

STUDY SERIES
(*Survey Methodology #2013-11*)

**Usability Testing of Alternative Design Features
for the 2005 National Census Test (NCT)
Internet Form: Methods, Results, and
Recommendations of Round-2 Testing**

Elizabeth Murphy¹
Kathleen Ashenfelter
Susan Ciochetto¹
Anshu Agarwal¹
Lawrence A. Malakhoff

¹ Former U.S. Census Bureau Employee

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

Report Issued: April 24, 2013

Disclaimer: This report is released to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.

Usability Testing of Alternative Design Features for the 2005 National Census Test (NCT) Internet Form: Methods, Results, and Recommendations of Round-2 Testing

Abstract: In preparation for the 2005 National Census Test and to inform future decision making on design issues, the Census Bureau conducted a second round of usability testing to evaluate the ease of gaining access to the online version of the Census Short Form, completing the form, responding to edit messages, and reviewing the entered data and making changes. This second round was conducted from November 30 to December 14, 2005. The key findings by screen, included participants having difficulty finding the Census ID number to log in, participants not reading the residence rules, and participants not reading the edit messages and being able to navigate from edit messages to correct the data.

Executive Summary

In preparation for the 2005 National Census Test and to inform future decision making on design issues, the Census Bureau conducted a second round of usability testing to evaluate the ease of gaining access to the online version of the Census Short Form, completing the form, responding to edit messages, and reviewing the entered data and making changes. This second round was conducted from November 30 to December 14, 2005. The key findings by screen, included participants having difficulty finding the Census ID number to log in, participants not reading the residence rules, and participants not reading the edit messages and being able to navigate from edit messages to correct the data. A list of all the usability issues that were identified is contained in Table 1.

Table 1: Summary of Prioritized Round-2 Usability Issues by Screen or Function

Screen/Function Name	Usability Issue	Priority
Housing Unit Identification (HUID) Screen	Difficulty of finding the Census ID Number	High
	Confusion in interpreting the graphic showing the Census ID Number	High
	Respondent burden in entering the Census ID Number	Medium
Before You Start Screen	Need for instruction not to use the browser's back button	High
	Confusion over the purpose of the sample screen, leading to possible confusion at the beginning of the form	High
Household Screen	Design alternatives to achieve clarity in the presentation of the residence rules (needs further evaluation)	High
	Effectiveness of capitalization for emphasis	Medium
Name Screen	[no issues identified]	N/A
Relationship Screen	Benefits of filling in actual name for "Person 1" as a point of reference	Medium
Sex/Date of Birth/Age Screen	Need to draw attention to babies less than one year old	Medium
Origin	Cognitive challenges in both grid and stacked designs	Medium
Race Screen	Cognitive difficulties caused by stacking response options	Medium
	Respondent tendency to skip instructions	Medium
Ancestry Screen	Overcoming the false assumption that only three ancestries can be entered (problem identified in Round-1)	Medium
	Instruction and countdown on maximum number of characters (design appears to work well)	Low
Review/Submit Screen	Lengthy instructions take up too much space in the display area (Round-1 issue resolved by redesign)	Medium
	Confusing effect of labeling ("Residence" vs. "Other Residence")	Medium
Help Function	Respondent assumptions about help depend on design	Medium
	How to encourage use of help (needs further testing)	Medium
	Section 508 (accessibility) issues	Medium
	Loss of context when scrolling down in the help file	Medium
	Some respondents not aware that help is available	Medium
	Relevance and practicality of help text (need for web-based style of writing)	Medium to High
Edit Function	Acceptability of revised wording of edit messages (needs further testing)	Medium
	Need to make edit messages as specific as possible to avoid respondent confusion/misinterpretation	Medium
	Respondent tendency to skim messages (not read completely) >> Need to keep messages very brief	High
	Low salience of instruction to "continue" despite edit, if the respondent does not know an answer or has given his/her best answer	Medium to High
	Disappearance of edit messages with scrolling	High
Use of color & contrast for functional (not decorative) purposes	Reading of labels discouraged by low visual contrast	High

	Respondent tendency not to read instructions (needs further research)	Medium
Navigation Function	Lack of support for respondent's leaving and returning to the Review/Submit screen	High
	Confusion caused by displaying edit messages when respondent wants to go back to a previous page	High

We also conducted appropriate accessibility testing to ensure the application complies with Section 508 accessibility requirements. With the cancellation of the Internet option for the 2010 Census, this report remained in draft form. Because planning is now underway for an Internet option in the 2020 Census, this report is being finalized to provide data on design alternatives that were tested for usability in the latter months of 2005.

Introduction

In preparation for the 2005 National Census Test (NCT), the U.S. Census Bureau conducted one round of usability testing to evaluate the ease of gaining access to the online version of the census short form, completing the form, responding to edit messages, and reviewing the entered data and making changes (Murphy, Ciochetto, Malakhoff, Coon, Ashenfelter, & Smith, 2013). The 2005 NCT online survey was field tested with a census day of September 15, 2005. Since the Internet was expected to be a method of data collection in the 2010 Census, the Internet Team decided to continue research not only on the specific issues identified in this first usability test, but also on broader issues that apply to possible future designs. Therefore, we planned and conducted an additional round of usability testing from November 30 to December 14, 2005, after the NCT had finished collecting data. Obtaining data to inform future decision-making on design issues was the goal of this second round of usability testing. We also conducted appropriate accessibility testing to ensure the online survey complied with Section 508 requirements.

Purpose

The purpose of the present document is to report on the second round of usability testing of the 2005 Census Internet Short Form. The purpose of automated accessibility testing was to identify any violations of the Federal accessibility regulations (Section 508 of the Rehabilitation Act as amended in 1998). This report documents the methods and procedures used during the usability studies, as well as the results and recommendations of the usability testing.

Scope

Usability testing focused on several key design areas in which it was possible to implement alternative designs. These included the style of language used for instructions, edit messages, and help text; use of color and contrast in response category labels; and selected issues identified in Round-1 usability testing. This second round of testing also investigated ways to reduce the cognitive burden of using an electronic version of the paper questionnaire.

The Round-2 Online Survey

The 2005 NCT topic-based, production survey was the baseline for changes for this second round of usability tests. Based on Census Bureau guidance, Z-Tech Corporation developed two versions of a topic-based online survey for Round 2. Unlike the paper form, which uses a person-based design (that is, the respondent completes all the demographic information for a person before completing all the information for the next person) the online form asks for everyone's name before then collecting everyone's age and soon for all data items.

Global changes to the topic-based online survey included the following:

- The blue bar separating the question and responses was lightened to match the color of the blue background;
- The text in the blue bar was changed to black text in a smaller font size;
- Labels were included in all blue bars (black text);

- Instructions were placed on a separate line instead of at the end of a question;
- A button was added to the left of the Next button that allowed the participant to “Return to Review/Submit.” This button was displayed only when the participant had been to the Review/Submit screen once; and
- The reference date was changed from September 15, 2005 to December 7, 2005.

Another global change involved using an Office of Management and Budget (OMB) number that reflected approval for usability testing instead of approval for a field test; however, this was not a variable of interest in the testing.

Based on observations and recommendations from Round-1 testing, changes to specific screens were as follows:

- **Before You Start:**
 Note: One version of the online survey had the revised screen while the other version had no Before-You-Start screen
 - The instruction at the top of the screen about tips for using the site was removed;
 - Information was added about using the Next/Previous button instead of the browser Forward/Back button. This instruction was displayed in a text box;
 - The title, “Navigating the Questionnaire,” and the Next and Previous instructions were removed;
 - The sentence that read, “The following sample screen...,” was changed to the title **SAMPLE SCREEN** (in bold);
 - The wording in the callouts on the Sample Screen was shortened;
 - Callouts for the Next and Previous buttons were added on the Sample Screen;
 - The text explaining the color of the tabs was removed; and
 - The text about Supported Browsers was removed;
- **Household:**
 Note: There were two versions of this screen. One version had the residence rules listed on the screen; the other used the links.
 - In both versions of this screen, the text introducing the residence rules was shortened;
 - Emphasis in the form of all CAPS was added to the word **USUALLY** in the first bulleted residence rule and **MOST OF THE TIME** in the second bulleted residence rule in both versions of the screen;
 - The second bulleted residence rule was reworded to read, “count them where they sleep” instead of “this is the place where they sleep” in an effort to use the same word throughout the residence rules (count) to possibly reduce cognitive burden in both versions of the screen;
 - The label on the link from “**INCLUDE/EXCLUDE** these people” was changed to “Additional guidelines” in an effort to convey the meaning that this is a link to more rules, and to make the content flow more logically;
 - The sentence, “We will count them at the other place.” was removed from both versions of the screens;

- The titles of the lists of rules were changed from “Include” and “Exclude” to “Count” and “Do NOT Count;” and
- Emphasis in the form of all CAPS was added to the word “ADDITIONAL” in the tenure question.
- Relationship:
 - Added “is [Person 1]’s” after the name of each person in the left margin in an effort to reduce cognitive burden; and
 - Moved the response categories down one line so that the names in the left margin did not run into the first response.
- Hispanic Origin:

NOTE: There were two versions of this screen. One removed the grid design and stacked the response options on top of each other for each person.
- Race:

NOTE: There were two versions of this screen. One removed the grid design and stacked the response options on top of each other for each person.
- Ancestry:
 - The three separate lines for the responses were combined into one text box.
 - A counter was added to the text box showing the maximum number of characters and the number of characters remaining.
 - The text of the question and example were broken into shorter lines in an effort to make them more easily scanned.
- Review/Submit:
 - Shortened the sentence asking the respondent to review the data.
 - Added a sentence noting that this is the final opportunity to make changes.
 - Shortened the instructions for changing an answer and remove the reference to completing an answer left blank.
 - Shortened the instruction for submitting the data.
 - Moved the Submit button at the top of the page to the right side of the page so it did not interrupt the eye movement down to the data.
 - Added light blue shading (the same as the question uses) to help guide the eye from the label to the data.
 - Added “(Hispanic, Latino, or Spanish)” or “(Not Hispanic, Latino, Spanish)” after the data for Origin.

There were two versions of the help text and two presentations of the help file. One version was the same as that used in the production survey (except that the reference date changed from September 15, 2005 to December 7, 2005 and the help timers were added). This version contained all of the help in one scrollable file that displayed the appropriate text according to which question invoked the help. The other version contained text revised to make visual scanning easier (as compared to the original). It also had only the text from the question in the window (i.e., item-specific help). Links on a left-navigation bar, in a “navigation tree” form, allowed the participant to move to other help text. This revised version of the application also included timers to indicate when the help file was

opened, when it was closed and the duration for which it was open; these data allowed us to track participants' use of the system help.

Two versions of the edit messages were tested. One version was the same as was used in the then-current production survey (except that the reference date changed from September 15, 2005 to December 7, 2005; and the age calculation reflected this change in reference date). The other had shorter, more easily scanned text that gave a more precise message about the nature of the problem. There was also a slight formatting change in that the final sentence of the edit message (“If you have provided your best answer...”) had a blank line above and below it in an effort to make it stand out better. On the Household page in the version that changes the edit messages, we anchored the screen display at the point of the first edit message so it was apparent that there had been a change to the screen when it redisplayed with an edit message. (Previously, the user had to scroll down to view the first edit message.)

The following table summarizes the differences between each version of the topic-based online survey.

Table 1. Differences between the Round-2 Test Versions of the 2005 National Census Test (Internet option)

Version L	Version M
New Before You Start screen	No Before You Start screen
Household screen with links to residence rules	Household screen with residence rules on screen
Grid format for race and origin questions	Non-grid race and origin questions
Production version of help text/format	Revised version of help text/format
Production version of edit messages	Revised version of edit messages
Household screen with edit message anchored at top of page (no change from production)	Household screen with edit message anchored at point of edit message

Methods

Participants were assigned to one version (L or M), and they completed two census forms (one using their own data and one using scenario-based data).

Independent and Dependent Variables

The following variables were manipulated for the usability portion of the test:

- Display of edit messages on household screen:
 - Screen is displayed at the point of the first edit message (e.g., if the first edit message is below the “fold,” that is where the focus is placed when the screen refreshes); or
 - The screen refreshes and displays at the top, even if there are no edit messages above the fold.

- Grid or no-grid on race and ethnicity (origin);
- Long-style edit messages or short-style edit messages;
- Revised spacing of the continue part of the edit message or no change in spacing;
- Long-style help messages or short-style help messages.

