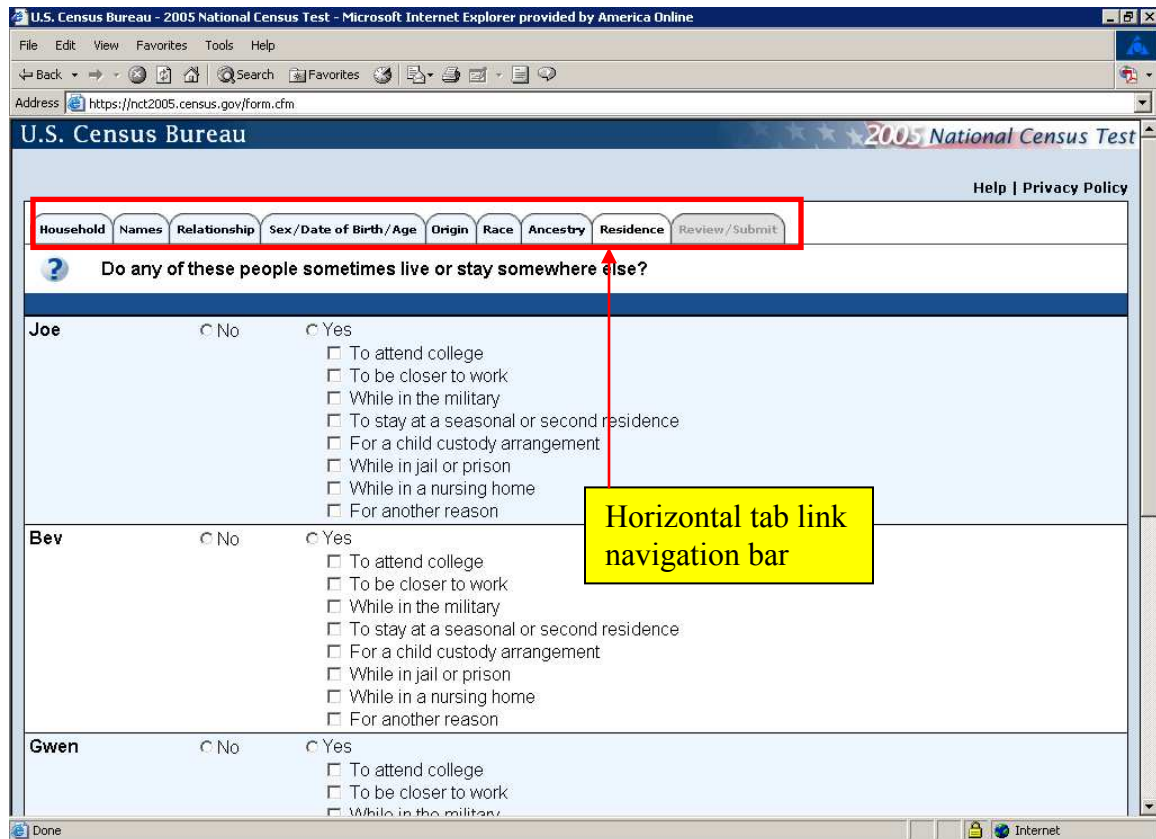


**Finding 3**:  **A skip link is needed for repetitive elements, such as navigation bars.**
Priority:  High

**Details**:  Blind and low vision users are entitled access equivalent to what a sighted user has available.

Recommendation:    Place a transparent graphical link preceding "Household" in tab order.  The target of the link follows "Review/Submit" in tab order.

