

RESEARCH REPORT SERIES
(*Survey Methodology* #2020-02)

**Usability Testing Results Evaluating the Decennial Census Race and Hispanic Origin
Questions Throughout the Decade: 2012-2020**

Erica L. Olmsted-Hawala
Elizabeth M. Nichols

Center for Behavioral Science Methods
Research and Methodology Directorate
U.S. Census Bureau
Washington, D.C. 20233

Report issued: June 26, 2020

Disclaimer: This report is released to inform interested parties of research and to encourage discussion of work in progress. Any views expressed are those of the authors and not those of the U.S. Census Bureau. The paper has been reviewed for disclosure avoidance and approved under CBDRB-FY20-178

Abstract

The decennial census questionnaire includes basic demographic questions about each person living in a household. Until 2010, the questionnaire had been on paper. Soon after 2010, the Census Bureau began planning for an online 2020 census questionnaire. In anticipation of that, there were multiple field tests of the online decennial form, and before each field test, there was usability testing. Usability testing involved a small number of participants pretesting the online questionnaire prior to the larger field tests. The sessions were one-on-one where each participant was asked to fill out the online questionnaire with answers that pertained to their real lives, while being observed by a researcher. The usability team assessed how the questions and the interface design worked with respect to that task.

Across the different decennial tests, the Census Bureau researched whether there were more effective ways to word and display the various questions and response options in the online survey. The usability team was interested in how the modifications to the questions, including wording of questions, instructions, response options and edit messages, as well as the layout of the questions on the screens, impacted question comprehension and the resulting answers real users gave.

This report shares the usability findings of the race and Hispanic origin questions across all rounds of the English language usability testing. One of the recurring findings was that combining the race and Hispanic Origin topics into one question was less confusing for participants than when those two topics were collected as separate questions on separate screens. Another recurring finding was how confusing and difficult it was to collect consistent detailed origin information, regardless of design. The third consistent finding was how the text used for this topic (whether in the question, in the instructions, or in the edit message) did not seem to have a great effect on how participants understood and answered these topics. While not all of the recommendations from the usability testing made it into the final version that was used on the 2020 Census questionnaire, this research provides a starting point for designing the race and Hispanic origin question for the next census.

Keywords: race, decennial census, usability

Suggested Citation: Erica L. Olmsted-Hawala, Elizabeth M. Nichols. (2020). **Usability Testing Results Evaluating the Decennial Census Race and Hispanic Origin Questions Throughout the Decade: 2012-2020.** *Research and Methodology Directorate, Center for Behavioral Science Methods Research Report Series (Survey Methodology #2020-02).* U.S. Census Bureau. Available online at <http://www.census.gov/content/dam/Census/library/working-papers/2020/adrm/rsm2020-02.pdf>

Table of Contents

1	Introduction.....	6
2	Race and Hispanic Origin Questions tested in the 2010 Census and in Decade leading up to the 2020 Census.....	6
3	Background on the Online Census Form	10
4	Usability Testing Methods.....	10
4.1	Participants.....	11
4.2	Devices.....	13
4.3	Procedure	13
4.4	Analysis.....	15
5	Design of Screens, Findings and Recommendations	15
5.1	Design of the 2012 National Census Test Race Screens.....	15
5.2	Usability Findings of the 2012 National Census Test Race Screens	18
5.2.1	Detailed race open-text fields closer to the race categories were answered more quickly than open-text fields at the bottom of the screen, which was farther away from the selected race category.	18
5.2.2	Error messages encourage participants to enter a specific detailed origin, even when they do not have strong ties to detailed origins, or do not know the detailed origin	18
5.2.3	Predictive text did not work as users anticipated	19
5.3	Design of the 2014 Census Test Race Screens	20
5.4	Design of the 2014 Census Test Race Error Messages.....	24
5.5	Usability Findings of the 2014 Census Test Race Screens	27
5.5.1	African American as race and as detailed origin	27
5.5.2	Separating Hispanic origin and race is confusing	28
5.5.3	Combined race and ethnicity question worked for participants.....	28
5.5.4	Feedback on instructions for selecting more than one race	28
5.5.5	Less initial item nonresponse when detailed origin is asked on its own screen, than when it is collected on the main race screen.....	28
5.5.6	Error messages could confuse users who are not expecting or not familiar with error messages	29
5.5.7	Error messages encourage participants to enter a specific detailed origin, even when they do not have strong ties to that detailed origin, or do not know the detailed origin.....	29
5.5.8	Debriefing feedback on error messages	31
5.6	Design of the 2015 Census Test – Optimizing the Race Screen for Mobile Devices.....	32
5.7	Usability Findings of the 2015 Census Test –Optimizing the Race Screen for Mobile Devices R1 and R2	35
5.7.1	Combined race and ethnicity question worked for participants.....	36
5.7.2	Some participants understood “origin” to mean “National origin” and not “Hispanic origin”	36
5.7.3	Issue with the abbreviation of African Am	37
5.7.4	Feedback on instructions for selecting more than one race	37
5.7.5	Race examples caused user comments.....	37
5.7.6	Wide variability in reporting detailed origins	38
5.7.7	Confusion with when to use “American”.....	38
5.7.8	Error messages encourage participants to enter a specific detailed origin, even when they do not have strong ties to that detailed origin, or do not know the detailed origin.....	39

5.7.9	The use of blue font to highlight certain words or parts of the question caused user problems	39
5.7.10	Gathering detailed origins on multiple screens caused additional respondent burden.....	40
5.8	Design of the 2015 National Content Test Race Screens.....	40
5.9	Usability Findings of the 2015 National Census Test Race Screens	48
5.9.1	Feedback on the MENA category	48
5.9.2	Confusion with when to use “American”.....	50
5.9.3	Feedback on question stem “race or ethnicity” versus “categories”	50
5.9.4	“Enter details” instruction on Race screen is confusing	51
5.9.5	Inconsistencies with “And so on” and “etc.”	51
5.9.6	Feedback on instructions for selecting more than one race	51
5.9.7	Open-text field on the detailed origins screen is not immediately apparent	52
5.9.8	Gathering detailed origins on multiple screens caused additional respondent burden.....	53
5.9.9	Preference on checkboxes vs open-text design (detailed information for race).....	54
5.9.10	Combined race and ethnicity question worked for participants – but a potential of fewer detailed origins for Puerto Rican participants.....	54
5.10	Design of the 2016 Census Test Race Screens	55
5.11	Usability Findings of the 2016 Census Test Race Screens	60
5.11.1	Participants appear to know what the race question is by looking at the response choices. The question stem without the word “race” was not what most expected but also did not cause issues.	60
5.11.2	Feedback on response choices to the combined question on race and ethnicity	61
5.11.3	Preference for African American to be spelled out, but abbreviation did not cause any measurement error	62
5.11.4	Feedback on instructions for selecting more than one race	62
5.11.5	Wide variability in reporting detailed origins	63
5.11.6	African American as race and as detailed origin	63
5.11.7	Use of checkboxes and open-text fields on the detailed race screen	64
5.11.8	Open-text field on the detailed origins screen is not immediately apparent	64
5.11.9	Feedback on the detailed origins question	65
5.12	Design of the 2017 Census Test Race Screens	66
5.13	Usability Findings of the 2017 Census Test Race Screens	67
5.13.1	Feedback on instructions for selecting more than one race	68
5.13.2	Political climate, including Executive Order 13769, potentially influences race reporting	68
5.13.3	Feedback on the detailed race question.....	68
5.13.4	Use of checkboxes and open-text fields on the detailed race screen	69
5.13.5	Error messages encourage participants to enter a specific detailed origin, even when they do not have strong ties to that detailed origin.....	69
5.13.6	Open-text field on the detailed origins screen is not immediately apparent	69
5.14	Design of the 2020 Census Race Screens	69
5.14.1	Design of the 2020 Census Usability Test Round 1 Hispanic origin and Race Screens.....	70
5.14.2	Design of the 2020 Census Usability Test Round 2 Hispanic origin and Race Screens.....	75
5.14.3	Design of the 2020 Census Usability Test Round 3 Hispanic origin and Race Screens.....	75
5.14.4	Final Design of 2020 Hispanic origin and Race Screens	78
5.15	Usability Findings of R1 2020 Census	78
5.15.1	Separating Hispanic origin and race is confusing	79
5.15.2	Open-text field on the detailed origins screen is not immediately apparent	79
5.15.3	Race Error May Cause Measurement Error	81

5.15.4	Challenges when Reporting Race as a Proxy.....	81
5.15.5	Spacing and formatting issues appearing on race screens.....	82
5.16	Usability Findings of R2 2020 Census	82
5.16.1	Challenges when Reporting Race as a Proxy.....	83
5.16.2	African American as race and as detailed origin	83
5.16.3	Separating Hispanic origin and race is confusing	83
5.17	Usability Findings of R3 2020 Census	83
5.17.1	Separating Hispanic origin and race is confusing	83
5.17.2	Spacing and formatting issues appearing on race screens.....	84
5.17.3	Feedback on instructions for selecting more than one race	85
6	Summary and Discussion.....	85
7	Acknowledgements.....	90
8	References.....	90
	Appendix A. General Device and Equipment used in Usability Tests 2012-2020	92
	Appendix B. Details of Smartphone, Tablet and Personal Computers used during testing 2012 – 2020...	93

1 Introduction

The U.S. Census Bureau conducts a census of the U.S. population once every 10 years, and from the beginning, some version of a race/ethnicity question has been on every census form. Each decade the data collection methodology is adapted based on current technology. This paper provides an overview of the research from 2010 to 2020 on how the race/ethnicity questions were adapted given the shifts in data collection methodology. More specifically, in the 2010 census, the main self-response mode was a paper questionnaire mailed to residential addresses. The 2020 Census will be the first census in which the main self-response mode is an online questionnaire. The development of an online census response option began shortly after the 2010 Census and involved years of developing, testing and refining the way the census questions were adapted for the online application. The research agenda over the decade included small-scale usability tests along with large-scale field tests, and testing was iterative; that is, results of one test fed into the design and analysis of the next test.

The Census Bureau's usability lab staff in the Center for Behavioral Science Methods was involved in the small-scale usability tests throughout the decade. The large-scale field tests were conducted in 2012, 2014, 2015, 2016, 2017, and 2018. The formal names of these large-scale tests were the 2012 National Census Test (2012 NCT), the 2014 Census Test (2014 CT), the 2015 Census Test (2015 CT), the 2015 National Content Test (2015 NCT), the 2016 Census Test (2016 CT), the 2017 Census Test (2017 CT). Usability tests were conducted prior to all but the 2018 field test. Usability tests were also conducted prior to the 2020 Census itself.

Usability testing involved a small number of participants pretesting the online questionnaire prior to these larger field tests. Usability testing occurred in one-on-one sessions where each participant was asked to fill out the online questionnaire with answers that pertained to their real lives while being observed by a researcher. The usability team assessed how the census questions and the interface design worked with respect to this task. Usability results and recommendations were provided to the sponsor after each round of testing. Sometimes those findings led to changes in the online questionnaire for the current test, but often the recommended changes were adopted in the next census test. Usability testing was conducted with both English and Spanish-speaking participants after 2012 and in some years, in other languages as well.

2 Race and Hispanic Origin Questions in the Census from 2010 to 2020

The 2010 Census included basic demographic questions about the characteristics of each person living in the household, including race and Hispanic origin based on federal standards set out by the Office of Management and Budget (OMB Standards for the Classification of Federal Data on Race and Ethnicity, 1997). Based on these standards, two distinct questions were asked in the 2010 Census: one on Hispanic origin followed by another on race, with the response categories defined by OMB Statistical Policy Directive No. 15. (See Table 1).