The following measures were collected:

- Observed behavior, for example, the number of instances of confusion or frustration in response to an edit message; looking for and using the residence rules; the number of questions/comments about whether it's OK to continue and leave an edit message unresolved.
- On the Popcount question, time from getting scenario-based household information from the test administrator until entering the population count.
- When edit messages are displayed, does the participant read to the end of the message; does the participant know what needs to be corrected?
- For those who receive edit messages, do they see the "continue" message; do they act on it? Do they know it's OK to continue without resolving an edit?
- When help is accessed, do participants read the complete content or at least enough to find the answer to their question? Duration of help usage recorded from opening to closing of a help message.

Participants

The Human Factors and Usability Research Group in the Center for Survey Measurement (CSM) recruited 17 participants from outside the Census Bureau and one new employee from within CSM to stand in for a participant who did not appear for a scheduled session. A standard set of screening questions was used to recruit suitable participants (Appendix A). To the extent possible, participants represented various kinds of households (e.g., small, large, complex households where everyone was related, complex households with unrelated people, and households with a new baby). Ethnic diversity was a goal of the recruiting effort. Participants were expected to have a minimum of one year of regular Internet experience. Outside participants were paid a \$40.00 honorarium.

The target population (i.e., respondents who were expected to use the Internet in 2010) can be expected to be at least moderately computer literate and well-versed in using the Internet. Table 2 shows the major characteristics of the participants as self-reported in response to the screening questions and to a questionnaire on experience with computers and the Internet.

Table 2. Characteristics of Participants in Usability Testing of Alternative Designs for the Census Internet Form (Numbers in parentheses are the participants' self-reported ratings on a 9-point scale, where 1 is very low and 9 is very high.)

Participant Number	Sex	Age	Race	Education	Computer Experience	Internet Experience
--------------------	-----	-----	------	-----------	---------------------	---------------------

1	F	68	black	2 years college	7-8 years; 4-5 hours/day (6)	3-4 years; 2-3 times/week (6)
2	F	54	black	some college	20 years (8)	12-15 yrs (6)
3	F	17	black	HS senior	4 years (5)	4 years (6)
4	M	21	black	HS grad	frequent user (7)	3-4 hours daily (9)
5	F	52	black	Masters	20 years (9)	10 years (7)
6	F	39	black	some college	uses daily in work (8)	mostly on Internet (8)
7	F	late 40s	white	3 years college	frequent user (7)	(7)
8	F	35	black	some college	daily (7)	daily (9)
9	F	25	black	some college	7 years (7)	7 years (8)
10	F		black		(7)	(7)
11	F	39	black	some college	daily (7)	3-4 hrs daily (7)
12	F	27	black	B.A.	several hrs daily (8)	several hrs daily (9)
13	M	20	black	college jr.	4 years (7)	2-3 hrs daily (7)
14	M	18	black	college freshman	2 years (7)	2-3 hrs daily (7)
15	F	late 20s	white	Ph.D.	(9) highly experienced	(9) highly experienced
16	M	65	white	Masters	frequent user (6)	(5)
17	M	58	white	B.A.	daily user; hand tremors & other disabilities (6)	(7)
18	F	23	white	B.A.	daily user (7)	mostly Internet (7)

The accessibility testing did not use participants. Instead, the online survey versions were run through an automated accessibility testing software package known as Insight/InFocus 4.2.2. Items flagged for manual checks were evaluated using the JAWS 7 (Job Access With Speech) screen reader.

Facilities and Equipment

The usability participants came to the Census Bureau's Usability Laboratory for the testing sessions. The test participant sat in one of 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. The workstation included a standard keyboard and mouse. Two microphones picked up sound in each testing room. The test administrator sat in the testing room with the test participant.

Usability Materials

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. Taping began when the participant signed the consent form.

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 went into more detail than the general recruiting screener, which asked how many years of experience they have had with the Internet.

Introduction to a Testing Session. A copy of the introduction read aloud to 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.

Satisfaction Questionnaire. A tailored version of the Questionnaire for User-Interaction Satisfaction (QUIS) appears in Appendix E. The core questions were the same as those used in Round-1 usability testing. Questions were added to assess user satisfaction with the edit messages and the help. The original QUIS was developed and validated at the University of Maryland (Norman, Shneiderman, Harper, and Slaughter, 1998).

Debriefing Questions. The kinds of questions that were asked during the retrospective portion of the test session appear in Appendix F. 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.

Usability Procedure

The test administrator (TA) welcomed the participant and gave a short introduction to the test session (Appendix D). The TA explained that we wanted the participant to think aloud while completing the census form. A short training in thinking aloud was given. Training consisted of asking the participant to use the Internet in some way that they

ordinarily would use it (e.g., to find the local weather, to check on sports scores, to check information at a club's site). We did not ask everyone to find the weather because we have found in past testing that some people have never used the Internet to check the weather and have no idea how to go about doing that. We wanted them to do something they felt comfortable with up front. Following the training in thinking aloud, the test administrator gave the participant a mailing package and a valid Census Housing Unit Identification Number to use in accessing the online form.

For this round of usability testing, the participant completed the online survey twice. The first time through the set of screens, participants provided data for their own households. They provided a think-aloud commentary while working on the form. Although the TA was in the room with the participant, there was minimal probing by the TA (only to remind the participant to think aloud).

On his/her second time through the set of screens, the participant entered data from a scenario prepared in advance by the test administrators (Appendix G). This was not strictly a data-entry task, however, because the TA played the role of a friend who needed help completing their census form. The participant asked the questions, and the TA/friend provided the predetermined responses. Some of these responses were designed to nudge the participant to look at help (for example the scenario included people who had another residence and were not core members of the household); others were designed to trigger edit messages (missing data). The TA did not probe the participant, but rather remained in the role of a friend who was being helped. After submitting the scripted data, the participant completed the QUIS.

After completing the census form and QUIS questionnaire, the participant was debriefed. The debriefing portion of the session often began with a discussion of any especially low or high ratings on the QUIS.

Finally, the TA thanked the participant, had him/her sign the payment voucher form and paid the participant. The one participant who was a federal employee was not paid for her participation.

Logging

Usability test sessions were logged either concurrently or at some later time. Logging was performed by a contractor (UserWorks, Inc.). Appendix H provides the set of codes used to capture the data of interest.

Data Analysis

The analysis of usability data provided the basis for deciding whether the design changes were effective.

Usability issues were identified by observation of the interactions between the participants and the online surveys. 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' success and satisfaction. If the participant used the "close window" link, the application was able to capture the time that the help file was open for each version of the instrument to see if one version took less time for the participant to use.

For logged behaviors, we calculated descriptive statistics (e.g., frequencies, mean number of occurrences). 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.

We considered behaviors such as confusion to be indicators of potential problems. In making comparisons, we compared indicators and counts between the two versions in a qualitative sense (Couper, October 14, 2005, personal communication). It was not possible to conduct formal hypothesis testing because of the small number of participants.

Results and Recommendations Based on Usability Testing


This section discusses the results of the usability testing. It is organized first by specific screens/items and then general issues. Screen shots of the two versions are included as well as screen shots from the production 2005 NCT online survey, where appropriate. The design goals are discussed within the treatment of each alternative feature. Finally, recommendations are made for future design or testing. Recommendations are prioritized as high, medium, or low depending on the extent to which a particular design feature will affect the respondent's ability to complete the online Census form.

Housing Unit Identification (HUID) Screen

Enter your Census ID number in the boxes below. Then click 'Submit'.

- - -

Your Census ID number is the 18-digit number in the address section of the materials we sent you in the mail.



Census ID Number: 00000-00000-00000-000

D-0000001 *****AUTO**3-DIGIT 112
TO RESIDENT AT
APT. 0007
2005 CENSUS AVENUE
SUITLAND, MD 20233-1234




Figure 1. HUID Screen – Version L and Version M

Instruction for finding the Census ID Number

The instruction on where to find the Census ID number was moved after Round 1 from above the graphic on the right to below the response boxes on the left. Several participants still did not see/read the instruction. One participant said that people are not going to know what their Census ID number is so there should be instructions. Another participant said that we were asking him to do something without giving him proper instructions.

One participant repeated the phrase “in the address section of the materials we sent you in the mail” several times trying to figure out what it meant. She and several other participants picked up the mailing envelope, saw that it was an open window and then picked up the return envelope and glanced at it. Not finding a bar code, they first searched the “front” of the questionnaire (the side with the questions). Again not finding a bar code, they turned over the questionnaire and finally found the Census ID number.

Recommendation: Use a step-wise set of instructions to guide the respondent:

- 1) You will need your Census ID number.
- 2) Here’s where to find your Census ID number: [Explicit directions on where to find the Census ID number]
- 3) Enter the ID.
- 4) Press submit.

Use short phrases as instructions. Place these instructions *above* the response box. Finding and entering the correct Census ID number is critical to the respondent's success in accessing the online form.

Priority: High

Graphic showing the Census ID Number

The red arrow was a strong visual cue that drew participants' eyes to the graphic. However, two participants did not realize that this was a sample graphic and thought that the ID number in the graphic was the one that they should use, even though it was all zeros.

Recommendation: Revise the graphic so that the sample ID number uses X's instead of zeros, thus making it more obvious that this is a sample. Make the graphic look more like what the respondent actually sees. Perhaps have a picture of the form in the mailing envelope. It is critical for the respondent to find and enter the correct Census ID number, and the design should do everything possible to assist.

Priority: High

Entering the Census ID Number

Once participants found the Census ID Number, their heads were down and they began entering their numbers. The first entry box did not have focus, so nothing happened. When they looked up and saw nothing in the boxes, they immediately knew what they had to do. Then they clicked on the first box and again entered the number. However, having to diagnose the problem and re-enter the number adds to respondent burden and to the respondent's level of annoyance.

As with the first round of usability testing, some participants liked the auto-tabbing and others did not expect it.

Recommendation: When the screen is initially displayed, give the box focus so the respondent can start typing without first having to click on the box. This will reduce respondent burden. Keep the auto-tabbing. Those who do not expect auto-tabbing will be able to recover if they tab past a data-entry field.

Priority: Medium

Before You Start Screen

⚠ Do NOT use your browser's Forward or Back buttons to navigate. Instead, use the Next or Previous button.

If you prefer to skip this screen, please click the Start button. Start

SAMPLE SCREEN

U.S. Census Bureau
2005 National Census Test

Help | Privacy Policy

Household
Names
Relationship
Sex/Date of Birth/Age
Origin
Race
Ancestry
Residence
Review/Start

? What is each person's sex, date of birth, and age?
Please report babies as age 0 when the child is less than 1 year old.

	Sex	Date of Birth (mm/dd/yyyy)	Age (on 12/07/2005)
Person 1	<input type="radio"/> Male <input type="radio"/> Female	<input type="text"/> / <input type="text"/> / <input type="text"/>	<input type="text"/> years
Person 2	<input type="radio"/> Male <input type="radio"/> Female	<input type="text"/> / <input type="text"/> / <input type="text"/>	<input type="text"/> years
Person 3	<input type="radio"/> Male <input type="radio"/> Female	<input type="text"/> / <input type="text"/> / <input type="text"/>	<input type="text"/> years
Person 4	<input type="radio"/> Male <input type="radio"/> Female	<input type="text"/> / <input type="text"/> / <input type="text"/>	<input type="text"/> years

<< Previous
Next >>

Go back one screen
Go forward one screen

Respondents are not required to respond to any information collection unless it has a valid approval number from the Office of Management and Budget. This test has been approved and given the following OMB number: 0607-0725

USCENSUSBUREAU

To begin, click the Start button. Start

Figure 2. Before You Start Screen -- Version L

Participants who used Version M did not see the Before You Start screen. They proceeded immediately to the household page. A couple of these participants used the browser's back button to return to a previous screen. The instruction not to use the browser's back button is one of the instructions on the Before You Start screen.

All Version L participants looked over the screen. Only one participant said she would normally skip the screen. One participant scrolled back up to use the Start button at the top of the page.

Only a couple of the participants who saw the Before You Start screen (see Figure 2) still did not understand that the sample screen was meant to give them an overview of how the screens in the online survey were laid out and how to navigate through the application.

One participant reported that “sample” meant “practice.” He tried to enter data on the screen. Another participant was not sure how to close the callouts. This is consistent with what we think respondents are expecting. They go to the application to enter their census data. They do not expect the first screen to be an instructional one.