Figure 5.  Checkboxes on the Hispanic origin page were inaccessible.
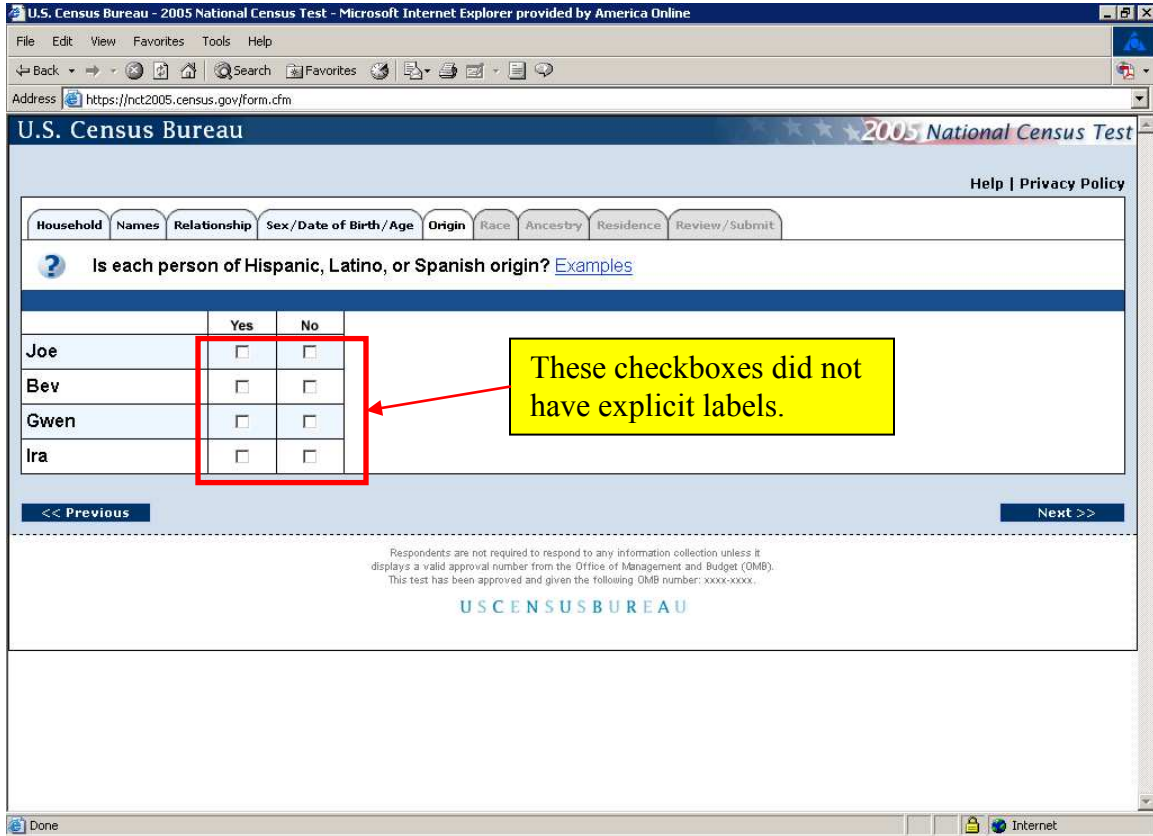


**Finding 4**:  **There were no labels for the checkboxes.**
Priority:  High

**Details**:  As the user tabbed through the checkbox table, no labels were detected.

Recommendation:    For each person, associate "yes" with the first checkbox, and "no" with the second checkbox.