Table 1. 2010 Census Hispanic origin and race questions

2010 Census	
Question Stem	Is Person 1 of Hispanic, Latino, or Spanish origin?
Response Options	<p>No, not of Hispanic, Latino, or Spanish origin</p> <p>Yes, Mexican, Mexican Am., Chicano</p> <p>Yes, Puerto Rican</p> <p>Yes, Cuban</p> <p>Yes, another Hispanic, Latino, or Spanish origin— <i>Print origin, for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.</i></p> <hr/>
Question Stem	What is Person 1’s race?
Response Options	<p>White</p> <p>Black, African Am., or Negro</p> <p>American Indian or Alaska Native—<i>Print name of enrolled or principal tribe.</i></p> <hr/> <p>Asian Indian</p> <p>Chinese</p> <p>Filipino</p> <p>Japanese</p> <p>Korean</p> <p>Vietnamese</p> <p>Other Asian—<i>Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on.</i></p> <hr/> <p>Native Hawaiian</p> <p>Guamanian or Chamorro</p> <p>Samoan</p> <p>Other Pacific Islander—<i>Print race, for example, Fijian, Tongan, and so on.</i></p> <hr/> <p>Some other race—<i>Print race.</i></p> <hr/>

In every census there are experiments embedded with the goal to improve future censuses. In 2010 there was an experiment on alternative ways to ask the race and Hispanic origin questions with the goals of decreasing item nonresponse and increasing reporting in the race and detailed origins screens, while also increasing the accuracy and reliability of the results. This was called the 2010 Census Race and Hispanic Origin Alternative Questionnaire Research (AQE) (Compton, Bentley, Ennis, & Rastogi, 2012). There were multiple variations in the questionnaires tested in the AQE, including four different versions of a single question that eliminated the Hispanic origin question and instead included Hispanic origin as a response option with the list of races. The question stem was also tweaked to include the words “or origin” (e.g., What is [Person1’s] race or origin.). This is known as the “combined question.” One version of the combined questionnaire (referred to as X3) reduced the number of Asian race categories and included a write-in field with instructions to print additional details on the origins after every response option¹. See Table 2 for a display of the

¹ Each race also listed out examples in small italic font. (For details see Compton et al., 2012.)

question wording. Results from the AQE found that across all of the different variations of the combined race and Hispanic origin questions, the combined question had a higher item response rate than when the questions were separate. The AQE also found that the combined question had fewer responses as “Some other race.” When the race and Hispanic origin questions are separate, the “Some other race,” originally intended to be a “small residual category,” became the third largest race group and was marked by the majority of people identifying as Hispanic. Finally, with the ability to write in additional details of race origins, 50% of people who marked White added additional details (e.g., Irish, German) and for people who marked Black or African American, 76% added in additional details (e.g., Jamaica, Nigerian) (U.S. Census Bureau, 2012). The Census Bureau said that the results from the AQE would be the basis for future research on race and Hispanic origin throughout the 2010-2020 decade.

Table 2. 2010 Decennial Census AQE question text on Hispanic origin and race (X3 experimental panel)

2010 Census Race and Hispanic Origin Alternative Questionnaire Experiment (AQE)	
Question Stem	What is person 1 's race or origin? White— <i>Print origin(s), for example, German, Irish, Lebanese, Egyptian, and so on.</i>
Response Options	<p>_____</p> <p>Black, African Am., or Negro <i>Print origin(s), for example, African American, Haitian, Nigerian, and so on.</i> _____</p> <p>Hispanic, Latino or Spanish origin <i>Print origin(s), for example, Mexican, Mexican Am., Puerto Rican, Cuban, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.</i> _____</p> <p>American Indian or Alaska Native <i>Print name of enrolled or principal tribe(s) for example, Navajo, Mayan, Tlingit, and so on.</i> _____</p> <p>Asian <i>Print origin(s) for example, Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on.</i> _____</p> <p>Native Hawaiian or other Pacific Islander <i>Print origin(s) for example, Native Hawaiian, Guamanian, or Chamorro, Samoan, Fijian, Tongan, and so on.</i> _____</p> <p>Some other race or origin <i>Print race(s) or origin(s)</i> _____</p>

Using the results of the AQE as a starting point, in field tests between 2012 and 2018, the Census Bureau researched whether there were more effective ways to word and display the race and Hispanic origin questions with the goal of improving the accuracy of the data. There were variations on formats and layouts, including versions that had not been on the 2010 Census or on the AQE. The census also continued to test the two separate questions format, where the Hispanic origin question was asked on one screen, with the race question asked on the following screen. The main vehicle for testing the different question versions was in the 2015 NCT. There was also a new race category that was added in the 2015 NCT: the Middle Eastern North African (MENA) category. As one of the focuses of this study was to obtain usability feedback on the new MENA category, we over recruited participants who may have had a MENA background so that we could learn how someone who may be of Middle Eastern or North African descent would answer the combined question on race and ethnicity with the new response choice of MENA. For more on the 2015 NCT results see Mathews et al., (2017).

With the intent of reducing item nonresponse, the online questionnaire also had validation checks for missing data programmed on the race and Hispanic origin question and on the combined question on race and ethnicity(s). The design of these checks (called error messages in the document) were also tweaked and retested across tests.

As plans for the 2020 Census were being developed, the Census Bureau needed to make a decision on the whether the testing was sufficient to warrant implementing a combined question on race and ethnicity(s), or whether to revert to the prior 2-question approach. This decision had to be made by December 31, 2017. The Census Bureau decided to continue using two separate questions

for collecting race and ethnicity in the 2018 End-to-End Census Test and 2020 Census. See the decision memo for more information on this (Memorandum 2018.02).

The aim of usability testing of surveys in general is to identify any issues with the design or question wording that causes confusion to users prior to the survey being seen by a larger audience. The intention of this report is to provide a detailed description of how English-speaking participants² interacted with and responded to the variations of the race and Hispanic origin questions, including the combined question, that were usability tested throughout the decade.

3 Background on the Online Census Form

In 2010, the decennial census was a short questionnaire that was designed to take approximately ten minutes to complete. During the test versions of the census throughout the 2010-2020 decade, the standard basic demographic questions consisted of:

- name
- sex
- age and date of birth
- Hispanic origin
- race
- home ownership or rental status
- the homeowner or renter's name
- relationship of the other household members to the owner or renter
- a count of the number of people living or staying in the home
- questions to make sure no one was added or left off the census in error.

These questions were on both the paper and online questionnaires.

In terms of design, with a few exceptions, the online self-response questionnaire included one question per screen. At the bottom of every screen there was a 'next' button to navigate forward to the next screen and a 'previous' button to return to prior screens. The online questionnaire also included other elements, such as error messages if the respondent did not provide an answer to a question, and context-sensitive help, accessible via a link on each screen.

Prior to the 2015 CT, the online version of the questionnaire was not optimized for mobile. Starting with the 2015 CT the questionnaire was designed to render optimally on personal computers (PC), mobile phones and tablets, called "PC and mobile optimization." Optimization means that the screens automatically readjusted in size and in design depending on the display size of the device. On smaller devices, the respondent would *not* have to zoom in to see the questions. The following sections describe the usability testing methods used, how the Hispanic origin and race question(s) were designed in each census test, what the usability issues and recommendations were by each usability test, as well as a final discussion on user issues identified across tests.

4 Usability Testing Methods

² The online questionnaire was field tested in Spanish, Korean and Chinese. CBSM conducted cognitive and usability testing in Spanish. DCMD conducted usability testing in 12 non-English languages.

Over the eight years of testing, three different software systems were used to develop the questionnaires. See Table 3 for specifics on the usability test, when they occurred and the software system that was used in user testing. In the figures in the document, often the overall look and feel differs between census tests and this very often had to do with the software system used to design the online questionnaire.

Table 3. Usability Tests, dates they occurred, by software system

Study Name	Dates of Usability Sessions	System Tested
2012 NCT	6/7 - 6/21, 2012	Centurion
2014 CT	4/2 - 4/10, 2014	Centurion
2015 CT R1	1/5 - 1/15, 2015	Centurion
2015 CT R2	2/19 - 2/27, 2015	Centurion
2015 NCT	7/15 - 7/31, 2015	Centurion
2016 CT	2/7 - 2/17, 2016	Primus
2017 CT	2/28 - 3/7, 2017	ECaSE-ISR
2020 R1	10/18 - 10/30, 2018	ECaSE-ISR
2020 R2	4/9 - 5/2, 2019	ECaSE-ISR
2020 R3	6/5/2019	ECaSE-ISR

4.1 Participants

Across all usability tests from 2012-2020, there were 150 English-speaking participants. The number of participants in any given test ranged from 5 to 43. All participants were familiar with using a laptop/desktop or a smartphone, tablet or all of those devices and had at least one year of Internet experience. For the studies that included testing on smartphones, participants had to have and use a smartphone for a year or more.

Participants were recruited from the metropolitan Washington DC area, Michigan and Illinois. Recruitment efforts included using advertisements on Craigslist.org, personal connections, email blasts distributed to all Census employees, through paper flyers posted at colleges as well as posted at local libraries, through some local neighborhood listservs including “Nextdoor,” and word of mouth. In addition, some participants were recruited using the usability lab recruitment database. A handful of participants were recruited using the intercept method at a library³. Participants identified with a wide range of races and ethnicities. See Table 4 for detailed information about participant characteristics from usability tests run in years 2012-2020.

³ In this method, the task administrator (TA) approached people at a library and asked them if they wanted to complete a 90-minute survey for the Census Bureau and that they would be paid \$60. The TA had her badge and a one-page flier with the census logo on it which explained the interview. The requirements were that the individual agreed to use his or her smartphone or tablet (if a mobile session), the individual was not a federal employee, used the Internet and that the individual had 90 minutes available in which to complete the usability session.

Table 4. English-speaking participant characteristics for usability testing for the years 2012-2020

Participant Characteristics	2012 NCT	2014 CT	2015 CT R1	2015 CT R2	2015 NCT	2016 CT	2017 CT	2020 R1	2020 R2	2020 R3
Sample size	20	11	15	15	43* **	11	10	10	10	5
Age: Mean (SD), range	41 (17) 20-69	41 (10) 22-56	39 (10) 27-57	45 (14) 25-68	36 (13) 19-70	39(13) 24-68	49 (16) 22-70	42 (23) 19-72	31 (16) 19-63	33(17) 21-62
Gender distribution										
Male	7	4	6	6	19	1	2	4	5	2
Female	13	7	9	9	23	10	8	6	5	3
Does not identify	0	0	0	0	1	0	0	0	0	0
Race										
White	10	7	4	7	23	3	8	4	2	1
Black	7	2	11	6	13	6	1	3	4	2
Asian	3	0	0	2	1	1	1	2	4	0
Multi-racial	0	2	0	0	6^	1	0	1	0	2
Hispanic Origin										
Yes	1	1	1	0	9	0	0	0	0	1
No	19	10	14	15	34	11	10	10	10	4
Education										
≤ Some college/ no response	7	3	8	8	7	1	0	4	7	3
Bachelor's degree	6	4	5	2	5	7	4	3	2	1
Post Bachelor's degree	7	4	2	5	11	3	6	3	1	1

* For purposes of this table, race was based on responses to the background questionnaire. However, many of the 2015 NCT participants later identified as MENA when asked during testing

** Education data not available for 20 2015 NCT participants

^Note that five of these participants did not identify with one of the races listed on the background questionnaire and said they fell into the category “Some other race” and wrote in Hispanic.

4.2 Devices

Throughout the decade, the usability tests of the census questionnaires were conducted on various devices including desktop computers, laptop computers, and mobile devices (smartphones and tablets). In some of the studies, participants were required to “Bring Your Own Device” (BYOD) and in other studies the Census Bureau provided the devices. In general, for mobile devices, BYOD was preferred so that we could see how users interacted with the census questionnaire on a variety of devices. However, depending on the software system constraints, such as with whether the questionnaire could be accessed behind the firewall, BYOD was not always possible.