Recommendation: Consider whether an instructional screen is, indeed, necessary. It was not clear whether or not our participants benefited from anything except the instruction not to use the browser back button. Perhaps this instruction could be moved to the first page, or perhaps this restriction on using the browser back button will be removed in future designs of the online survey. If it is critical for the respondent not to use the browser back button, this instruction needs to be displayed prominently.

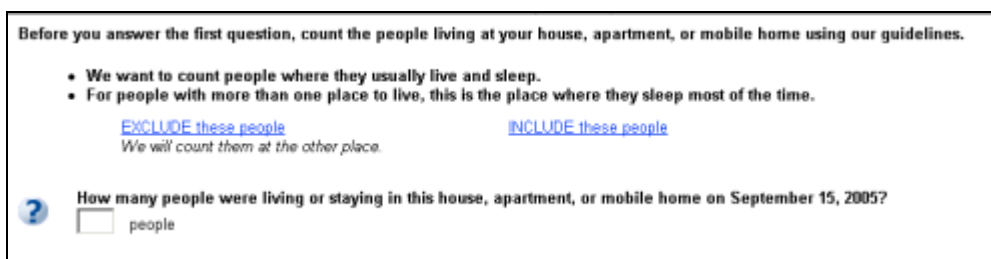
Priority: High (for keeping the instruction not to use the browser back button)

If the Before You Start page is included, do not use an actual screen shot in the graphic. A wireframe, which is just a skeletal view of the screen, may present enough information about layout and navigation without implying there is another purpose to the screen. Consider different terms for “navigate” and “sample screen.” Add instructions to the screen to make respondents aware that this is an informational screen only. It is important not to confuse respondents to the point that they decide not to continue.

Priority: High (to avoid confusion at the beginning of the form)

Household Screen – Residence Rules and Popcount

Figure 3 shows how the residence rules were presented in the 2005 NCT production version. Figures 4 and 5 show the presentation of the residence rules in versions L and M, respectively, as presented in this round of usability testing.



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.


[EXCLUDE these people](#) [INCLUDE these people](#)
We will count them at the other place.

How many people were living or staying in this house, apartment, or mobile home on September 15, 2005?
 people

Figure 3. Residence Rules – 2005 NCT Production Version

Count the people living at your house, apartment, or mobile home using these guidelines...

- Count people where they **USUALLY** live and sleep.
- For people with more than one place to live, count them where they sleep **MOST OF THE TIME**.
- [Additional guidelines](#)

 How many people were living or staying in this house, apartment, or mobile home on December 7, 2005?

people

Figure 4. Residence Rules -- Version L

Count the people living at your house, apartment, or mobile home using these guidelines...


- Count people where they **USUALLY** live and sleep.
- For people with more than one place to live, count them where they sleep **MOST OF THE TIME**.

Do NOT Count:

- College students who live away from this address most of the year
- Armed Forces personnel who live away
- People who, on December 7, 2005, were in a:
 - Nursing home, mental hospital, etc.
 - Jail, prison, detention facility, etc.

Count:

- Babies and children living here, including foster children
- Roommates or boarders
- People staying here on December 7, 2005 who have no other permanent place to live

 How many people were living or staying in this house, apartment, or mobile home on December 7, 2005?

people

Figure 5. Residence Rules -- Version M

Using the links for detailed residence rules versus displaying all of the rules on the screen.

All but one Version L participant read the introductory sentence and the two bulleted sentences before going to the first question (see Figure 4). Participants who read the introductory sentence seemed to think that these two bullets gave them a good rule for whom to count as a household member. Only one participant clicked on the “Additional guidelines” link when filling out his or her own data, and that person used the link when she went back to the household screen from the Review/Submit screen. She wanted to tell us that another son was a family member but did not live in the house on the reference day because he was in college.

On the other hand, when interviewing the TA for the scenario, two participants used the links to see how college students should be counted. One participant asked the TA/Ethel if she needed additional guidelines and clicked on the link when she said yes. This indicates that the participant understood the use of the link as a way to get additional information.

This understanding of the link was not apparent in the first round of usability testing (see Figure 3 for the layout of the screen with residence rules in the production application). The change in the title of the link in Version L was definitely an improvement.

When doing the scenario interview, one participant realized that the person in has she had indicated on the Overcount question, and who only came home on the weekends, should not have been included in the Overcount question. Initially she did not count him, but when she scanned the categories in the undercount question and saw “people staying here temporarily” she went back and changed the Popcount number to include him. This demonstrates that our respondents do not always know when they need additional guidelines. They assume that their understanding of the composition of their family matches the Census Bureau’s residence rules.

Near the beginning of the scenario, one participant correctly entered six people for the scenario household. However, when asked whether it makes a difference that Sam is in college, he replies, “not yet,” indicating a less-than-complete understanding of the residence rules.

Version M presented the residence rules on the screen (see Figure 3). Although some participants read or scanned through all of the rules, a couple went straight to the first question without even reading the introductory sentence or even the first two bulleted guidelines. These actions are not totally unexpected because people are known to scan quickly or skip text when completing an online form.

One participant did not read the residence rules, but asked the TA/Ethel, “How many people are living or staying in your house?” When questioned about whether to include the older son, the participant stated that he did not count because he was in college. However, the son commuted to college and lived at home; so he should have been counted as living in the household. The Popcount was recorded as five, which was one short.

Those who read the displayed guidelines may have been attracted to the short phrases. Those who skipped them entirely may have seen the large amount of text and may not have wanted to read it but rather get on to the task of answering questions. Not reading any of the guidelines did not give the participant any rules on which to base their response.

Even those who read the residence rules, at least to some extent, came to erroneous conclusions about the number of people in the scripted household. When questioned about the son in college, one participant stated that he is in college, so not counted. The participant then continued to count the remaining family members and recorded the Popcount as four when it should have been six.

Thus, simply scanning the guidelines may be problematic. For the most part, participants who used Version L (guidelines accessed by the link) read the bullet that said to count people where they sleep most of the time. They were then probed by the test

administrator to determine if this is where the older son slept most of the time and correctly included him in the household roster. On the other hand, two participants who used Version M (guidelines listed on the screen) scanned the list saw “Do NOT Count” and “College Students” and said that the son should not be listed on the roster because he attended college. Thus, the only participants who incorrectly counted the people in the scenario household were the ones who used the version of the application that had the guidelines presented on the screen.

The logs also indicate an extremely low rate of looking for the residence rules, as shown in. In the logs, the observers noted that no one using Version M used the residence rules in Round 2; however in the usability testing notes, there was indication that some participants used the rules in Version M.

Recommendation: Displaying all of the residence rules on the screen does not guarantee that a respondent will read them. In this round of testing, the residence rules were correctly applied more often when the rules were NOT on the screen. However, we need to inform respondents about the rules that apply to them. Perhaps a middle ground between not displaying any of the detailed rules and displaying all of the details would be to list the categories on the screen with a title that indicates these situations need special attention and then put the details behind a link. For example,

Special living situations (click for instruction):

[College Students](#)

[Armed Forces](#)

[Nursing Home](#)

[Foster Children](#)

[Jail](#)

[Boarders](#)

This design would also encourage respondents to seek help at a very early stage in their completion of the online form. See the discussion in Section 10b about participant use of help.

Priority: High

Using capitalization to emphasize words

We wanted to see whether participants would emphasize the words in all capitalized text as they read questions or instructions aloud. Almost half of the participants emphasized the words that were written in all capitalized letter when they read the word aloud.

One participant stressed the capitalized words in the instruction, “Count people where they USUALLY live and sleep...count them where they sleep MOST OF THE TIME.”

Recommendation: Capitalization seemed to be an effective way to emphasize text. It should be used sparingly to emphasize an isolated word or phrase. If overused, the effectiveness of this technique is lost.

Priority: Medium

Name Screen

When filling out data for her own household, one participant said everyone in the household had the same last name even though she entered “Jr.” as part of P1’s name. When she got to the names screen, she removed the Jr. Similarly, in the scenario, the test administrator said everyone had the same last name on household page, but when the names page came up, she said the nanny had a different last name. All participants who said everyone had the same last name were easily able to make changes to the nanny’s last name. One participant questioned whether or not she could change the name but tried it and was pleasantly surprised that it was an easy change to make.

Recommendation: None

Relationship Screen

Figure 6 presents the content and layout of the relationship question as displayed in both Version L and Version M tested for usability in Round 2.


Household	Names	Relationship	Sex/Date of Birth/Age	Origin	Race	Ancestry	Residence	Review/Submit		
<p> How is each person related to John?</p>										
			Related			Not Related				
<p>Jane is John's</p> <table style="width: 100%; border: none;"> <tr> <td style="width: 33%; vertical-align: top;"> <input type="radio"/> Husband or wife <input type="radio"/> Biological son or daughter <input type="radio"/> Adopted son or daughter <input type="radio"/> Stepson or stepdaughter <input type="radio"/> Brother or sister <input type="radio"/> Father or mother <input type="radio"/> Grandchild <input type="radio"/> Parent-in-law <input type="radio"/> Son-in-law or daughter-in-law <input type="radio"/> Other relative </td> <td style="width: 33%; vertical-align: top;"> <input type="radio"/> Roomer or boarder <input type="radio"/> Housemate or roommate <input type="radio"/> Unmarried partner <input type="radio"/> Foster child or foster adult <input type="radio"/> Other nonrelative </td> </tr> </table>									<input type="radio"/> Husband or wife <input type="radio"/> Biological son or daughter <input type="radio"/> Adopted son or daughter <input type="radio"/> Stepson or stepdaughter <input type="radio"/> Brother or sister <input type="radio"/> Father or mother <input type="radio"/> Grandchild <input type="radio"/> Parent-in-law <input type="radio"/> Son-in-law or daughter-in-law <input type="radio"/> Other relative	<input type="radio"/> Roomer or boarder <input type="radio"/> Housemate or roommate <input type="radio"/> Unmarried partner <input type="radio"/> Foster child or foster adult <input type="radio"/> Other nonrelative
<input type="radio"/> Husband or wife <input type="radio"/> Biological son or daughter <input type="radio"/> Adopted son or daughter <input type="radio"/> Stepson or stepdaughter <input type="radio"/> Brother or sister <input type="radio"/> Father or mother <input type="radio"/> Grandchild <input type="radio"/> Parent-in-law <input type="radio"/> Son-in-law or daughter-in-law <input type="radio"/> Other relative	<input type="radio"/> Roomer or boarder <input type="radio"/> Housemate or roommate <input type="radio"/> Unmarried partner <input type="radio"/> Foster child or foster adult <input type="radio"/> Other nonrelative									
<p>Joe is John's</p> <table style="width: 100%; border: none;"> <tr> <td style="width: 33%; vertical-align: top;"> <input type="radio"/> Husband or wife <input type="radio"/> Biological son or daughter <input type="radio"/> Adopted son or daughter <input type="radio"/> Stepson or stepdaughter <input type="radio"/> Brother or sister <input type="radio"/> Father or mother <input type="radio"/> Grandchild <input type="radio"/> Parent-in-law <input type="radio"/> Son-in-law or daughter-in-law <input type="radio"/> Other relative </td> <td style="width: 33%; vertical-align: top;"> <input type="radio"/> Roomer or boarder <input type="radio"/> Housemate or roommate <input type="radio"/> Unmarried partner <input type="radio"/> Foster child or foster adult <input type="radio"/> Other nonrelative </td> </tr> </table>									<input type="radio"/> Husband or wife <input type="radio"/> Biological son or daughter <input type="radio"/> Adopted son or daughter <input type="radio"/> Stepson or stepdaughter <input type="radio"/> Brother or sister <input type="radio"/> Father or mother <input type="radio"/> Grandchild <input type="radio"/> Parent-in-law <input type="radio"/> Son-in-law or daughter-in-law <input type="radio"/> Other relative	<input type="radio"/> Roomer or boarder <input type="radio"/> Housemate or roommate <input type="radio"/> Unmarried partner <input type="radio"/> Foster child or foster adult <input type="radio"/> Other nonrelative
<input type="radio"/> Husband or wife <input type="radio"/> Biological son or daughter <input type="radio"/> Adopted son or daughter <input type="radio"/> Stepson or stepdaughter <input type="radio"/> Brother or sister <input type="radio"/> Father or mother <input type="radio"/> Grandchild <input type="radio"/> Parent-in-law <input type="radio"/> Son-in-law or daughter-in-law <input type="radio"/> Other relative	<input type="radio"/> Roomer or boarder <input type="radio"/> Housemate or roommate <input type="radio"/> Unmarried partner <input type="radio"/> Foster child or foster adult <input type="radio"/> Other nonrelative									
<p><< Previous</p>										

Figure 6. Relationship Screen -- Version L and Version M

Tailoring the relationship screen to ask the relationship of Person X to Person 1 specifically seemed to help all participants. When entering the data for the scenario, one

participant told the TA, “the world revolves around Ethel” (Person 1). This is definitely the idea that we wanted to get across. Adding the additional text (“is [Person 1]’s”) was attention grabbing and also helped when the participant had to scroll down and the question/reference person was no longer visible on the screen.