Figure 6.  Checkboxes on the Race page were inaccessible.
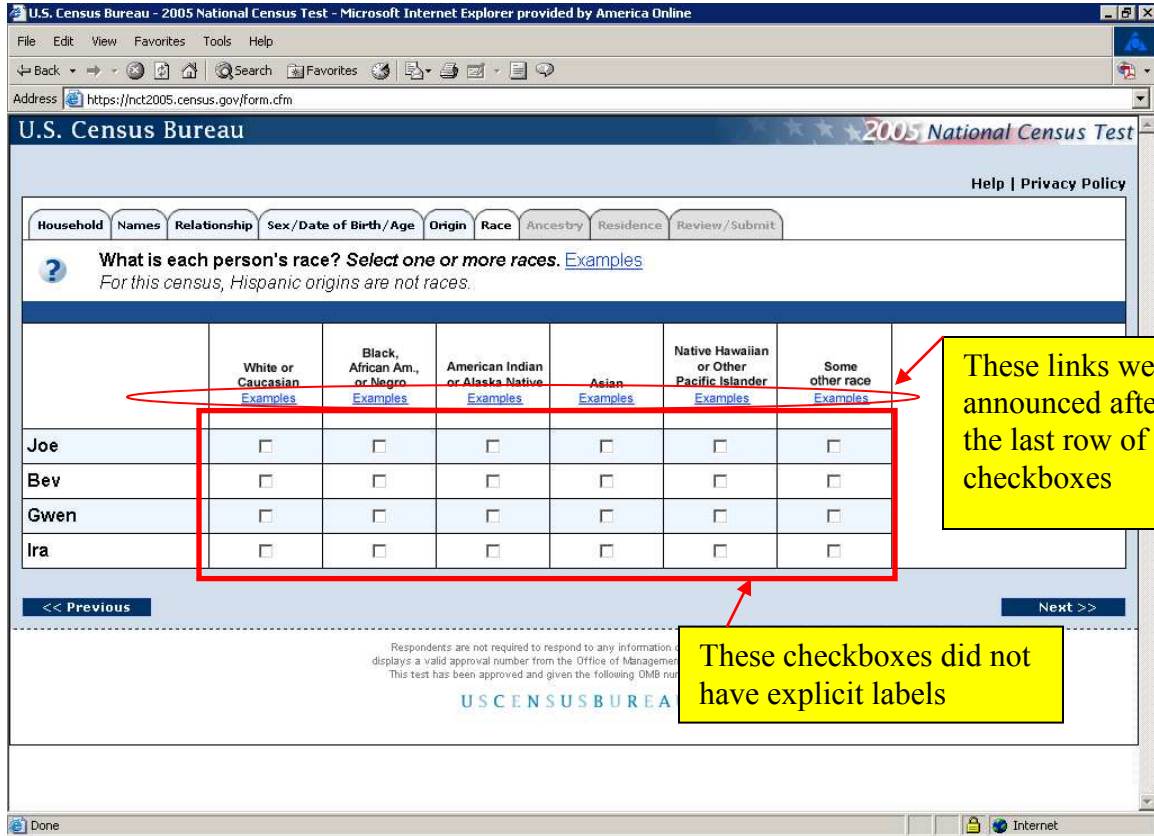


**Finding 5**:  **There were no labels for the checkboxes.**
Priority:  High

**Details**:  As the user tabbed through the checkbox table, no labels were detected.

Recommendation:    For each person, associate the appropriate race group name with the corresponding checkbox.

**Finding 6**:  **The visual order of the race link examples did not match the tab order.**
Priority:  High

**Details**:  As the user tabbed through the checkbox table, the examples links were not heard until after the last checkbox.

Recommendation:    For each race group, the example link should immediately follow the race group name.

Figure 7. Data entry fields for the household roster did not have explicit labels.



**Finding 7**: **There were no labels for the data entry fields.**
Priority: High

**Details**: As the user tabbed through the roster table, no labels were detected.

Recommendation: For each person, associate "First Name", "MI", and "Last Name" with the appropriate data entry fields.