For every session, the test administrator (TA) would audio and video record the session so that after the session the TA could review what had occurred, primarily as a memory aid. The recording captured the audio of the participant and TA and the video of the screen (computer or mobile screen) that had the survey appearing on it, *not* a video of the participants’ face. All of the census surveys were Internet based and consequently required an Internet Browser in order to access the online questionnaire. For these studies, the Internet Browser that was loaded on the government personal computer (PC) was either Internet Explorer (IE) or Firefox. For the BYOD mobile devices test sessions, the user determined which browser to use (they were told to open the one they typically use). For the sessions where government-provided mobile devices were used, the Safari or Chrome browser was used.

Tests included either the desktop computer (pre 2015 tests) with Windows 7 operating system or a laptop computer (post 2015 sessions) again with Windows 7. The laptop was a 13.5-inch Dell Precision M4800 or Dell Latitude E6540. For sessions that included eye tracking, the X120 or T120 Tobii eye tracking hardware and software were used for the desktops and the TOBII X2-60 or the SMI Red 250 were used for laptop or mobile devices. Information in Appendix A includes the type of device(s) used, how the sessions were recorded, the browser that was used to access the Internet survey as well as whether eye tracking was used during the session by usability study across the decade. Appendix B details which type of smartphone, tablet and PC were used during testing.

4.3 Procedure

Each usability session, aside from those in 2012⁴, lasted approximately 90 minutes. To offset the costs of parking and travel, participants received an incentive of \$40 (2012 and 2014) or \$60 (2015-2020) for their participation. Upon arriving in the usability lab or in the remote location (e.g., library) the participant, working one-on-one with a TA was asked to sign a consent form that referenced the pre-approved OMB number, the confidentiality of the session, the voluntary nature of the study, and that the session would be recorded. Once the consent form was signed the recording was started. At that point the participant completed a brief background questionnaire on his/her demographics, computer experience, Internet use, and from 2015 on, mobile device use.

⁴ In 2012, the sessions took approximately 60 minutes.

After completing the background questionnaire, participants were instructed to think aloud⁵ and given a practice task that required them to think aloud. The think-aloud technique was used to understand the participant's cognitive processes as they interacted with the interface. Think aloud is modeled on Ericsson and Simon's (1996) approach to collecting verbal protocols, which is used to maintain a running verbal commentary of the participants' expectations and reasoning.

Some of the usability tests included eye tracking which tracked where the participant looked at the screen during the session. For the usability tests that included eye tracking, the participants' eyes were calibrated by having them follow a red ball as it moved across the laptop screen. At this point the main part of the session began and the participant was told to complete the online Census test while being observed by a TA. Participants were instructed to answer the questions using real-life information (e.g., name, date of birth, race, etc.). The role of the TA was to observe what the participant was doing and give non-intrusive probes (e.g., keep talking) when appropriate.

On the screens related to Hispanic origin, race, and detailed origins, the TA observed and noted any spontaneous comments about the question, response categories, and instructions. The TA took note about the ability to select the desired response choice (if more than one race was selected), the participants' reaction to the error messages (if triggered), and if they left any response fields blank.

After answering the census online questionnaire, participants answered a set of satisfaction questions. None of the satisfaction questions ask specifically about race, Hispanic origin or detailed origins so we do not report the satisfaction ratings here. Towards the end of the decennial testing cycle, participants also answered a short quiz that asked questions about their general census knowledge. One of the questions on this knowledge check asked if they thought someone could select more than one race on the census questionnaire. Another question asked if they could leave a question blank and move onto the next question, both of which pertain to the combined question on race and ethnicity, and detailed origins screens and which we report in this paper.

After the satisfaction questionnaire, participants were asked a set of vignettes and debriefing questions. The vignettes were not related to the combined question on race and ethnicity, or to the detailed origins screens, so are not described here. After the vignettes were completed, there was a final debriefing.

In the final debriefing, the TA showed the participant screen shots of specific screens from the census questionnaire and asked debriefing questions about them. This allowed the TA to capture additional cognitive feedback about the participants' impressions of the questions and any other additional comments they may have about the screen. The questions could be as general as "Did you have any difficulty with this screen? [If yes] What was that?" or more specific. For the screens related to race and ethnicity, there were specific questions based on the question wording, layout and design. There were specific probes if the participant left the detailed origin open-text field blank, if there were spontaneous comments made while completing the screens, and if we had

⁵ Aside from in 2014 when participants answered the census questionnaire in silence.

expected the participant to select more than one race and he or she did not. Depending on the census test, there were also specific follow-up questions about the changes to the screen. At the conclusion of the debriefing, the participant was thanked for his/her time, paid and the session ended.

4.4 Analysis

Immediately after each usability session, the TA summarized the usability and cognitive findings. Analysis of the combined questions on race and ethnicity and the detailed origins screen(s) focused on observations of the sessions (both from real time and from the audio and video recording), spontaneous think-aloud comments, debriefing probes and, when it was available, eye-tracking data. For the 2012 CT, we also had efficiency metrics where we timed participants completing the combined question on race and ethnicity.

Eye-tracking data were analyzed with gaze plots of individual participants and with counts of fixations across participants. A gaze plot shows where an individual looked on the screen and in the order they looked. A fixation is where the eyes were relatively still, and focused on an object (e.g. word, image, etc.) giving it their visual attention. This information told us what parts of the screen and question were attended to and which parts were not.

5 Design of Screens, Findings and Recommendations⁶

In the next section, we share the results and feedback from participants on the race and Hispanic origin screen and/or the combined question on race and ethnicity, as well as on the detailed origins screens from the usability tests.

5.1 Design of the 2012 National Census Test Race Screens

The 2012 NCT tested two different race questions in a split-ballot experiment. Both versions were similar to what had been in the AQE using a single combined question listing “Hispanic origin” as a response option (Compton et al., 2012) however, the 2012 NCT used an online questionnaire. One version had seven checkboxes with an open-text field for details about each main response choice (Figure 1). The other had seven checkboxes followed by a separate question with three open-text fields to include details about the race (Figure 2). Respondents remained on the screen and received an error message if the combined question on race and ethnicity was left blank and the next button selected as seen in Figure 3.

Predictive text was used in the open-text fields, so that when the user would type, specific detailed origins appeared based on the letters that were typed. The objective was to reduce burden on respondents so they would not have to type in the entire race/ethnic origin. The list of preloaded predictive text candidates was extensive and represented anticipated categories, however it was also understood that users could enter free text or additional text that was not in the preloaded list.

⁶ Screen shots contain fictional data only

https://nct2012-butle344.centurion-qa.ssd.census.g...

What is Jane M Doe's race or origin? Select one or more boxes **AND** enter the specific race(s) or origin(s).

White — Enter origin(s), for example, German, Irish, Lebanese, Egyptian, and so on.

Black or African Am. — Enter origin(s), for example, African American, Haitian, Nigerian, and so on.

Hispanic, Latino, or Spanish origin — Enter origin(s), for example Mexican, Mexican Am., Puerto Rican, Cuban, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.

American Indian or Alaska Native — Enter name of enrolled or principal tribe(s), for example, Navajo, Mayan, Tlingit, and so on.

Asian — Enter origin(s), for example, Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on.

Native Hawaiian or Other Pacific Islander — Enter origin(s), for example, Native Hawaiian, Guamanian or Chamorro, Samoan, Fijian, Tongan, and so on.

Some other race or origin — Enter race(s) or origin(s).

Figure 1: 2012 NCT Combined question Version 1: Open-text field after each Category

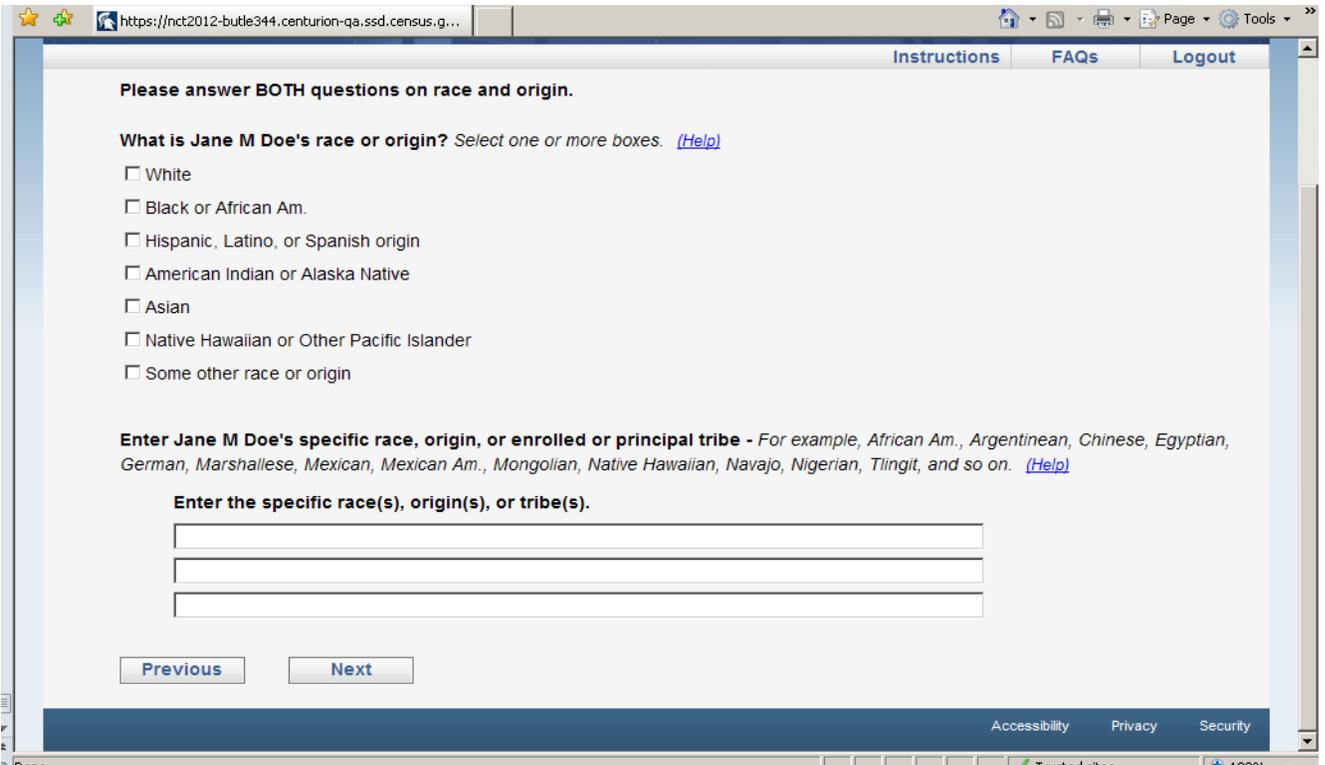


Figure 2: 2012 NCT Combined question Version 2: Open-text fields after all Categories