Recommendation: Continue to use the name of Person 1 as the point of reference in the left-hand column.

Priority: Medium

Relationship column headers

There are two columns of relationship categories in this question- “Related” and “Not Related”. The term “Related” was vocalized by the screen-reader for entries in both columns. In order to meet accessibility requirements, “Related” should be associated with column 1. “Not Related” should be associated with column 2.

Recommendation: The response options in column 2 should be vocalized with the phrase “Not Related” to be equivalent to what persons with normal vision experience.

Priority: Medium

Sex/Date of Birth/Age Screen

As shown in Figure 7, both versions used the same presentation for the question on sex, date of birth, and age.

	Sex	Date of Birth (mm/dd/yyyy)	Age (on 12/07/2005)
Mike	<input type="radio"/> Male <input type="radio"/> Female	<input type="text"/> / <input type="text"/> / <input type="text"/>	<input type="text"/> years
Mary	<input type="radio"/> Male <input type="radio"/> Female	<input type="text"/> / <input type="text"/> / <input type="text"/>	<input type="text"/> years
Mark	<input type="radio"/> Male <input type="radio"/> Female	<input type="text"/> / <input type="text"/> / <input type="text"/>	<input type="text"/> years

Figure 7. Sex/Date of Birth/Age Screen -- Version L and Version M

Instructions for the Age/Date of Birth Screen

We placed the instruction to report babies as age 0 on a separate line in an effort to make it more visible to participants. However, about half of the participants never read the instruction. This estimate is supported by the observation that in the scenario interview (the second time the participants saw the screen), when the test administrator/Ethel asked

the participants to change the age of the baby from the zero that was calculated to one year old, most participants complied. A couple of participants did comment that the word “years” was after the entry box, so “it seems like they aren’t counting months.” Those who saw the instruction told the test administrator “the instructions say he counts as zero.”

Recommendation: It might help to put the word BABIES and the phrase, LESS THAN 1 YEAR OLD in all capitals to draw the respondent’s attention.

Priority: Medium

Estimating Age and Date of Birth

An issue for participants in the scenario was whether it was acceptable to guess or estimate age and/or date of birth. The scenario did not give an age or date of birth for Natasha, who was Ethel’s mother-in-law. Similarly, the information in the scenario indicated that the nanny was “about 20” but did not give a date of birth.

One participant was not sure whether she could guess on date of birth and age for Natasha; so she consulted help on this topic. The participant skimmed through most of the help but skipped over the one sentence that instructed her to estimate age if she does not know the exact age. She decided to leave some answers blank and made some comments about the purpose of the edit message.

When asked if it was OK to estimate an age for Natasha, one participant replied that she did not know, stating, “It doesn’t say anything about estimating.” Later on, this participant commented that if it is acceptable to estimate, it should be explained in the question itself instead of being stated only in the edit message.

Another observation about estimation is that people assumed that if it was acceptable to estimate on age, then they had a responsibility to determine a correct year of birth to match the estimated age. People did not seem to make a distinction between estimating age and estimating year of birth. In their minds, the two seemed to go together inseparably. Their thinking seemed to follow the rule that if you are allowed to estimate age, you should give a corresponding year of birth. It did not seem to make sense to them to estimate age and then leave year of birth blank. For these participants, putting in an age, even an estimated age, seemed to REQUIRE filling in a year of birth.

Hispanic Origin Screen

As shown in Figures 8 and 9, the two versions used two different approaches to laying out the response options to the Hispanic origin question.

Household	Names	Relationship	Sex/Date of Birth/Age	Origin	Race	Ancestry	Residence	Review/Submit
<p>? Is each person of Hispanic, Latino, or Spanish origin? Examples</p>								
		Yes	No					
John		<input type="checkbox"/>	<input type="checkbox"/>					
Jane		<input type="checkbox"/>	<input type="checkbox"/>					
Joe		<input type="checkbox"/>	<input type="checkbox"/>					
<p><< Previous</p>								

Figure 8. Hispanic Origin Screen -- Version L (grid)

We investigated an alternative to using a grid in the Hispanic origin question because grids or matrices can be cognitively challenging because of the complexities of the format (Dillman, Smyth, and Christian, 2009). Our alternative design stacked the responses vertically (see Figure 9). In the scenario with six people in the household (one of whom was Hispanic), some participants wanted to put the yes response in first and then enter no for everyone else. For them, there was no visual reinforcement that they had done this correctly, as there was in the grid design. They often had to concentrate a little harder to make sure their answer was correct. Accessibility testing revealed column headers used as labels for the yes/no checkboxes in Figure 8 were not associated with these checkboxes, making these fields inaccessible. In summary, the grid layout design for the Hispanic origin question was inaccessible.

Household	Names	Relationship	Sex/Date of Birth/Age	Origin	Race	Ancestry	Residence	Review/Submit
<p>? Is each person of Hispanic, Latino, or Spanish origin? Examples</p>								
<p>Origin</p>								
Mike		<input type="checkbox"/> Yes	<input type="checkbox"/> No					
Mary		<input type="checkbox"/> Yes	<input type="checkbox"/> No					
Mark		<input type="checkbox"/> Yes	<input type="checkbox"/> No					
<p><< Previous</p>								

Figure 9. Hispanic Origin Screen -- Version M (non-grid)

Recommendation: If an alternative to the grid design is necessary, explore other options. Both designs seemed prone to error.


Race Screen

Household	Names	Relationship	Sex/Date of Birth/Age	Origin	Race	Ancestry	Residence	Review/Submit
<p>? What is each person's race? Examples Select one or more races. <i>For this census, Hispanic origins are not races.</i></p>								
Race(s)								
	White or Caucasian Examples	Black, African Am., or Negro Examples	American Indian or Alaska Native Examples	Asian Examples	Native Hawaiian or Other Pacific Islander Examples	Some other race Examples		
John	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Jane	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Joe	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<< Previous								

Figure 10. Race Screen -- Version L (grid)

Alternative grid design

As with the Hispanic origin screen, we investigated an alternative to using a grid in the race question (see Figure 11). One Version-M participant did not make a selection for the first person on the roster. She did not see her mistake until the edit message was displayed. This case was the only indication that this design has the potential for error. However, we noticed that stacking the response choices made the task more cognitively demanding than using the grid design.

 What is each person's race? Examples <i>Select one or more races.</i> <i>For this census, Hispanic origins are not races.</i>	
Race(s)	
Mike	<input type="checkbox"/> White or Caucasian <input type="checkbox"/> Black, African Am., or Negro <input type="checkbox"/> American Indian or Alaska Native <input type="checkbox"/> Asian <input type="checkbox"/> Native Hawaiian or Other Pacific Islander <input type="checkbox"/> Some other race
Mary	<input type="checkbox"/> White or Caucasian <input type="checkbox"/> Black, African Am., or Negro <input type="checkbox"/> American Indian or Alaska Native <input type="checkbox"/> Asian <input type="checkbox"/> Native Hawaiian or Other Pacific Islander <input type="checkbox"/> Some other race
Mark	<input type="checkbox"/> White or Caucasian <input type="checkbox"/> Black, African Am., or Negro <input type="checkbox"/> American Indian or Alaska Native <input type="checkbox"/> Asian <input type="checkbox"/> Native Hawaiian or Other Pacific Islander <input type="checkbox"/> Some other race

[<< Previous](#)

Figure 11. Race Screen -- Version M (non-grid)

Recommendation: Since the stacking of the response options seemed to make it more cognitively difficult for our participants, we recommend either keeping the grid design or exploring other designs.

Priority: Medium

Accessibility testing with the JAWS screen-reader revealed the Person 1 name was being associated with the Person 2 Race checkbox response options.

Recommendation: If the Version M (non-grid) Race option is selected for production, accessibility testing should verify the Person 1 name is associated with the first group of checkbox response options, the Person 2 name is associated with the second group of response options, and so on.

Instructions for the Race question

As we did with the instruction in the age question, we placed instructions associated with the race question on separate lines to help make them stand out to someone who would scan the question. These were instructions to select one or more races and the instruction that Hispanic origins are not races for this census. Note that the Hispanic origin instruction was also in italicized text. As with the age instruction, it was evident that

several participants did not read the instructions. When the TA said that one of the family members in the scenario identified with two races, these participants did not know whether they could make more than one selection. Most did not know how to handle the scenario household member who was Hispanic. Many participants listed the nanny as being of “Some other race” because the only information given about her was that she was Hispanic (from Argentina).

Recommendation: This is probably a case of people not realizing that they need any instruction or explanation. It may also be that their eyes were drawn to the graphical layout of the response options and away from the questions/instructions. For this question, it might be helpful to put the instructions in each of the person-name boxes to the left of the response options. The instructions would then be repeated for each person.

Priority: Medium

Ancestry Screen

Figure 12. Ancestry Screen -- Production Application

Figure 13. Ancestry Screen – Version L and Version M

Text area

Since the three text boxes in the 2005 NCT production application (see Figure 12) sent a strong signal to respondents that only three ancestries could be entered, both versions of the application in the second round of testing included a free-text area in which the participant could type a response. None of the Round-2 participants said, “Oh, you only

want me to enter three,” as many did in the first round of usability testing when we tested the production screen design.

From looking at the examples, participants seemed to understand that nouns were required (as opposed to full sentences). Some typed multiple entries all on one line, while others used multiple lines. In the previous round of usability testing, one participant tried to enter the name of her tribe, but it did not fit in the text box. She stopped typing at the end of the box because she did not realize she could continue. Although we did not have any participants with a long name to enter in this second round of testing, the alternative design would have alleviated any problem.

Recommendation: Keep the free-text area instead of the three text boxes.

Priority: Medium

Text area instruction

Below the text area was a short instruction that displayed the maximum number of characters for the response and counted down the remaining characters (see Figure 13, above). Although none of the participants made any remarks about the instruction, they all seemed to use it. Evidence of this is the observation that participants used conjunctions and punctuation when recording multiple ancestries. In other words, they were not concerned with the length of their response. No participant came close to using the maximum number of characters.

Recommendation: Keep the instruction under the text box.

Priority: Low

Accessibility of text area instruction

Screen reader users were not given an indication of how much they could type into the ancestry dialogue box. First, there was no notice of how much space was available before text could be entered, and second, the “characters remaining” function was not accessible. There was no equivalent access to the information about remaining space for JAWS users.

Recommendation: Leave the text positioned as is, but change the tab order so this text is vocalized before the ancestry dialogue box. Since JAWS users cannot see how many characters remain in the ancestry dialogue box when they are typing, a chime should be sounded when the space is used up.

Priority: High

Review/Submit Screen

Figure 14 shows the top of the Review/Submit screen in the production application but without the tabs or banner at the very top of the screen.

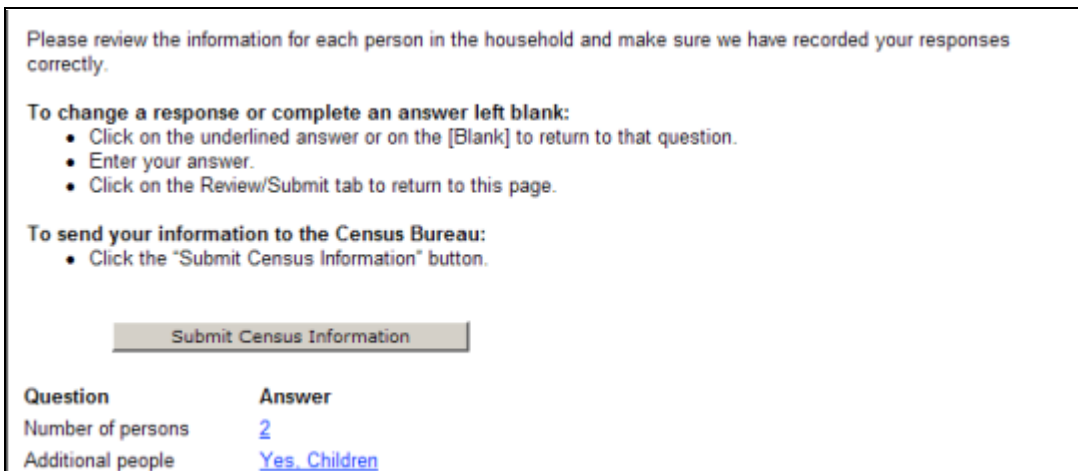


Figure 14. Review/Submit screen -- Production application

Figure 15 shows the alternative design for the Review/Submit screen of the topic-based application evaluated in this round of usability testing. This alternative was used in both Version L and Version M.