Figure 8. "Related" and "Non Related" were not associated with their corresponding relationship category.
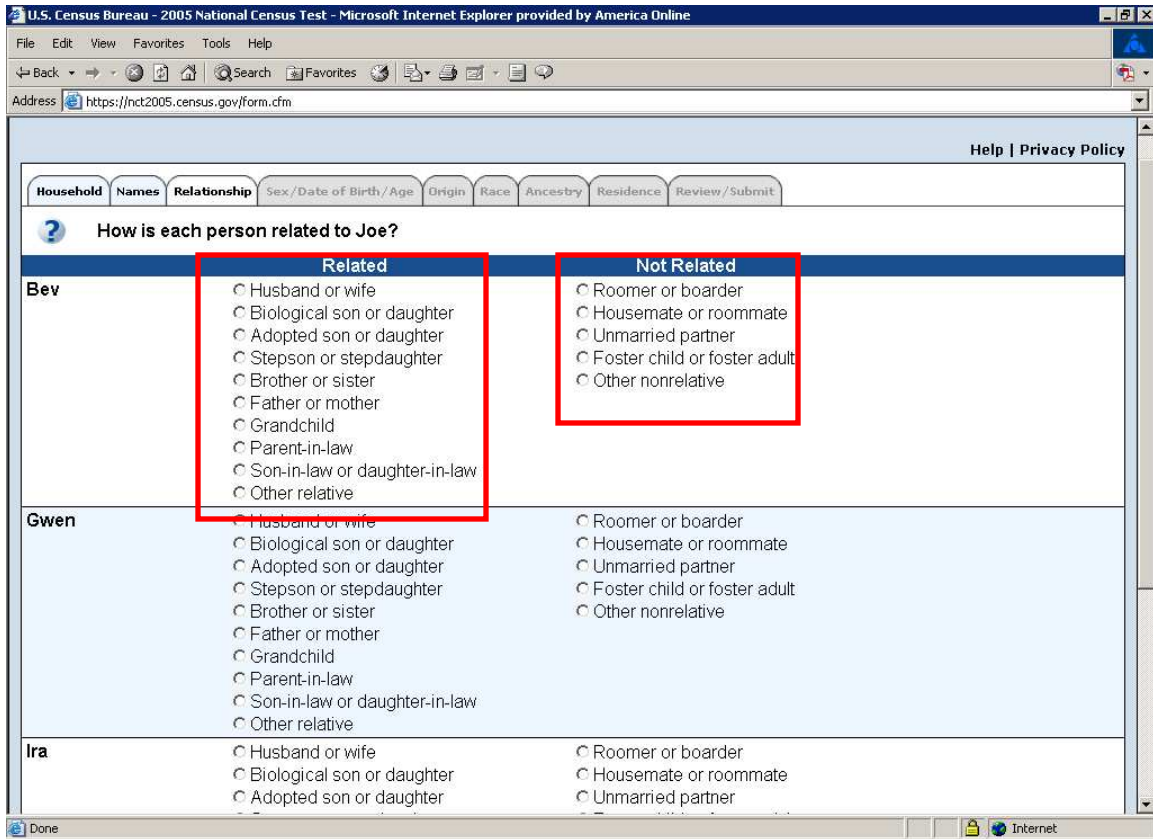


**Finding 8**: **The relationship sub-types were not associated with their corresponding relationship.**
Priority: Low

**Details**: As the user down-arrowed through the relationship options, the relationship sub-type was not announced. Even though this was rated as low priority it is mentioned because the relationship subtypes work in the person-based version of this form.

Recommendation: Modify the functioning person-based code and apply it here.

Figure 9. Date of birth and age data entry fields were inaccessible.



**Finding 9: There were no labels for month, day, year, and age data entry fields.**
Priority: High

**Details**: As the user tabbed through the table, no labels were detected.

Recommendation: For each person, associate month, day, year, and age with its corresponding data entry field.

Figure 10.  Ancestry data fields require explicit labels.



**Finding 10**:  **There were no labels for the three ancestry data fields.**
Priority:  High

**Details**:  As the user tabbed through the table, no labels were detected.

Recommendation:    For each person, associate a label, as shown above, with each ancestry data entry field.

**Finding 11**:  **The labels for the "Don't Know" checkboxes were not associated.**
Priority:  High

**Details**:  As the user tabbed through the screen, the label was heard after "blank".

Recommendation:    For each person, associate a label directly with the checkbox.

Figure 11.  The labels for the owner's name were not associated with the corresponding data entry field.



**Finding 12:  The labels for First Name, MI, and Last Name data entry fields were inaccessible.**
Priority:  High

**Details**:  As the user tabbed through the table, no labels were detected.

Recommendation*:    For each person, associate a label, as shown above, with each ancestry data entry field.

Figure 12.  This error message appeared after requesting more time.