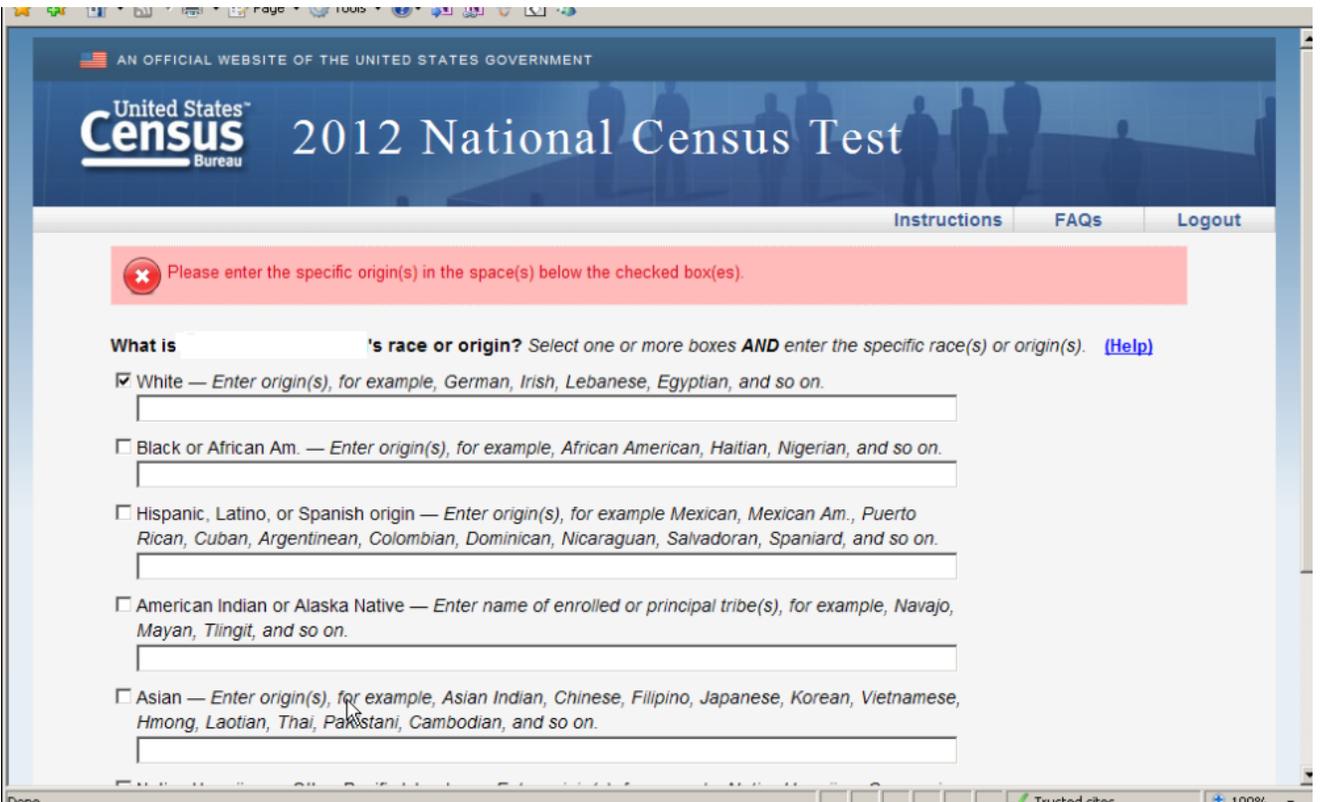


Figure 3: 2012 NCT Combined question Version 1 with error message.

5.2 Usability Findings of the 2012 National Census Test Race Screens

5.2.1 Detailed race open-text fields closer to the race categories were answered more quickly than open-text fields at the bottom of the screen, which was farther away from the selected race category.

On average, participants completed the race question with the open-text fields directly below the race categories slightly quicker than the race question with the open-text fields at the bottom of the screen. For example, for the first person in the household (which would be the first time the participant saw the design), it took on average 49 seconds to answer the question with the open-text field below the screen compared to an average of 55 seconds when the open-text fields were at the bottom of the screen. One possible explanation for the increase in time was that participants had to read more text to understand what was being asked of them in the design when the open-text fields were not near the initial race answer.

5.2.2 Error messages encourage participants to enter a specific detailed origin, even when they do not have strong ties to detailed origins, or do not know the detailed origin

For both versions of the race question, participants selecting “White” often did not initially specify a detailed origin and therefore received an error message when they attempted to go to the next screen in the survey (see Figure 3 for an example of the error message). The error message prompted them to enter a detailed origin. Some participants commented that they initially left the field blank because they did not connect to any detailed origin. For instance, one participant said, “We’ve lived here [in the U.S.] a long time; I’m not sure if it [the question] is for people who have come more recently to the U.S. It’s [the question] not very relevant to my family.” The issue she experienced was not so much with the question layout as with the question of whether she should report detailed origins since she did not feel connected to any specific one. Another participant made a similar comment that “specifying ancestry” seemed like it was more for “people who had not been in the U.S. that long.”

Most participants who selected White initially left the open-text field blank. They said they had no real connection to detailed origins, received the error message, and entered a detail, such as of “European descent.” One participant had seen the example “German” listed and commented that she did not normally see that specificity in this type of a question so, she said, it did help to imply that she needed to put in specific details (she said this only after receiving the error message).

There was one participant who misunderstood the error message. She had initially checked “White” but upon receiving the error message ended up adding the word “Chinese” because she said there was an Asian woman who was a caretaker whom she was close to when she was younger. Some participants who selected “Black/African American” on both versions of the combined question commented that they had to answer the same question twice – they would check the “Black or African Am” response option and then type in “African American” in the open-text field. For example, one participant who had left blank the open-text field said he had to enter African American again. After receiving the error message, he typed in “African American” and then was able to go forward. Another participant, though, who checked “Black/African American” tried to move forward without entering her detailed origins. At the error message screen, she

decided to add in her Caribbean (Montserrat) detailed origin. She also, at this point, said she had Native American heritage on her mother's side. She selected the AIAN race but then, after reading the help text, which explains that a person of AIAN descent should "maintain tribal affiliation or community attachment" decided that she would remove Native American because, while she knew she had Native American in her ancestry, she said she was not affiliated with any specific tribe. During the debriefing, the participant explained, "I took off Native American because I don't think it qualifies, [but] that almost seems wrong cause it is in my history."

So for this participant, the error message worked when it encouraged the participant to be more specific about her African-American detailed origins. Yet it is unclear whether the help text led to a correct answer, as far as her interpretation of the help messaging on the Native American race category.

In response to the user feedback and behavior on the race screens, the team developed additional wording on the error screen that would be used in the next test. This new wording aimed to inform those who do not feel a strong affinity for detailed origins that they would be able to move forward without entering a response. The error message would still pop open on the screen but it would include text in green at the top of the screen that said, "If you would like to provide more detail, please enter the specific origin(s) in the space(s) below the checked box(es). If not, use the 'Next' button to continue."

5.2.3 Predictive text did not work as users anticipated

Sometimes the text that users entered was not among the list of words that was preloaded in the system. For example, a participant who selected Asian as the main race category initially started to type "Chinese" in the open text field, the word "Chinese" popped open below the field, but when he specified "Chinese/Japanese" he noticed that the "Chinese" predictive text went away after he entered the slash. (The list had the term "Chinese" and the term "Japanese," but it did not contain the words "Chinese/Japanese"). The participant asked what was the correct way to "indicate more than one origin?" He also asked if what he did was "wrong." Some participants who selected "White" also entered more than one detailed origin and the predictive text went away after they started entering the second detailed origin as well.

The team recommended testing versions of the question that instruct respondents to use a comma or space for multiple origins and then having the prefills work accordingly. Additionally, the team recommended that common combinations such as "Chinese/Japanese" that were typed by respondents during the 2012 NCT (and all future census field test) into the open-text fields could be programmed into the next survey. We also recommended testing a version of the question that remembers the origin information for the first person in the household if household members were related. The questionnaire would then offer the same prefills in the open-text field if the character string that had just been entered starts to be entered again. For instance, if someone selected "White" and then entered "Hungarian Polish German," the next person might be their biological son or daughter and so the same origins might be entered. Once "Hungarian P" is typed, "Hungarian Polish German" should be in the list of response options for the second person in the household. This recommendation was not implemented for the next round of testing, and from 2016 testing and later, predictive text features for the race and detailed origin screens were

removed. While the intent was to reduce burden, it appeared that the predictive text often caused confusion because there are too many iterations and idiosyncrasies for how respondents enter their detailed origins. For more specifics on the results of the 2012 NCT usability test, see Ashenfelter, Olmsted-Hawala, Lakhe and Malakhoff, 2013.

5.3 Design of the 2014 Census Test Race Screens

The 2014 CT included three different versions of the race question in a split-ballot experiment. TVersion 1 the same as Version 1 in the 2012 NCT -- the combined question on race and ethnicity had checkboxes and open-text fields on the same screen (Figure 4). . Version 2 included a combined question on race and ethnicity with checkboxes and on the next screen an open-text field for the detailed origins (Figure 5 and Figure 6). Version 3 included separate Hispanic origin and race questions, similar to the design used in the 2010 Census, but in the online questionnaire the two questions were displayed on separate screens where the second question appeared after the participant clicked the “next” button (Figure 7 and Figure 8).

The screenshot shows the '2014 Census Test' interface. At the top, it says 'AN OFFICIAL WEBSITE OF THE UNITED STATES GOVERNMENT' and 'United States Census Bureau'. The main heading is '2014 Census Test'. There are links for 'Instructions', 'FAQs', and 'Logout'. The question is: 'What is Jane Doe's race or origin? Select one or more boxes AND enter the specific race(s) or origin(s). (Help)'. The options are: 'White', 'Hispanic, Latino, or Spanish origin', 'Black or African Am.', 'Asian', 'American Indian or Alaska Native', 'Native Hawaiian or Other Pacific Islander', and 'Some other race or origin'. Each option has a checkbox and an associated open-text field for entering specific details. At the bottom, there are 'Previous' and 'Next' buttons. The footer contains 'Form Number: DA-1 OMB No.: 0607-0970 Approval Expires: 3/31/13' and links for 'Accessibility', 'Privacy', and 'Security'.

Figure 4: 2014 CT Version 1 - Combined question with checkboxes and open-text fields on same screen

AN OFFICIAL WEBSITE OF THE UNITED STATES GOVERNMENT

United States
Census
Bureau

2014 Census Test

[Instructions](#) [FAQs](#) [Logout](#)

What is Jane Doe's race or origin? *Select one or more boxes* [\(Help\)](#)

- White — *For example, German, Irish, English, Italian, Lebanese, Egyptian, and so on.*
- Hispanic, Latino, or Spanish origin — *For example, Mexican or Mexican American, Puerto Rican, Cuban, Dominican, Salvadoran, Colombian, and so on.*
- Black or African Am. — *For example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, and so on.*
- Asian — *For example, Chinese, Filipino, Asian Indian, Vietnamese, Korean, Japanese, and so on.*
- American Indian or Alaska Native — *For example, Navajo Nation, Blackfeet Tribe, Muscogee (Creek) Nation, Mayan, Doyan, Native Village of Barrow Inupiat Traditional Government, and so on.*
- Native Hawaiian or Other Pacific Islander — *For example, Native Hawaiian, Samoan, Guamanian or Chamorro, Tongan, Fijian, Marshallese, and so on.*
- Some other race or origin

[Previous](#) [Next](#)

Form Number: DA-1 OMB No.: 0607-0970 Approval Expires: 3/31/13 [Accessibility](#) [Privacy](#) [Security](#)