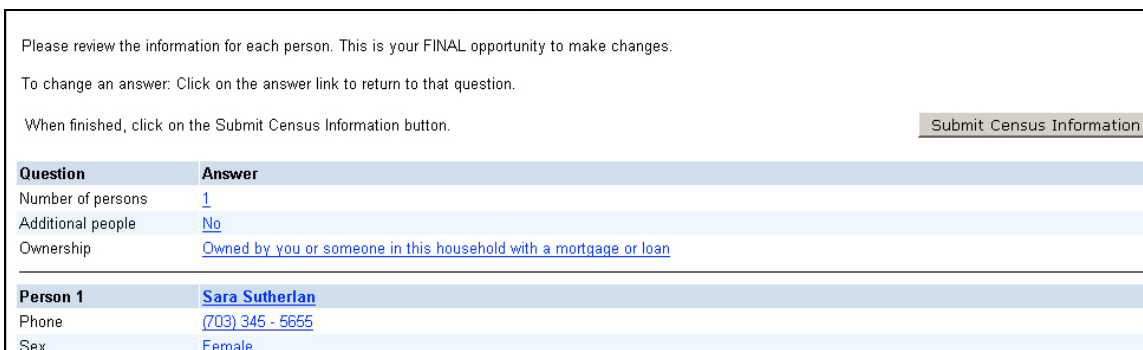


Figure 15. Review/Submit screen – Version L and Version M

Instructions on Review/Submit Screen

Participants read or scanned all of the instructions at the top of this page. As part of the scenario, they were asked to change a response. Several participants scrolled back up to the top of the page to find out exactly how to do that because they remembered seeing an instruction on how to change an answer. In the first round of usability testing, participants were confused by the instruction to “click on the underlined answer” presumably because they could not see any of the data since the instruction took up the display area of the screen. By reducing the amount of real estate, the instructions took and shortening the verbiage, the alternative design (Figure 15) seems to have alleviated this problem.

Recommendation: Keep the instructions on the Review/Submit screen short and concise so that some of the respondent’s data appears “above the fold.” Replace the comma in

the third line (“When finished, ...”) with a colon to match the formatting of the second line. This may make it easier to scan.

Priority: Medium

Labels on the Review/Submit Screen

The label on the Review/Submit screen for Residence was changed to read “Other Residence.” Unlike the behavior of some participants during the first round of usability testing (when the label just read Residence), no one went back to the residence screen to see what this answer meant. When discussing the change after the first round of usability testing, there was some concern that the label would not match the tab and that this would cause confusion. We did not observe any such confusion.

Recommendation: Keep “Other Residence” as the label.

Priority: Medium

Help Text

As shown in Figures 16, 17, and 18, we compared two ways of presenting help in this second round of usability testing.

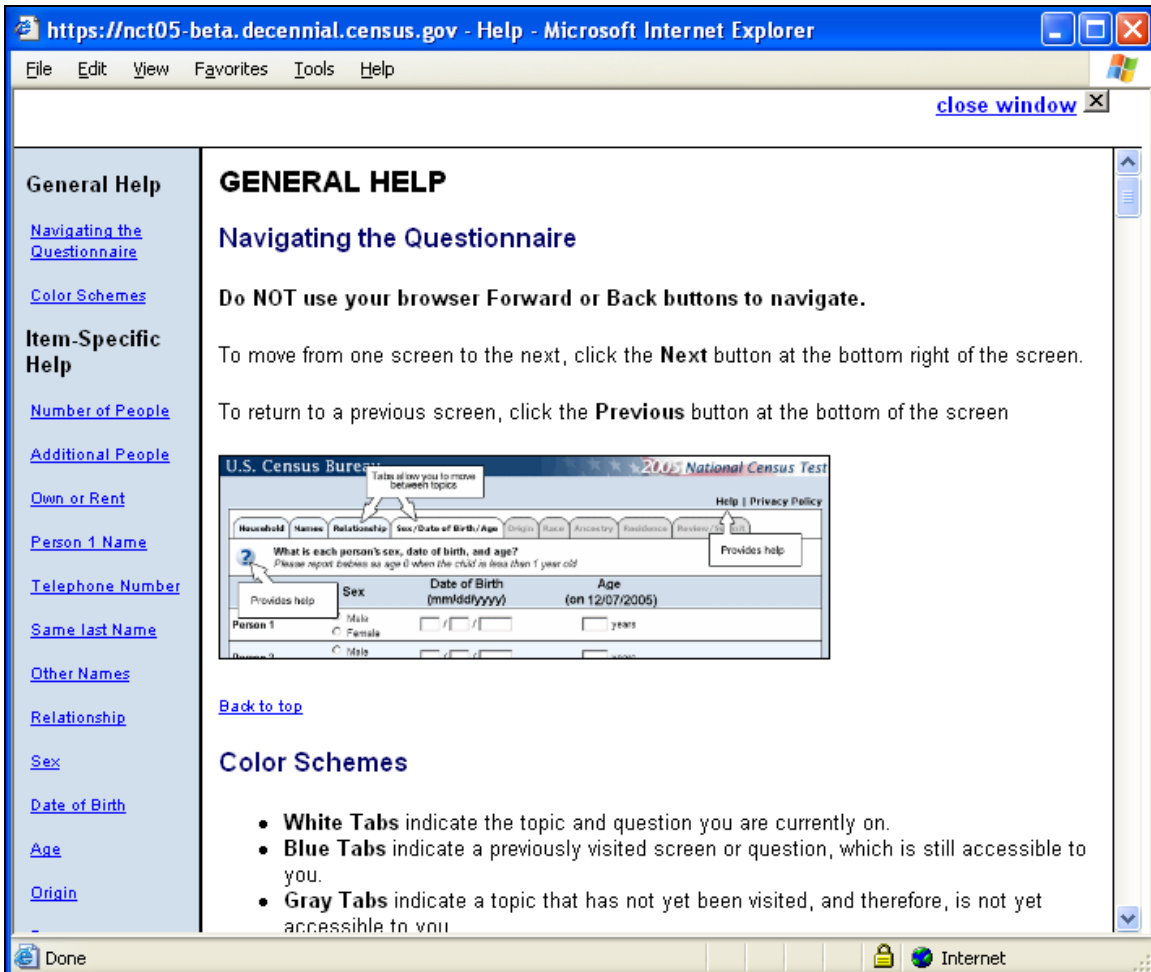


Figure 16. Design of Help text (top of file) -- Version L

Figures 16 and 17 illustrate the approach used in Version L, which was to present help in one, long scrollable file.

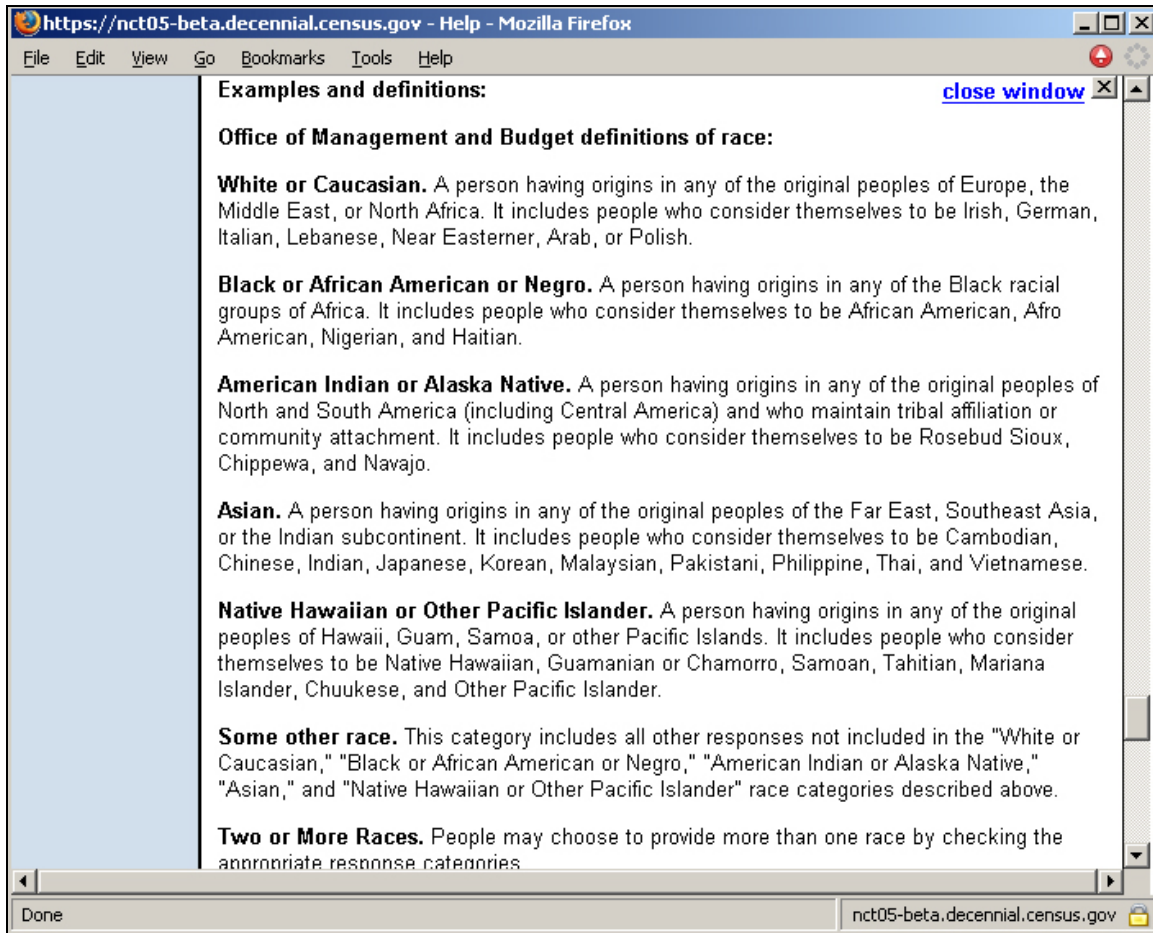


Figure 17. Design of Help text (middle of file) -- Version L

The following example illustrates a design issue with presenting help in a scrollable file:

One participant was searching for additional guidelines about her son at college, but she began reading in the middle of the section so that she could see only the guidelines on people to exclude. The test administrator prompted the participant to go to help and scroll up. The participant began to read the complete set of guidelines on who to include/exclude in the household and said, “Oh...I didn’t go up far enough.”

Figure 18 shows the approach taken in Version M, which was to make help item-specific. One participant looked to help to find out if the nanny’s origin should be considered Spanish/Latino/Hispanic. She quickly read the Version-M help and found that she should select “yes” for this person (“Rosa”).