Figure 5: 2014 CT Version 2 - Combined question with checkboxes for race

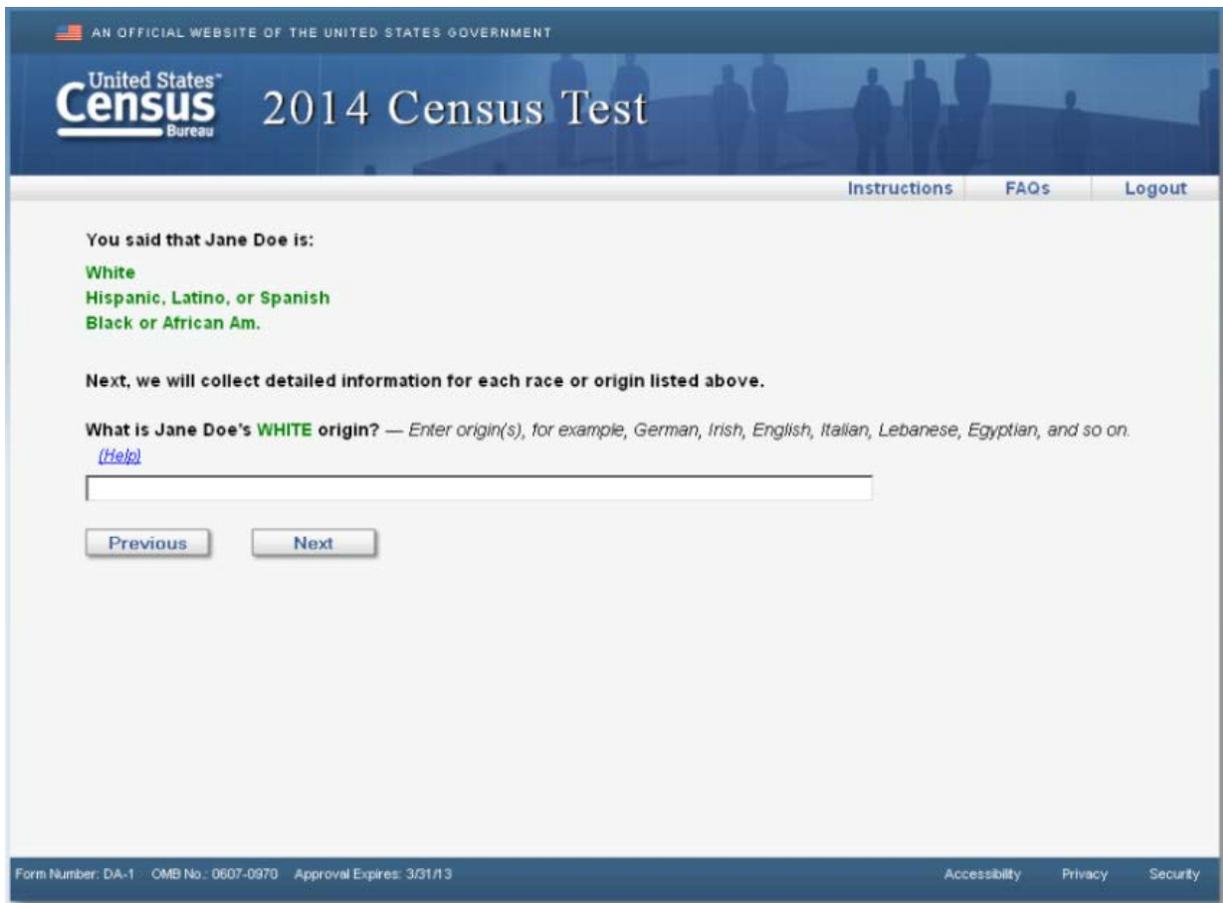


Figure 6: 2014 CT Version 2 PC screenshot for the additional ethnic details for the White category.

AN OFFICIAL WEBSITE OF THE UNITED STATES GOVERNMENT

United States
Census
Bureau

2014 Census Test

[Instructions](#) [FAQs](#) [Logout](#)

Is Jane Doe of Hispanic, Latino or Spanish origin? *Select one or more boxes.* [\(Help\)](#)

- No**, not of Hispanic, Latino, or Spanish origin
- Yes, Mexican, Mexican Am., Chicano
- Yes, Puerto Rican
- Yes, Cuban
- Yes, another Hispanic, Latino, or Spanish origin — *Enter origin(s), for example, Salvadoran, Dominican, Guatemalan, Colombian, Spaniard, Ecuadorian, and so on.*

Form Number: DA-1 OMB No.: 0607-0970 Approval Expires: 3/31/13 [Accessibility](#) [Privacy](#) [Security](#)

Figure 7: 2014 CT Version 3 - Hispanic origin question similar to what was used in 2010 Census

AN OFFICIAL WEBSITE OF THE UNITED STATES GOVERNMENT

United States[™]
Census
Bureau

2014 Census Test

Instructions FAQs Logout

What is Jane Doe's race? *Select one or more boxes. For this survey, Hispanic origins are not races.* [\(Help\)](#)

White — *Enter origin(s), for example, German, Irish, English, Italian, Lebanese, Egyptian, and so on.*

Black or African Am. — *Enter origin(s), for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, and so on.*

American Indian or Alaska Native — *Enter name of enrolled or principal tribe(s), for example, Navajo Nation, Blackfoot Tribe, Muscogee (Creek) Nation, Mayan, Doyan, Native Village of Barrow Inupiat Traditional Government, and so on.*

Chinese

Filipino

Asian Indian

Vietnamese

Korean

Japanese

Other Asian — *Enter origin(s), for example, Pakistani, Cambodian, Hmong, Thai, Laotian, Bangladeshi, and so on.*

Native Hawaiian

Samoan

Guamanian or Chamorro

Other Pacific Islander — *Enter origin(s), for example, Tongan, Fijian, Marshallese, Palauan, Tahitian, Papua New Guinean, and so on.*

Some other race — *Enter race(s) or origin(s).*

Figure 8: 2014 CT Version 3 - Race question similar to what was used in used in 2010 Census

5.4 Design of the 2014 Census Test Race Error Messages

Three experimental error messages for the combined question were also tested in the 2014 CT. One of the error messages appeared when a respondent did not provide any race or detailed origins in at least one open-text field but selected next. The first experimental error message was in green with an icon that had a lowercase letter “i” within a circle (e.g., as a cue to indicate information). It instructed the respondent to answer the origin question but it also included a message that said that the respondent should select the next button if he or she did not identify with an origin (Figure 9). The second error message was identical to the previous version but was red instead of green and with the icon of a lowercase “x” within a red circle (e.g., as a cue to indicate stop and fix the problem) instead of the “i” icon (Figure 10). The third experimental error message included everything that was in the first error message except the message indicated that the respondent should select the next button if they did not identify with an origin (Figure 11).

AN OFFICIAL WEBSITE OF THE UNITED STATES GOVERNMENT

United States[™]
Census
Bureau

2014 Census Test

[Instructions](#) [FAQs](#) [Logout](#)

 Please provide a specific response in the space below. If this person does not identify with a specific origin, use the "Next" button to continue.

You said that jane doe is:

- White
- Hispanic, Latino, or Spanish
- Black or African Am.

Next, we will collect detailed information for each race or origin listed above.

What is jane doe's WHITE origin? — Enter origin(s), for example, German, Irish, English, Italian, Lebanese, Egyptian, and so on.
[\(Help\)](#)

[Previous](#) [Next](#)

Form Number: DA-1 OMB No.: 0607-0970 Approval Expires: 3/31/13 [Accessibility](#) [Privacy](#) [Security](#)

Figure 9: 2014 CT Error 1 - "Green long" error message, with "i" icon



Please provide a specific response in the space(s) below the checked box(es). If this person does not identify with a specific origin, use the "Next" button to continue.

What is Jane Doe's race or origin? Select one or more boxes **AND** enter the specific race(s) or origin(s). [\(Help\)](#)

White — Enter origin(s), for example, German, Irish, English, Italian, Lebanese, Egyptian, and so on.

Hispanic, Latino, or Spanish origin — Enter origin(s), for example, Mexican or Mexican American, Puerto Rican, Cuban, Dominican, Salvadoran, Colombian, and so on.

Black or African Am. — Enter origin(s), for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somalian, and so on.

Asian — Enter origin(s), for example, Chinese, Filipino, Asian Indian, Vietnamese, Korean, Japanese, and so on.

American Indian or Alaska Native — Enter name of enrolled or principal tribe(s), for example, Navajo Nation, Blackfeet Tribe, Muscogee (Creek) Nation, Mayan, Doyan, Native Village of Barrow Inupiat Traditional Government, and so on.

Native Hawaiian or Other Pacific Islander — Enter origin(s), for example, Native Hawaiian, Samoan, Guamanian or Chamorro, Tongan, Fijian, Marshallese, and so on.

Some other race or origin — Enter race(s) or origin(s).

[Previous](#)

[Next](#)

Figure 10: 2014 CT Error 2 - "Red long" error message, with "x" icon

AN OFFICIAL WEBSITE OF THE UNITED STATES GOVERNMENT

United States **Census** Bureau

2014 Census Test

Instructions | FAQs | Logout

i Please provide a specific response in the space(s) below the checked box(es).

What is Jane Doe's race or origin? Select one or more boxes **AND** enter the specific race(s) or origin(s). [\(Help\)](#)

White — Enter origin(s), for example, German, Irish, English, Italian, Lebanese, Egyptian, and so on.

Hispanic, Latino, or Spanish origin — Enter origin(s), for example, Mexican or Mexican American, Puerto Rican, Cuban, Dominican, Salvadoran, Colombian, and so on.

Black or African Am. — Enter origin(s), for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somalian, and so on.

Asian — Enter origin(s), for example, Chinese, Filipino, Asian Indian, Vietnamese, Korean, Japanese, and so on.

American Indian or Alaska Native — Enter name of enrolled or principal tribe(s), for example, Navajo Nation, Blackfeet Tribe, Muscogee (Creek) Nation, Mayan, Doyan, Native Village of Barrow Inupiat Traditional Government, and so on.

Native Hawaiian or Other Pacific Islander — Enter origin(s), for example, Native Hawaiian, Samoan, Guamanian or Chamorro, Tongan, Fijian, Marshallese, and so on.

Some other race or origin — Enter race(s) or origin(s).

Form Number: DA-1 OMB No. 0607-0970 Approval Expires: 3/31/13 Accessibility Privacy Security

Figure 11: 2014 CT Error 3 - "Green short" error message, with "i" icon

5.5 Usability Findings of the 2014 Census Test Race Screens

In this round of testing, five participants saw Version 1, (Figure 4). Three participants saw Version 2 (Figure 5). Three participants saw Version 3 (Figure 7 and Figure 8.) Participants were also exposed to the different versions of the edit messages.

5.5.1 African American as race and as detailed origin

One participant triggered both the red short and the red long error message while completing the online questionnaire. This participant received the race screen similar to version used in the 2010 Census (Version 3). He initially did not select a race at all. After clicking on "next" he received the short red message. He then selected one category (Black or African Am.) and did not enter a detailed origin. He then received the red long error message and he typed in "African American" after looking at the categories. See Figure 12. During the debriefing, this participant used the word "confusion" to describe how he was feeling after seeing the different error messages. The behavior of leaving the open-text field empty after checking the box "Black or African Am." and then, after receiving an error message typing in "African American" was also seen in the usability tests of the 2012 NCT as well as future usability studies.

5.5.2 Separating Hispanic origin and race is confusing

One participant (not of Hispanic origin) who received the separate Hispanic origin and race questions in Version 3 (Figure 7) said she was offended when asked for the Hispanic origin because she said that Census had no interest in her ethnic background, only had interest in Hispanics.

One participant was asked by the TA on Version 3 (Figure 8) how she would answer the race screen. The participant responded, “This is so difficult, I am White and Black. How can I explain? I have to write White Italian, and Black African American. But it doesn’t include Hispanic, Yes? [I’m also] Dominican.” She wanted to be able to indicate White, Black and Hispanic all on one screen but the separate question did not feel adequate for what she wanted.

It appeared that Version 3 (with Hispanic origin first and on a separate screen from the race question) was problematic for some users because they could not see ahead to know that they would be asked for race after the Hispanic origin question. This version appears to give priority to one origin over the other by the order of the question. The combined question on race and ethnicity was less problematic in that regard.

5.5.3 Combined race and ethnicity question worked for participants

During usability testing we noticed that participants did not encounter problems with the combined race and ethnicity question. That is, participants did not have any issues with “Hispanic, Latino or Spanish origins” as a response option, and made no spontaneous comments about the choice. This was not the case for when the question was separated into two questions (see finding 4.2.5).

5.5.4 Feedback on instructions for selecting more than one race

One participant did not realize he could select more than one race on the version that separated Hispanic origin and race (Version 3-Figure 8). That participant wrote in the ‘Some other race’ response field “half white and African American.” When asked during debriefing, not everyone understood that they could enter more than one race, but it was not clear whether that differed by the version of the race screen they were provided.

5.5.5 Less initial item nonresponse when detailed origin is asked on its own screen, than when it is collected on the main race screen

A few participants did not fill in their detailed origins into the provided open-text field when that field was on the main race screen as it was in Version 1 (Figure 4) and Version 3 (Figure 8). In contrast, when the detailed origins question was on a separate screen as it was in Version 2, that question was always completed by participants (Figure 6). It is likely that the screen was completed because it appeared as its own question on a separate screen.

One participant, who inadvertently answered the Census survey twice⁷, accessed two different versions of the race question (Version 1 and Version 2). She provided different answers to the race

⁷ She had originally been given a Census ID number and started her census that way. Midway through the session, participants are asked to imagine what they would do if they had to leave the survey. This participant re-entered the

question in each of her Census submissions. For Version 1 (Figure 4) the participant only checked the box “Hispanic origin” and left the open-text field blank. For Version 2 she checked the box “Hispanic origin” on Figure 5 and then on the next screen that asked for detailed origins (Figure 6) she entered the specific Hispanic origin. In this case, it appeared that Version 2 collected more detailed race information than Version 1.

5.5.6 Error messages could confuse users who are not expecting or not familiar with error messages

Few error messages in the combined race and ethnicity question were invoked during the actual sessions. Only three out of eleven participants received them. One participant received the long red error message on the combined race and ethnicity question with checkboxes and open-text field on the same screen (Version 1) because she did not enter a detailed origin (see Figure 10). She mistook the message to mean she could not select more than one race. She had originally selected “White” and “Hispanic,” but did not type in a detail for the White race. After she received the error message, she navigated back and unselected the “White” response option. This participant had identified herself as White, Black, and Hispanic on the paper demographic questionnaire completed at the beginning of the session, but in the 2014 CT we only collected that she was “Hispanic.” During the debriefing, she commented on the error message that she had seen: “In the first one I was a little nervous, the error that came up. The red box.... that made me feel insecure so the only thing I could do was check Hispanic.”

In 2014, the open-text field did not highlight which might have led to her misunderstanding of what the error message referenced. However, this same design was used in 2012 and we did not observe anyone changing his or her race because of the message.

5.5.7 Error messages encourage participants to enter a specific detailed origin, even when they do not have strong ties to that detailed origin, or do not know the detailed origin

One participant interacting with Version 3 (Figure 8) of the race screen selected a race (White) but then did not specify an origin. He moved forward by clicking on the “next” button and then received the green long error message. At this point, he said that he did not really know what to put because in his family history there were so many different origins. In the end, he typed in “English.”

5.5.7.1 *Eye tracking on race error messages suggests participants did not read the messages*

Two participants out of the three that triggered an error message had eye-tracking (ET) data. The ET data highlights that they did not actually read the entire message – especially when it was the longer message. As shown in the right image of Figure 12, only the first few words of the long message were attended to. In the image on left in Figure 12, the participant appeared to skim the

survey but inadvertently navigated down the non-ID path of the questionnaire. She completed all the screens in the non-ID version, including the screens she had previously completed in the ID path. This created two returns for her, the original partially completed return with the ID and the fully completed non-ID return.

entire short message. In the image in Figure 13, we observe the same pattern of eye fixations only on the first part of the error message.

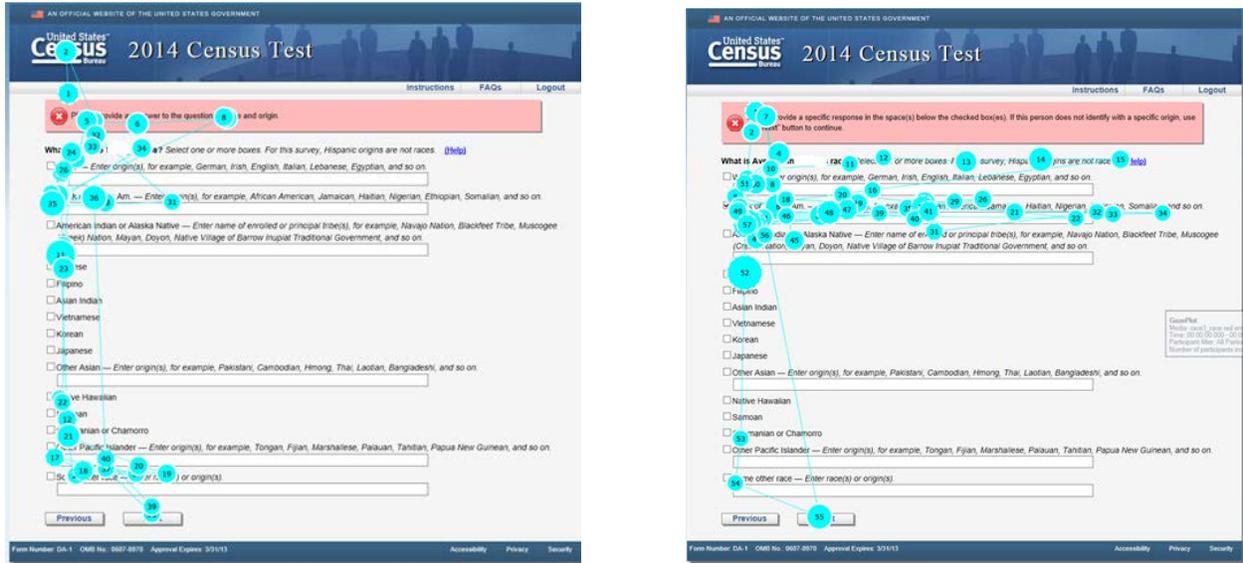


Figure 12: Gaze plot of one participant who received the default short red error message on the race question because he did not choose a race. Then after choosing a race he received the long red error message because he left detailed origins blank.

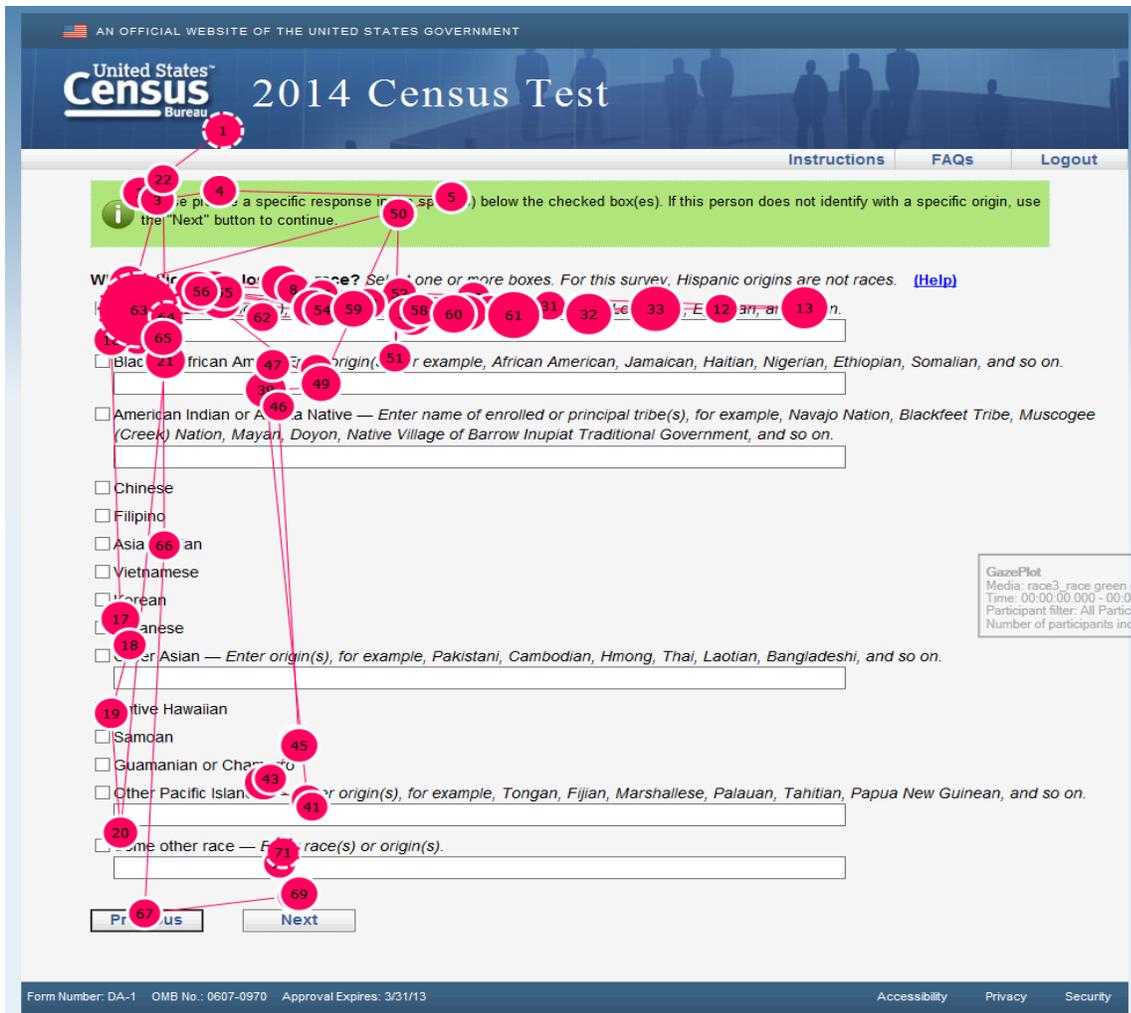


Figure 13: Gaze plot of one participant who received the green long error message on Version 3 of the Race question

5.5.8 Debriefing feedback on error messages

During the debriefing, the TA asked participants about the error messages – and their responses indicate some confusion with the messages. Responses were mixed about what the message was telling them to do. While many participants said that the two long messages meant they could ignore them if they want (particularly if it was green), the instruction may not be clear enough to tell the participant what they have to do. One participant said it needed to clarify that the first sentence indicates something needs to be typed into the box. She said she thought she had already provided a specific response. During debriefing, participants were also asked about the color of the error messages. One participant said that the color did not mean anything to him, a few other participants reported that green meant that it could be ignored. Most participants commented that the red icon of the “x” was off-putting.

Participant comments on the red color of the error message included:

- “Use the red color but change the icon to the exclamation point.”

- “The red long message was, kind of a conflicting image. Red tells you, you skipped something, but the message says it’s okay to go on. [I’d] most likely fill something in on this one.”
- “It’s in red. [It] alerted me there is an issue. I’ve made an error. I clicked too many, perhaps there is something I am missing. Red means: Warning. Re-do. It means mistake.”
- “I feel this is mandatory. It probably won’t let go any further. If you did you might be thrown out.”

Participant comments on the green color of the error message included:

- “The green I could tell meant, ‘read me’ and I also know that it’s positive and I haven’t done an error but I knew I had to read that. My sense is that I cannot proceed without filling out the page, I have to fill it out or the process is stopped.”

Participant comments on the long green error message:

- “Hey we’d appreciate it but not end of the world [if it was ignored].”

Participant comments on the short green error message:

- “I liked it - was simpler with less to read, but I would feel more compelled to put in my origin.”
- “Makes it easier to bypass the specifics. [I] prefer for it to be more direct. People might choose not to get more specific.”
- “A mixture because it says to provide a response but the green makes it seem less angry. I don’t think you could skip.”
- “This is my favorite one. I like that it’s larger. It looks easier. Just ahh it’s less wordy. It’s visually more pleasing.”

One participant, when asked to compare the two colors said, “Red might make you think you did something wrong. I like green better.” While all participants had a comment on color when directly asked during debriefing, the few who triggered a race error message while completing their online census questionnaire did not appear to give meaning to the color during the session as they had no spontaneous comments about the color.

5.6 Design of the 2015 Census Test – Optimizing the Race Screen for Mobile Devices

The 2015 CT was primarily focused on getting the screens working and optimized for the mobile display. There was only one version of the race question tested, which was similar to Version 2, from the 2014 CT. It was the combined question (shown for a PC in Figure 14) and the detailed origins were collected on subsequent screen(s), after the participant had clicked the “next” button (Figure 15). The race or races that the participant checked on the first screen determined what subsequent screen(s) were shown. In Figure 14, a person would have to select the boxes in front of “Native Hawaiian or Other Pacific Islander” and “Some other race or origin” response to get the screen shown in Figure 15. In this test and in all future tests, the screens would change dynamically to display on a mobile phone, which is referred to as “mobile optimized.” See Figure 16 for a mobile view.