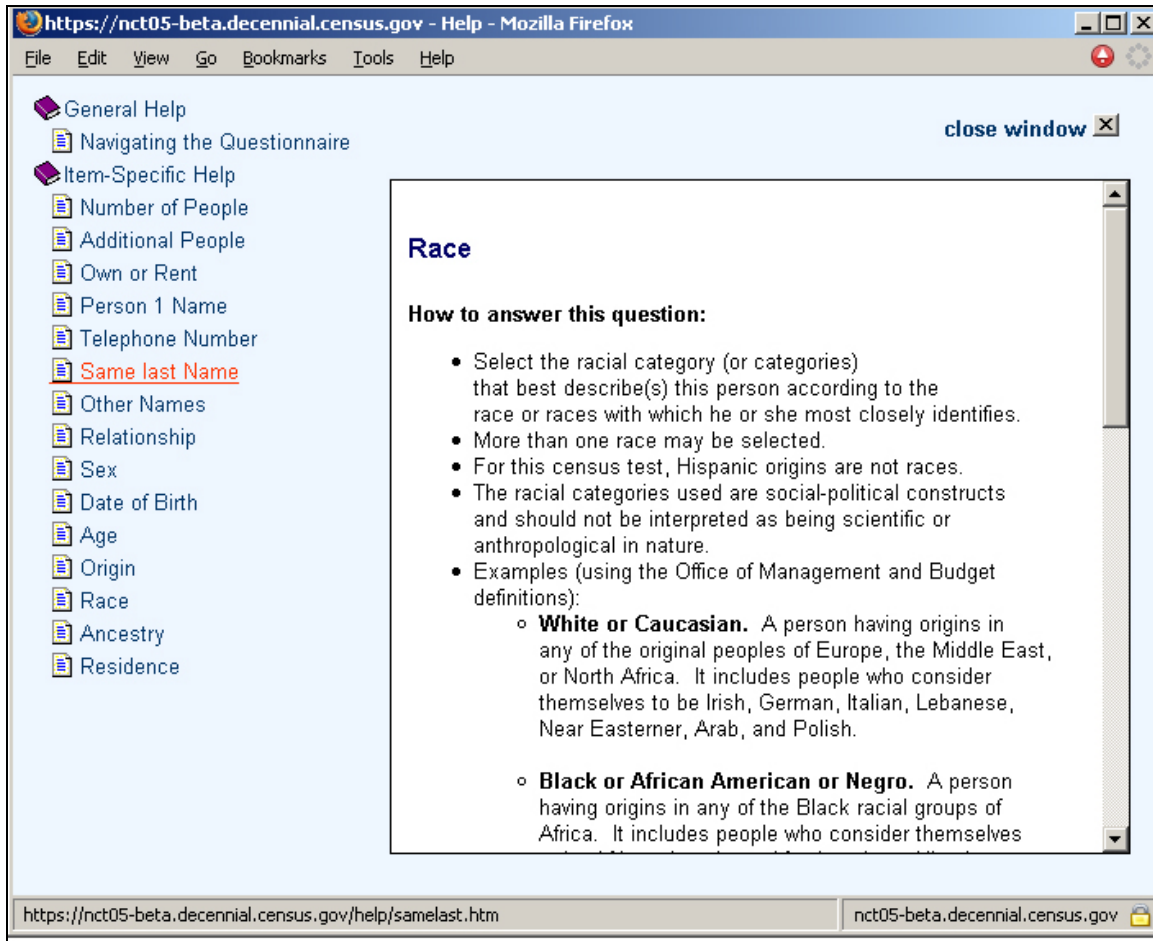


Figure 18. Design of Help Text -- Version M

Design of help text

The help file for Version L was one long scrollable text in the right side of the help window. The left side had links that would take the respondent to specific topics in the file (see Figure 16). However, these links were at the top of the file and scrolled off the screen when the respondent was further down in the file (see Figure 17). For the Version M help file, the links for specific topics were also on the left side but they were always visible (Figure 18). The text appeared in a separate window on the right side of the screen.

Another difference in the design of the two help files involved where clicking on the example link (residence rules, origin or race screens) would take the respondent. In Version L, clicking on the example link anchored the respondent at the point in the help text that showed the examples. In Version M, a similar click brought the respondent to the beginning of the help for that topic. A few participants did not see the examples in the origin help screen because it was just a couple of sentences at the very top of the page followed by white space and the title for the race help text. Because the race title was more eye-catching, participants saw that and seemed to conclude that the application did not take them to where they expected to be.

Recommendation: The Version M design with the links always visible and each topic on a separate screen seemed to work better for the participants. This design, however, may not be Section-508 compliant. Further design and testing is needed. If we do use a scrollable help file, there should be white space between the top of the page and the first line of text that the respondent will read so that participants do not assume that the text is part of the “banner” and ignore it. In any case, when the user is taken to help, the top of the page should display the beginning of the appropriate help content.

Priority: Medium

Accessibility of Help Text

After the help window opened, the correct topic was displayed, but the screen reader vocalized all topic links and the corresponding help text before reaching the specific help text needed. JAWS users heard information on topics unrelated to their help request with the current design.

Recommendation: A separate window is needed for every help topic. Eliminate the index listing and scrolling help window.

Priority: Medium

General use of help

Although participants were encouraged to find help during the scenario interview, many still resisted. For example, when trying to decide the relationship to Person 1 of someone referred to as the nanny, the TA asked if there was a definition of roomer or boarder or some of the other “not related” categories. More often than not, the participant said that help was not available and proceeded to make up his/her own definitions.

We did notice, however, that if the participant went to help and the information answered his/her question, then he/she was more inclined to go to help for other questions. For example, a Version M participant used the Examples link on the Origin screen when entering her data to determine if one of her housemates should be considered Hispanic. The roommate’s mother was a Jewish refugee who grew up in Mexico and speaks Spanish. The participant skimmed the bulleted help and saw “trace family roots to Spanish-speaking country” and “regardless of race” and decided that her housemate was of Hispanic origin. She also used the examples link on the next screen to quickly skim the definitions of the race categories.

One Version M participant was encouraged to find help a couple of times; so when she got to the Hispanic origin question and saw the example link, she clicked on it. She scanned the information and was very pleased with what she read. After that, she tried to go back and find help on the relationship screen. She had a hard time finding how to get help, but once she found it, she scanned the text until she came upon the unrelated category definitions. While she did not find the answer to her question (how to

categorize a nanny), she still seemed interested in the help. Since Version M of the help had the topics listed on the left side, she clicked on a couple of the links to look at the help text for other topics.

Recommendation: We need to further design and test help so that respondents know that they need it and know that it is available (in addition to being easy to use). Perhaps listing the types of residence rules on the screen that may be different from respondent expectations would assist users. This approach would let all respondents know that the Census Bureau has rules that may differ from everyday usage and get them into the habit of seeking help early in the questionnaire.

We also can try putting additional help on the screen (to the extent that it does not bog down readability). For example, the tenure question currently explains “free and clear” with a parenthetical phrase that says “(without a mortgage or loan).” This phrase seemed to help participants understand which category to mark when the TA/Ethel said that she did not have a mortgage but did have a small equity loan. Making similar parenthetical annotations in other places could be helpful to respondents in interpreting the intention of various questions and response options.

Priority: Medium

Help buttons/links

U.S. Census Bureau 2005 National Census Test Help | Privacy Policy

Household Sex/Date of Birth/Age Origin Race Ancestry Residence Review/Submit

Count the people living at your house, apartment, or mobile home using these guidelines...

- Count people where they USUALLY live and sleep.
- For people with more than one place to live, count them where they sleep MOST OF THE TIME.
- [Additional guidelines](#)

? How many people were living or staying in this house, apartment, or mobile home on December 7, 2005?
1 people

? Were there any ADDITIONAL people staying here December 7, 2005 that you did NOT include in the question above?
 No Yes

- Children, such as newborn babies or foster children
- Relatives, such as adult children, cousins, or in-laws
- Nonrelatives, such as roommates or live-in baby sitters
- People staying here temporarily

Figure 19. Household screen showing available help links -- Version L

Help was available from a link at the top of each page (above and to the right of the tabs), via the question mark icon in front of each question, or from selected links (residence rules, origin, race) in the online survey, as shown in Figure 19.

Two participants who used Version L used the question mark icon to access help. Version L had the Before-You-Start screen that informed the participant about the question mark icon being a help link. Others who accessed help did so by using the examples link or the help link at the top of the page.

Generally, help was accessed during the scenario interview at the prompting of the TA/Ethel about whether any additional guidance was available. This sent participants in search of help. On screens such as Relationship and Other Residence, the question and banner and, therefore, the help had scrolled off the screen and were no longer visible. On these screens, the participant looked for help at the response level. In one case, the participant scrolled to the top on the Relationship screen and found help. When she was finished, she entered her answer in the space for Person 2 instead of Person 6. She noticed this when she went to Person 3 and realized she had already answered this question. She corrected her response for Person 1 and scrolled down to Person 6. She obviously did not remember scrolling up to get help and expected that when she returned from help she would be where she needed to be to enter her response. Another participant also remarked that when she came back from help she was at Person 1 instead of Person 6.

Recommendation: If possible, we should try to make the help more obvious for respondents. One way to remove the problem of losing context when scrolling down the page is to lock the top of the screen so that the question (and, consequently the help) is always visible. One participant suggested that the help link at the top of the page be moved to be lower on the page. Putting it in the blue background makes it blend in with the “banner” and, therefore, more susceptible to being overlooked. Perhaps move it into the white space near the tabs. Screen resolution will have an impact on how this looks because, at lower resolutions, the tabs take up more of the screen. Perhaps we should also test a design in which the participant can click on the response or question to get help with it. The challenge would be how to alert the respondent that help is available. A couple of participants understood that the question mark icon was a link to help when they read the Before-You-Start screen.

Priority: Medium

Content of help text


One Version-L participant went to help to see if she should guess at year of birth. She read “how to answer this question” for year of birth, then saw the “why we ask this question” discussion and skipped it (calling it “government mumbo jumbo”). She started reading how to answer this question for age and got to the sentence about counting babies as zero. Since this was all a repeat of what she already knew, she skipped reading the last sentence that actually contained her answer. The point is that there was a lot of text here for her to sift through and she reached her limit before getting her answer. During the scenario interview, when the TA/Ethel said she did not know the age or year of birth of one person, this same participant said she had two choices – she could leave it blank and get the edit message or use help to see what to do. She chose to leave it blank because she said she could not deal with help.

The streamlined design of the Version-M help text met the needs of the participant who was looking for how to classify her housemate’s origin. This design was consistent with the way in which this participant wanted to use the help.

Recommendation: The help text must be revamped to make it relevant and to answer questions that respondents actually have. We also recommend writing it in a style suitable for the web (e.g., short phrases and bullet formatting). Just the quick effort made for this test seemed to benefit the participants who used help. The help should be short and to the point to assist respondents toward their goal of finishing the form as quickly as possible.


Priority: Medium to High

Edit messages

 There is a problem with the following:

- John - Please select male or female.
- John - Please enter the date of birth as a two-digit month, two-digit day and four-digit year.
- John - Please be sure you entered the age as of December 7, 2005. If you don't know the exact age, please provide an estimate.
- Jane - Please select male or female.
- Jane - Please enter the date of birth as a two-digit month, two-digit day and four-digit year.
- Jane - Please be sure you entered the age as of December 7, 2005. If you don't know the exact age, please provide an estimate.
- Joe - Please select male or female.
- Joe - Please enter the date of birth as a two-digit month, two-digit day and four-digit year.
- Joe - Please be sure you entered the age as of December 7, 2005. If you don't know the exact age, please provide an estimate.


If you have provided your best answer or you don't know the answer, please continue to the next question.



What is each person's sex, date of birth, and age?
Please report babies as age 0 when the child is less than 1 year old.


	Sex	Date of Birth (mm/dd/yyyy)	Age (on 12/07/2005)
→ John	<input type="radio"/> Male <input type="radio"/> Female	<input type="text"/> / <input type="text"/> / <input type="text"/>	<input type="text"/> years
→ Jane	<input type="radio"/> Male <input type="radio"/> Female	<input type="text"/> / <input type="text"/> / <input type="text"/>	<input type="text"/> years
→ Joe	<input type="radio"/> Male <input type="radio"/> Female	<input type="text"/> / <input type="text"/> / <input type="text"/>	<input type="text"/> years

Figure 20. Sample Edit Message -- Version L

 Please check the following:

- Mike - Please select male or female.
- Mike - Enter the date of birth.
- Mike - If you don't know the exact age, enter your best guess.
- Mary - Please select male or female.
- Mary - Enter the date of birth.
- Mary - If you don't know the exact age, enter your best guess.
- Mark - Please select male or female.
- Mark - Enter the date of birth.
- Mark - If you don't know the exact age, enter your best guess.

If you don't know the answer or have provided your best answer, continue to the next question.

 **What is each person's sex, date of birth, and age?**
Please report babies as age 0 when the child is less than 1 year old.

	Sex	Date of Birth (mm/dd/yyyy)	Age (on 12/07/2005)
→Mike	<input type="radio"/> Male <input type="radio"/> Female	<input type="text"/> / <input type="text"/> / <input type="text"/>	<input type="text"/> years
→Mary	<input type="radio"/> Male <input type="radio"/> Female	<input type="text"/> / <input type="text"/> / <input type="text"/>	<input type="text"/> years
→Mark	<input type="radio"/> Male <input type="radio"/> Female	<input type="text"/> / <input type="text"/> / <input type="text"/>	<input type="text"/> years

Figure 21. Sample Edit Message -- Version M

In both versions, the data suggest that for the scenario, many participants read the entire edit message or the key part of the message. Very few participants did not read to the end of the message. Fewer edit messages were triggered while participants were answering for their own personal information and we see a similar pattern. From the logs there does not appear to be a large difference in how participants react to the edit messages between the two versions, with about the same number of participants having a negative response to an edit message and participants continuing without resolving the edit failure.