AN OFFICIAL WEBSITE OF THE UNITED STATES GOVERNMENT

United States[™]
Census
Bureau

2015 Census Test

[Instructions](#) [FAQs](#) [Save and Log Out](#)

What is Jane A Doe's race or origin? [Select one or more boxes](#) [\(Help\)](#)

- White — For example, German, Irish, English, Italian, Lebanese, Egyptian, and so on.
- Hispanic, Latino, or Spanish origin — For example, Mexican or Mexican American, Puerto Rican, Cuban, Dominican, Salvadoran, Colombian, and so on.
- Black or African Am. — For example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somalian, and so on.
- Asian — For example, Chinese, Filipino, Asian Indian, Vietnamese, Korean, Japanese, and so on.
- American Indian or Alaska Native — For example, Navajo Nation, Blackfeet Tribe, Muscogee (Creek) Nation, Mayan, Doyon, Native Village of Barrow Inupiat Traditional Government, and so on.
- Native Hawaiian or Other Pacific Islander — For example, Native Hawaiian, Samoan, Guamanian or Chamorro, Tongan, Fijian, Marshallese, and so on.
- Some other race or origin

Form Number: DA-1 OMB No.: 0607-0979 Approval Expires: 5/31/2015 [Accessibility](#) [Privacy](#) [Security](#)

Figure 14. Screenshot from the 2015 CT

AN OFFICIAL WEBSITE OF THE UNITED STATES GOVERNMENT

United States
Census
Bureau

2015 Census Test

[Instructions](#) [FAQs](#) [Save and Log Out](#)

You said that Jane A Doe is:

- Native Hawaiian or Other Pacific Islander
- Some other race or origin

Next, we will collect detailed information for each race or origin listed above.

What is Jane A Doe's **NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER** origin?

— Enter origin(s), for example, Native Hawaiian, Samoan, Guamanian or Chamorro, Tongan, Fijian, Marshallese, and so on. ([Help](#))

[Previous](#) [Next](#)

Form Number: DA-1 OMB No.: 0607-0079 Approval Expires: 5/31/2015 [Accessibility](#) [Privacy](#) [Security](#)

Figure 15. Screenshot from the 2015 CT that asks for additional ethnic details

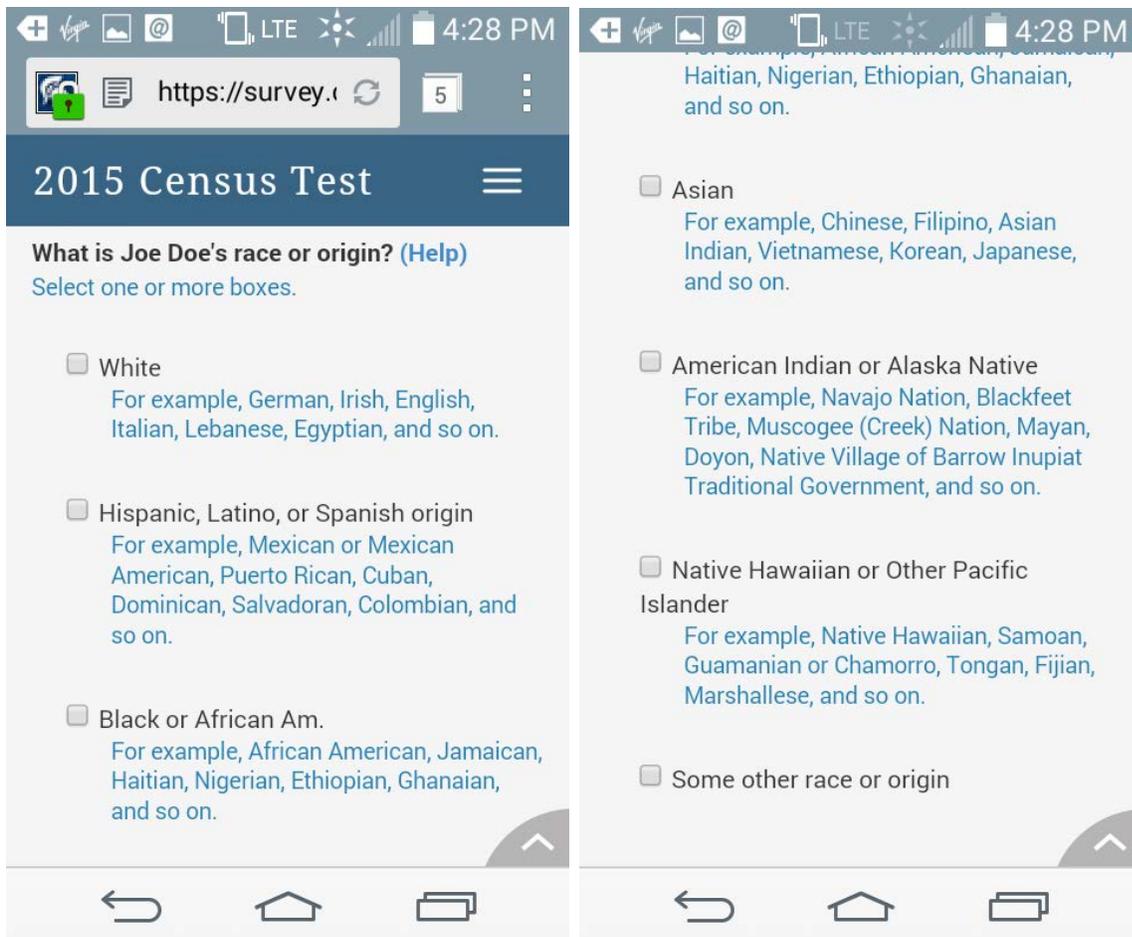


Figure 16. Screenshot from the 2015 CT.

5.7 Usability Findings of the 2015 Census Test –Optimizing the Race Screen for Mobile Devices R1 and R2

The race screens changed from Round 1 to Round 2 only insofar as there was additional spacing added between the response options. This was a global change across all screens to address the issue of “fat finger.” “Fat finger” is the phenomenon where the user touches an inaccurate response because there was not enough space for the finger to select the correct answer. In addition, italic font for instruction text was changed back to regular font and left aligned. Italic font on smaller devices is more difficult to read than regular font, and generally left aligning text is easier to read than centered text. The question wording and response option choices did not change between Round 1 and Round 2. See Figure 17.

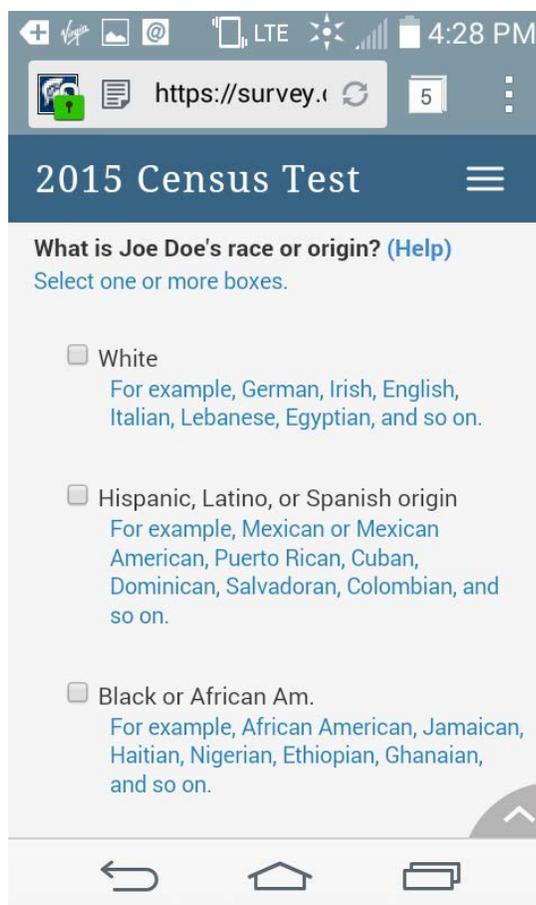
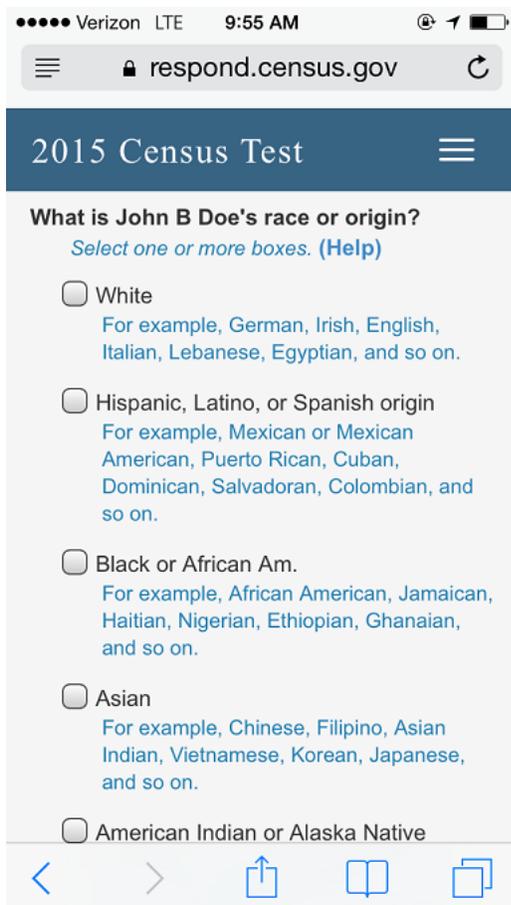


Figure 17. 2015 CT Round 1 and Round 2. On left Round 1, on right Round 2. Round 2 has more space between response options and no italic font in the instruction “Select one or more boxes.”

5.7.1 Combined race and ethnicity question worked for participants

Similar to the 2014 CT usability testing, participants did not encounter problems with the combined race and ethnicity question in the 2015 CT usability testing. That is, participants did not have any issues with “Hispanic, Latino or Spanish origins” as a response option, and made no spontaneous comments about the choice.

5.7.2 Some participants understood “origin” to mean “National origin” and not “Hispanic origin”

The race question read, “What is your race or origin?” While no one expressed confusion with the word *origin* in the race question, participants indicated they were thinking the word meant “national origin” rather than “Hispanic origin.” At least one participant said there was no way to select the origin (as she attempted to click on the blue examples under the race response choices). One participant said she thought race and origin were actually two different questions, but there was no way to select her origin on the screen, only her race. The next question following the

combined race and ethnicity asked “what is your <RACE⁸> origin?” One participant spontaneously said as she read the screen, “What is my white origin? I guess that means was I born in the U.S. My ‘white origin’ is a little confusing.” A modification to the initial question stem could simply be to remove the word ‘origin’ e.g., “What is your race?” The usability team’s recommendation was to add additional context to “origin”, such as ancestry, ethnic details, etc. in the detailed origins question; the recommendation was not accepted.

5.7.3 Issue with the abbreviation of African Am

When confronted with the word label “African Am.” to indicate African American, one participant said aloud “A.M.” This participant also read aloud the example which had the word American spelled out in African American. We know from previous 508 compliance testing that “Am.” reads incorrectly in the JAWS screen reader. Usability staff were informed that the use of the abbreviation is from earlier paper questionnaires when there was not enough space to spell out the word American. Eventually, in the 2017 CT, following usability recommendations, the “Am.” was changed to “American.”

5.7.4 Feedback on instructions for selecting more than one race

When probed during debriefing, many participants said they thought they could only select one race. None of these participants self-identified as multi-racial, even though at least two participants had parents of different races. The recommendation coming out of Round 1 was to left-justify the instruction directly below the question because participants seemed to understand what the question was asking for, and based on prior eye-tracking data, they were not necessarily reading the entire question and getting to the instruction. This design change was made prior to Round 2 usability testing. However, during Round 2 it was still not clear whether participants noticed or processed the ‘*Select one or more races*’ instruction. It was not clear, based on the limited feedback we received, whether the change to left justify the instruction worked any better than the initial design. Only one participant selected multiple races for her child in Round 2.