Introductory statement wording

In contrast with the first round of usability testing, no participant in this round reacted negatively to the initial statement in the Version-L edit message, which said there was a “problem” (see Figure 20). Similarly, no one reacted negatively to the initial statement in the revised message asking him/her to check the data (see Figure 21).

Recommendation: Further testing is needed to see if the revised wording is acceptable.

Priority: Medium

Specific nature of edit message

The revised, Version-M edit messages that were more specific about the problem helped participants when it was unclear what the problem was. For example, the last name of the household in the scenario was O’Brien. The system, however, does not accept special characters, such as an apostrophe. The edit message in Version L asked the participant to verify the name and make necessary corrections whereas the message for Version M specifically said the field did not accept special characters. It is possible that the participants did not know what was meant by the phrase “special characters” as well.

In one example of a Version-M participant responding to an edit message, an edit message appeared after a participant entered the last name of “O’Brien.” After reading the phrase, “special characters,” he correctly concluded that the problem is the apostrophe. The participant then removed all the apostrophes from the names and made the B lowercase before continuing, in effect doing some of the work that the computer would have done for him.

Even though they were computer literate, a couple of Version-M participants did not know what a “special character” was. Use of this technical terminology could lead to unnecessary confusion and measurement error.

One participant asked the TA if the spelling was correct. When asked what he thought “special characters” meant, he replied, “the person does not exist.” He reported that the reason for the edit is that the person may not be present in the census records (implying that we are checking names against a master list as they were entered) or that the data provided did not match with the name. When asked what to do next, the participant decided to change Person 1 to Ethel’s husband. This confusion could be avoided by using non-technical terminology instead of the phrase “special characters.”

When the same message appeared on the Names screen, this participant said it meant extra characters are not needed and removed the apostrophe. Other participants who did not know what a “special character” was deduced that it was the apostrophe and removed it. [Note: Better not to risk unnecessary confusion. Use commonly understood wording.]

On the other hand, Version-L participants (without the specific instruction about special characters) looked at the name with the apostrophe, decided it was correct, and continued. If they noticed on the Names screen that the apostrophe was removed, they added it to the name again. Again, they got the edit message to verify the name and again they continued. (Each time the participant re-entered the apostrophe, the application accepted no change and moved on because these are just soft edits.) When these participants got to the Review/Submit page and the apostrophe was again removed, they went back to the Names screen and once again added it back into the name. It was not until they got back to the Review/Submit page and the apostrophe was not part of the name that they realized the system was not going to allow an apostrophe. This showed persistence on the part of the participants, but this level of confusion could be avoided by stating the problem more clearly in the edit message.

Recommendation: To the extent possible, edit messages should be crafted so that they are informative to the respondent about what, exactly, the problem is.

Priority: Medium

Length of edit message

If participants read the edit messages at all, they did not read the entire message, whether it was the messages in Version L or Version M. Several participants read the first couple of words and, if they got the gist of the message, they attempted to fix the problem. In the scenario, the TA/Ethel said that she did not know the year of birth and age of Person 5 and did not know the year of birth for Person 6 but guessed at an age. The edit message for Version-M participants was shorter and more specific than was the message given to Version-L participants. In Version M, the message said, “enter year of birth,” while the Version-L message said, “please enter the date of birth as a two digit month, two digit day and four digit year.” Despite receiving shorter, more specific edits, the Version-M participants may have found the box with several edit messages to have been overpowering because they could have potentially received several different edit messages in a row instead of just one single message.

As with problems caused by skimming the residence rules, we noted that skimming the edit message for age and date of birth could be problematic. Several participants scanned the edit message and saw that they could guess at the age. Therefore, once they had an age, they calculated a year of birth (as discussed previously), counter to the actual rules.

Recommendation: Although the edit messages for Version M were shortened, we think they still need to be even shorter. The point of the message needs to be conveyed in the first couple of words.

Priority: High

“Continue” message

Several participants did not get to the statement at the bottom of the edit message telling them that they could continue if they did not know the answer or had provided their best answer. They assumed that when they clicked the Next button and went to the next screen, the system must be satisfied with the answer they provided regardless of whether they fixed an edit correctly (or did not even attempt to fix it). One participant who had seen edit messages on several screens finally read the complete edit message on the Sex/Date of Birth/Age screen. She said, “Oh, that screen was just to make us stop and think about what we put in, but it didn’t stop us.”

Recommendation: The statement that informs the respondent that he/she can continue if he/she has provided his/her best answer or does not know the answer, still needs to be more prominent. It needs to stand out from the content of the edit messages. Perhaps some use of highlighting (e.g., bolding or all capitals) would work to draw attention to it. The caution here is that only short phrases of two or three words should be highlighted. If too much is highlighted, the purpose of highlighting is lost.

Priority: Medium to High

Anchoring of edit messages

The household screen contained several questions on a page that scrolled. Edit messages appeared above each question on this page. This meant that, in Version L, when the page refreshed with an edit message for a question “below the fold,” the message was not visible without scrolling. Assuming that they were reading something new, several Version-L participants just started reading the residence rules at the top of the household page again until they finally realized they had already been to this page. One participant remarked, “didn’t we just do this?” She scrolled down saying that she wanted to hit Next. She briefly saw the edit message while scrolling down to get to Next, but she did not stop to read it. She just hit Next.

Please enter a name. If you don't know this person's name, continue to the next question.

→ If an owner or renter of this house or apartment lives here, enter the person's name below. If an owner or renter does not live here, enter the name of any adult living here. This is Person 1. What is Person 1's name?

First Name MI Last Name

Please provide a telephone number. If you don't know the telephone number, continue to the next question.

→ What is Person 1's telephone number?

Area Code +Number

() - .

Does everyone in this household have the same last name?

Yes

No

Next >>

Respondents are not required to respond to any information collection unless it displays a valid approval number from the Office of Management and Budget (OMB). This test has been approved and given the following OMB number: 0907-0725.

USCENSUS BUREAU

Figure 22. Version M -- Anchoring of edit message on household page.

Version M anchored the page so that the question with the edit message appeared at the top (see Figure 22). For the most part, Version-M participants saw the edit message. A couple of participants did not actually see the message but, instead, saw the shading in the answer box. We think that this happened because the message was at the very top of the page and may have blended into the banner.

Recommendation: Anchor the household page at the point of the edit message so that it is easier for participants to see. Insert some white space between the top of the page and the edit message so that the message does not blend in with the banner.

Priority: High

Use of Color and Contrast in the User Interface Design

The screenshot shows a web browser window titled "U.S. Census Bureau - 2005 National Census Test - Microsoft Internet Explorer". The page header includes "U.S. Census Bureau" and "2005 National Census Test". A navigation bar contains tabs for "Household", "Names", "Relationship", "Sex/Date of Birth/Age", "Origin", "Race", "Ancestry", "Residence", and "Review/Submit". The main content area features a question: "What is each person's sex, date of birth, and age? Please report babies as age 0 when the child is less than 1 year old." Below this is a table with four rows, each for a person. The table has three columns: "Sex", "Date of Birth (mm/dd/yyyy)", and "Age (on 09/15/2005)".

	Sex	Date of Birth (mm/dd/yyyy)	Age (on 09/15/2005)
Person 1	<input checked="" type="radio"/> Male <input type="radio"/> Female	<input type="text"/> / <input type="text"/> / <input type="text"/>	<input type="text"/> years
Person 2	<input type="radio"/> Male <input checked="" type="radio"/> Female	<input type="text"/> / <input type="text"/> / <input type="text"/>	<input type="text"/> years
Person 3	<input type="radio"/> Male <input type="radio"/> Female	9 / 9 / 9999	<input type="text"/> years
Person 4	<input type="radio"/> Male <input type="radio"/> Female	05 / 04 / 1975	30 years

Navigation buttons: "<< Previous" and "Next >>".

Disclaimer: Respondents are not required to respond to any information collection unless it displays a valid approval number from the Office of Management and Budget (OMB). This test has been approved and given the following OMB number: xxxxx-xxxx.

USCENSUSBUREAU

Figure 23 -- Production application screen: Dark-blue, horizontal bar with capitalized column headings and white letters in the bar

Response category labels

Observations made in Round 1 of usability testing indicated that participants did not see labels in the dark blue horizontal bar, separating the question from the responses on the topic-based application (see Figure 23). The alternative design, illustrated in Figure 24, lightened the blue bar and improved the visual contrast by changing the font to black. It appeared that more Round-2 participants read the labels than Round-1 participants did. For example, Round-2 participants would say their name (as displayed in the left column of the response section), read the label aloud, and then continue up to read the question.

Of course, this is not the order in which we expect people to read. We expect them to start with the question. By observation, however, the respondent's attention is drawn to the names in the left column, especially to their own name. In trying to understand what they are being asked, they then look up slightly and perceive the labels in the lighter blue bar, and then move up a little further to the question. Our general impression is that

respondents look first at the response options and look at the question (and any instructions) only if what they are being asked is not apparent to them. This observed behavior runs counter to expected respondent behavior as modeled in the research literature (Tourangeau, Rips & Rasinski, 2000). Our observation suggests that anything we want the respondent to read should be displayed close to, perhaps in parallel with, the response options.

The screenshot shows a web form from the U.S. Census Bureau for the year 2005. At the top, there is a light-blue horizontal banner with the text "U.S. Census Bureau" on the left and "2005" on the right. Below the banner is a navigation bar with buttons for "Household", "Sex/Date of Birth/Age", "Origin", "Race", "Ancestry", "Residence", and "Review/Submit". The main question is "What is each person's sex, date of birth, and age?" with a sub-instruction: "Please report babies as age 0 when the child is less than 1 year old." Below the question is a table with three columns: "Sex", "Date of Birth (mm/dd/yyyy)", and "Age (on 12/07/2005)". The first row of the table is for a person named "Sara". Under "Sex", there are radio buttons for "Male" and "Female". Under "Date of Birth", there are three input boxes separated by slashes. Under "Age", there is an input box followed by the word "years". At the bottom left of the form is a button labeled "<< Previous". At the bottom center, there is a disclaimer: "Respondents are not required to respond to any information collection unless it displays a valid approval number from the Office of Management and Budget (OMB). This test has been approved and given the following OMB number: 0607-0725." Below the disclaimer is the text "USCENSUSBUREAU".

Figure 24. Version-L and Version-M screen with light-blue, horizontal banner containing capitalized column labels and black text

Recommendation: Keep the lighter blue bar with the black font for labels.

Priority: High

Make questions more distinctive

In this round of testing, we examined whether lightening the blue bar that separated the question and response influenced how often the participant read the question. In the previous round of usability testing, we noted that participants often missed the instructions associated with a question because they got the gist of the question from the responses and did not read the question area of the screen. We cannot say whether lightening the blue bar helped with this issue or not. The grid design is a strong visual attention grabber. For example, on the Names screen, one participant saw Person 2 in the left column and said that was his mother. He did not read the question until after he entered her name. On the Other Residence screen, another participant said “Joseph” (the name of Person 1 in the left column), “other residence,” and then went up to the question to read it. We can view this as similar to reading the title or subject of a document before reading the text. It gives context. The problem comes in, however, if the participant does not go up and read the question and the other instructions associated with it.

Recommendation: Since the blue bar between the question and the response option is not a required element, it might be worth removing the bar entirely to see if there is any effect. Another suggestion is to move any instructions closer to the response option area so that people are more likely to see them. Further research is needed on the issue of the order in which respondents actually read the elements of a question-and-response area.

Priority: Medium

Navigation

Return to Review/Submit

A button was added for Round-2 usability testing to allow the participant to quickly return to the Review/Submit page. It was located between the Previous and Next buttons and was only visible once the participant had been to the Review/Submit page (see Figure 25). This button was added because some respondents who used the 2005 NCT Internet application said they wanted a hot key to get them back to the Review/Submit screen.

The screenshot shows a survey form with a navigation bar at the top containing tabs for 'Household', 'Names', 'Relationship', 'Sex/Date of Birth/Age', 'Origin', 'Race', 'Ancestry', 'Residence', and 'Review/Submit'. The 'Review/Submit' tab is active. Below the navigation bar, there is a question: 'What is each person's sex, date of birth, and age?' with a sub-instruction: 'Please report babies as age 0 when the child is less than 1 year old.' The form contains three rows of input fields for 'Sex' (radio buttons for Male and Female), 'Date of Birth' (mm/dd/yyyy), and 'Age' (on 12/07/2005). The first row shows 'Male' selected, '01 / 01 / 1970', and '35 years'. The second row shows 'Female' selected, '02 / 02 / 1970', and '35 years'. The third row shows 'Male' selected, '03 / 03 / 1990', and '15 years'. At the bottom of the form, there are three buttons: '< Previous', 'Return to Review/Submit >>', and 'Next >>'. Below the buttons, there is a disclaimer: 'Respondents are not required to respond to any information collection unless it displays a valid approval number from the Office of Management and Budget (OMB). This test has been approved and given the following OMB number: 0607-0725.' and the USCENSUS BUREAU logo.

Figure 25. Return to Review/Submit button added for ease of navigation.

Most test participants saw the button and used it to get back. One participant hit the Previous button. To her, “previous” apparently meant the preceding screen she had been on, not the arrangement of screens in the application. When that did not take her where she wanted to be, she used the Review/Submit tab. Another participant hit the Next button several times. She seemed to stop and think that there had to be an easier way to do this, saw the Return to Review/Submit button and used it.

Recommendation: Keep the Return to Review/Submit button. Perhaps display it in different colors so that it pops out at respondents.

Navigating using the Previous button

The application was designed such that navigating away from a page that had missing or “unacceptable” data brought up the edit messages. However, this meant that if a respondent got to a page and wanted to go back to the previous page and look at something, the edit messages came up. This confused several participants because it was not what they were expecting. One participant said that since she got an edit message, it must mean that she could not go back to the previous page.

Recommendation: Allow the respondent to go back to the previous page by hitting the Previous button (or clicking on a tab) without invoking the edit messages.

Priority: High

Satisfaction Results

In general, participants using both Version L and Version M gave extremely high ratings for satisfaction with aspects of the topic-based application. Looking at the individual ratings, we found just a sprinkling of ratings below the scale mid-point: there were four ratings of 4 for getting started; one rating of 3 for information displayed on the screens; and one rating of 4 for making changes to answers. Table 3 includes the average rating for the different components measured.

Table 3. Satisfaction Ratings on a 9-Point Scale, where 1 = extremely unfavorable and 9 = extremely favorable (from modified QUIS, N = 18)

1	Overall reaction	7.78
2	Getting started	7.44
3	Screen layouts	8.11
4	Information displayed on the screens	7.89
5	Overall experience of entering info	8.28
6	Order of presenting questions	8.61
7	Going back to previous questions	8.17
8	Making changes to answers	7.56
9	Use of terminology	8.33
10	Content of edit messages	8.00
11	Accessing help	7.94
12	Content of help	8.19
13	Presentation of edit messages	7.89

The means indicated a high level of satisfaction across this group of participants. They reflect the fact that all participants were successful in completing the online census form using both their own data and the scenario data, as supplied by the TA/Ethel.

However, these satisfaction results do not match the session logs that showed that there was quite a bit of confusion expressed by participants while entering data. This was the case in both versions of the instrument. Table 4 provides the average number of instances of confusion, frustration, and need for assistance from the test administrator by version by data entry task. Thus, participants did not seem to take their own confusion into account when rating their satisfaction with the questionnaire.

Table 4. Comparison of Version L and Version M on Confusion, Frustration, and Need for Assistance from the Test Administrator in Round-2 Usability Testing. (Values given are means of frequency counts.)

Version L (n=9)	Confusion	Frustration	Need for Assistance from TA
Personal information	4.56	0.00	0.22
Scenario	9.00	0.22	0.00 (not applicable)
Version M (n = 9)			
Personal information	2.33	0.00	0.22
Scenario	5.22	0.56	0.00 (not applicable)

Summary

Because preparations for testing leading up to the 2020 Census are underway, this report, which had remained in draft form for years, has become relevant. Issues with questions such as how to display instructions for residence rules, how to display and navigate between edit messages and the data of interest, and how to organize help are again being discussed. The findings and recommendations in the report are still relevant to future versions of the census questions.

References

Couper, M. (2005). Personal communication dated October 14, 2005.

Dillman, D.A., Smyth, J.D., and Christian, L.M. (2009). *Internet, Mail, and Mixed Mode Surveys – The Tailored Design Method*. Hoboken, NJ: John Wiley & Sons, Inc.

Murphy, E.D., Ciochetto, S., Malakhoff, L., Coon, D., Ashenfelter, K., & Smith, N. (2013). 2005 Census Internet Prototype Applications: Usability and Accessibility Testing (Round1). Center for Survey Measurement Report, U.S. Census Bureau.

Section508.gov. (n.d). *Section 508 Standards Guide*. Retrieved March 8, 2013, from <http://www.section508.gov/index.cfm?fuseAction=stds>

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The Psychology of Survey Response*. New York, NY: Cambridge University Press.

Appendix A. Screening Questionnaire

What is your name?

Age?

18-25

26-33

34-41

42-49

50-57

58-65

66+

Education?

Less than high school

High school graduate or GED

Technical training

Community college

4-year college

Post-college education (# of years)

Spanish/Hispanic/Latino Origin?

Race? (use OMB categories)

Any disabilities? If yes, please describe briefly.

Use any special settings when at the computer? If yes, please describe custom settings.

Vision?

Wear corrective lenses?

Color blind?

Computer experience?

Number of years:

Frequency of use (# of hours per day):

Internet experience?

Number of years:

Frequency of use (# of hours per day):

Size of household?

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 for ease of use in providing Census data. Your reactions and comments will inform design of a prototype Census Internet form for the 2010 Census.

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. Only staff involved in this design research will have access to the tapes. 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. I certify that I am over 18 and in good health.

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?

(Check 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 some design concepts that we are considering for use in the online form for 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.

[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 were 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; limit = 5 minutes]

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 about.

Do you have any questions?

During this session, you were filling out the Internet census form a couple of times, completing a questionnaire about your experience, and then we'll talk about it.

Part I. OK, now, we'll have you get started. Here's the mailing package you would have received? Please open it and look it over. You have an option to respond by paper or

Internet. Here's the URL for the Internet form. Please enter it and begin with the first screen. Answer the questions for your own household.

[Hand participant the mailing package (Version L or Version M) and the card with the URL.]

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

Part II. Scenario-Based Form Completion

During this part of the session, we're going to do something a little different. I were a friend of yours who needs help completing my Census form online. You will ask me the questions, and I will give you the answers. You will enter my answers into the form.

At end, administer the QUIS. Ask them to think aloud as they do the QUIS so we get an idea of their thinking.

Look over QUIS and pick something rated especially high or low to start a discussion.

Discuss their experience of completing the form for themselves versus for the other household. What worked well; what did not work so well?

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
1 2 3 4 5 6 7 | wonderful
8 9 | not applicable |
| 2. Getting started: | difficult
1 2 3 4 5 6 7 | easy
8 9 | not applicable |
| 3. Screen layouts: | confusing
1 2 3 4 5 6 7 | clear
8 9 | not applicable |
| 4. Information displayed on the screens: | inadequate
1 2 3 4 5 6 7 | adequate
8 9 | not applicable |
| 5. Overall experience of entering information: | difficult
1 2 3 4 5 6 7 | easy
8 9 | not applicable |
| 6. Order of presenting questions: | confusing
1 2 3 4 5 6 7 | clear
8 9 | not applicable |
| 7. Going back to previous questions: | impossible
1 2 3 4 5 6 7 | easy
8 9 | not applicable |
| 8. Making changes to answers: | difficult
1 2 3 4 5 6 7 | easy
8 9 | not applicable |
| 9. Use of terminology throughout the questionnaire: | inconsistent
1 2 3 4 5 6 7 | consistent
8 9 | not applicable |
| 10. Content of edit messages: | informative
1 2 3 4 5 6 7 | confusing
8 9 | not applicable |
| 11. Accessing help: | difficult
1 2 3 4 5 6 7 | easy
8 9 | not applicable |
| 12. Content of help: | not at all helpful
1 2 3 4 5 6 7 | very helpful
8 9 | not applicable |
| 13. Presentation of edit messages: | attention grabbing
1 2 3 4 5 6 7 | not noticeable
8 9 | not applicable |
| 14. Length of time to complete the form: | seemed very long
1 2 3 4 5 6 7 | seemed very short
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 were asked about during the debriefing. Any probes based on observation of the participant were part of the debriefing. The areas we want to probe about are listed below:

1. Did you notice the tabs, color indicators, etc. How about the Return to Review/Submit button?
2. Did you think it was okay if someone marked more than one race.
3. What did you think of the help? Helpful?
4. What did you think of the edit messages? When an edit message appeared, how certain were you that you knew what the problem was and how to correct it?
5. Was it easier to enter your own data or the data for your friend? Why?
6. Compare the experiences of entering your own data versus entering data for your friend. What worked well in each time through the form? What did not work as well?

Appendix G -- Scenario Data

Scenario for R2 Testing

Question:	Response:
Popcount	The following people live in the household: <ul style="list-style-type: none"> ➤ Respondent ➤ Husband ➤ Son of husband by first marriage, in college (if asked, still lives at home) ➤ Newborn son ➤ Mother of husband ➤ Nanny
Tenure	Paid off mortgage but have small home equity loan
Name	P1: Respondent – Ethel G. O’Brien P2: Husband – Fred G. O’Brien P3: Son #1 – Sam G. O’Brien P4: Son #2 – Adam G. O’Brien P5: Mother-in-law – Natasha G. O’Brien P6: Nanny – Rosa H. Lopez
Relationship	(encourage to go to help for P6)
Sex/DOB/Age	P1: 03/23/1973 P2: 03/23/1963 P3: 03/23/1985 P4: 03/23/2005 (change age = 1) P5: 03/23/?? DK age P6: 03/23/?? Age = 20
Origin	Nanny is from Argentina
Race	P1: white P2: white P3: white and American Indian P4: white P5: white P6: Hispanic
Ancestry	Ask – how many generations should I go back? P1: German, Hungarian P2: English, Scottish P3: English, Scottish, Cherokee P4: English, Scottish, German, Hungarian P5: English P6: Spanish
Residence	P3: college student (living at home) P6: lives with a friend on her days off

**Appendix H Behavior Codes for Logging Test Sessions -- 2005 Internet Form
(Round 2)**

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.
Emphatic reading (E)	Participant emphasizes or stresses words in all caps when reading aloud.
Skip reading question (S)	Participant goes right to the response options without reading the question.
Read whole question (W)	Participant reads the whole question.
Read part of question (Q)	Participant reads just part of the question.
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? Indicators of confusion are a puzzled look, a long pause in actions or in thinking aloud, a statement about being confused; no emotional content, just baffled. (If there is emotional affect, code it as frustration.)...
Screen navigation problem (V)	Participant has problems navigating between screens. Navigating is inefficient. This includes using the browser back button.
Need assistance from administrator to continue (D)	Note that administrators assist as a last resort. Thus, other behavior codes may be logged for this situation.
Left blank (O)	Participant left an item blank for any reason; record the reason if it is evident
Technical/System issue (T)	A server error or other software issue occurs.
Reading edit message (Y, N, P)	Participant does (Y), does not (N) read to the end of an edit message or reads part of the message, the key information, but not the whole message (P)
Negative response to edit	Participant is served an edit message and they don't understand the

message (G)	edit message, don't see it, say they don't need to read it, don't know what to do, etc.
OK to continue (K)	Participant asks whether it's OK to continue without resolving an edit message or participant comments that it's OK to continue (logger indicates whether it's a question or comment)
Continue with edit unresolved (X)	Participant leaves data unchanged and goes on to the next screen.
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.
Reading help content (H for Yes, Z for No)	Participant reads enough to find the answer (Y/N)
Stop reading help content (B)	Participant stops reading help and goes back to the questionnaire
Frustration (U)	The following are indicators of frustration: heavy sighs, not moving forward in the form, looking around (away from display), twisting in chair, saying that he/she feels frustrated, colorful language. Frustration has emotional content (emotional affect), whereas confusion does not.
Use of Residence Rules (R)	Participant looks for and reads residence rules (i.e., who to count or not count as a member of the household)