Another participant, when answering for a person in her household, with whom she was unrelated, selected “Some other race.” She said she knew the person was multi-racial, but was not sure which races to select. Still, no one who was multi-racial commented that they had to select only one choice.

5.7.5 Race examples caused user comments

Two participants during the first round of testing asked how the Census Bureau came up with the examples for each race. The first user commented spontaneously while completing his online census questionnaire for his household. This participant identified as Ugandan and asked why Nigerians were there as examples. Another participant commented during the debriefing that her origin was not listed. She asked, “Who came up with these examples?” She said that some people would be offended if their origin was not listed. Based on this feedback the team recommended

⁸ Here <RACE> would fill with the race or ethnicity that the participant had selected on the previous screen.

that the help text include additional examples as well as explain why certain ethnicities were listed in the order they appeared. (During Round 1 debriefing we did not ask explicit probes about the help text.)

For Round 2, the race help text was modified to include an explanation for the order of the examples, (e.g., “examples listed for each response category are listed in order of population size...” as shown in Figure 18). Still even with this update to the help text, we continued to hear mixed comments on the race examples. Some said the inclusion of the examples implied an overly complicated question and that it was not a “typical” design. However, other users said they liked to see the examples listed out. When asked about the help text during the Round 2 debriefing, most users appeared to understand it, although one participant said the language was complex. While usability staff suggested conducting research to see if removing the examples for the race response categories would change the response distribution, this suggestion was not implemented.

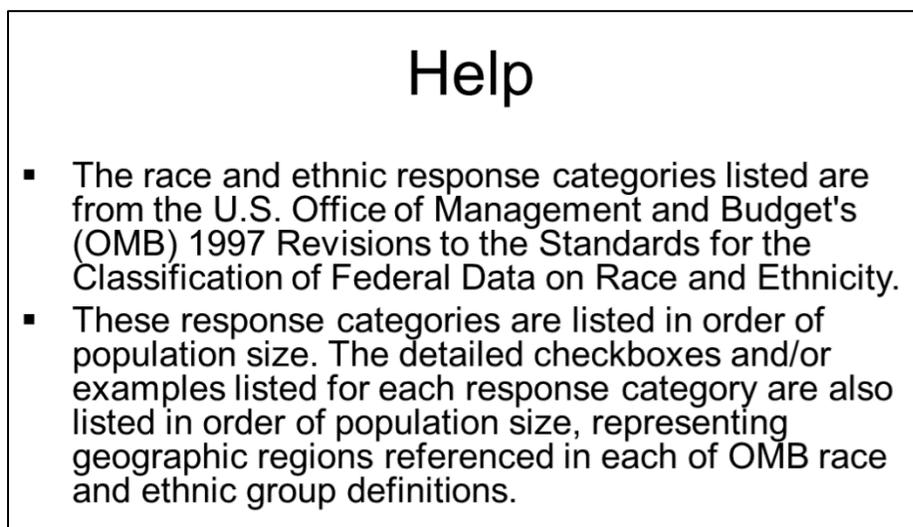


Figure 18. Race help shown to participants during Round 2 debriefing.

5.7.6 Wide variability in reporting detailed origins

One participant wondered aloud how far back in his ancestry he should go to answer the race question. He decided to go only two generations for his detailed origins. Another participant who identified himself and his wife as being both African American and Native American marked only African American for his biological children. We recommended addressing the uncertainty on what the census means by detailed origins within the help section of the screen but that recommendation as not accepted.

5.7.7 Confusion with when to use “American”

One foreign-born participant, as he was filling out the race question, asked what to put for his American born son. He said his son was “an American kid.” He eventually typed Ugandan for his son. Another user put *American* for everyone in her household including herself because she was not sure of their origins and she called herself a “Euro-mix.” She said she felt like *American* would be considered an accurate response because it came up as an option in the predictive text.

We recommended addressing the uncertainty on what the census means by detailed origins within the help section of the screen but that recommendation was not accepted. We continued to see this phenomena of answering “American” for foreign born participants and others who did not associate with any particular detailed origin in subsequent census usability tests.

5.7.8 Error messages encourage participants to enter a specific detailed origin, even when they do not have strong ties to that detailed origin, or do not know the detailed origin

At least one person who left the detailed origins blank for an unrelated household member said he did not know what the roommate’s detailed origin was and navigated forward. This action triggered an error message, where he said “okay so ‘don’t know’ is not an option.” (See screen with error messages in Figure 19.) He then made a guess of what the roommate’s detailed origin was and navigated forward.

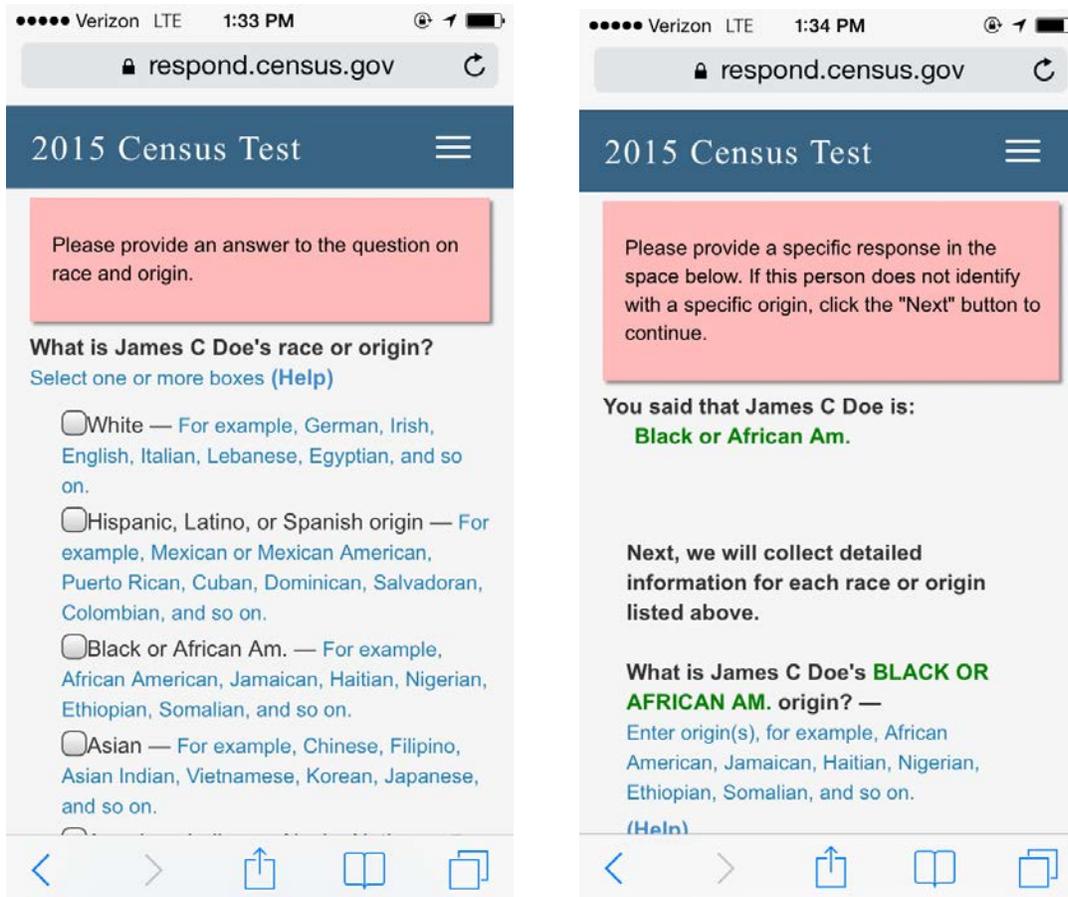


Figure 19. 2015 CT. Error message appears on screen if user attempts to navigate forward without selecting a response.

5.7.9 The use of blue font to highlight certain words or parts of the question caused user problems

The example texts were in blue font on the combined question on race and ethnicity and the detailed origins screen. During the session while answering the combined question on race and

ethnicity, one user tried to click on the blue example text “African American.” Another user tried to click on “Irish” on the detailed origins screen saying she wanted it to fill the field. Another user tried to click on the blue text to select the detailed origin. During the debriefing, one participant said that blue usually means a link, but she knew the examples were not links. Based on this feedback, and the convention that blue is used to indicate hyperlinks, the team recommended that the example text return to a black color as they were in earlier tests. This recommendation was accepted and by the 2016 Census Test, the screens no longer used blue text except when the text was a hyperlink.

5.7.10 Gathering detailed origins on multiple screens caused additional respondent burden

Race details were collected on separate screens, so if a participant selected two races, then that participant would receive two subsequent detailed origins screens, one for each of the races.

While selecting multiple races was rare during the usability sessions, during the second round of testing, one participant selected two races for her daughter. She then navigated forward to the first detailed origins screen where she entered detailed origins for both races. When she navigated forward to the following screen, she realized that she should have separated the detailed origins. She went back to the previous screen, deleted the second detailed origin, then navigated forward to the next screen, and typed that detailed origin again. Thus, she was able to fix the problem. However, it was not initially clear to her that there would be multiple screens asking for all the detailed origins that she had checked, and this caused additional burden for her. During debriefing, the same participant recommended that we allow all the detailed origins for the races she had selected to be listed on the same screen. We shared this recommendation with the development team, and while it seemed likely to solve the issue, it was not adopted. Paradata analysis could show how often respondents that had selected multiple races went back and forth on the navigation path to put the detailed origins on the correct screens and thus give an indication of the magnitude of the issue.

5.8 Design of the 2015 National Content Test Race Screens

The 2015 NCT tested the combined race and ethnicity question as well as the 2010 Census version of separate race and Hispanic origin questions. This census test also added the MENA category. In the combined question on race and ethnicity, there were several variations of the question wording, instructional text, and categories in a split-ballot experiment including:

- A panel used the words “race or ethnicity” in the question, “What is <NAME’s> race or ethnicity?”
- A panel used the words “categories” in the question, “Which categories describe <NAME?>”
- A simple instruction, “Select one or more boxes AND enter ethnicities.”
- A longer instruction, “Select one or more boxes AND enter ethnicities. Note you may report more than one group.”

See examples of the screens with different variations in PC format in Figure 20 through Figure 22, and mobile versions Figure 25 and Figure 26 below.

The screen following the combined question on race and ethnicity collected detailed origins and was designed in two different formats:

- A list of examples for the previously selected race or ethnicity group and an open-text field.
- A list of checkboxes for the six most frequently occurring detailed origins in that race or ethnicity group in the U.S. based on American Community Survey data. There was also an open-text field if none of the choices applied to the respondent.

See examples of the screens with different variations on the PC in Figure 23 and Figure 24 and for the mobile version in Figure 27. For more on the goals of the 2015 NCT and the outcomes of the testing of the race question, as well as all variations of the race and ethnicity screens tested see (Mathews et al., 2017).

The screenshot displays the 2015 National Content Test interface. At the top, it features the United States Census Bureau logo and the title "2015 National Content Test". Navigation links for "Instructions", "FAQs", and "Save and Log Out" are visible. The main question is "What is John K Doe's race or ethnicity? (Help)", with instructions to "Select one or more boxes AND enter ethnicities." Below this, there are seven checkbox options, each with a list of example ethnicities: "White" (German, Irish, English, Italian, Polish, French, etc.), "Hispanic, Latino, or Spanish origin" (Mexican, Puerto Rican, Cuban, etc.), "Black or African Am." (African American, Jamaican, Haitian, etc.), "Asian" (Chinese, Filipino, Asian Indian, etc.), "Middle Eastern or North African" (Lebanese, Iranian, Egyptian, etc.), "American Indian or Alaska Native" (Navajo Nation, Blackfeet Tribe, etc.), and "Native Hawaiian or Other Pacific Islander" (Native Hawaiian, Samoan, etc.). A "Some other race or origin" option is also present. At the bottom of the form are "Previous" and "Next" buttons. The footer contains OMB No.: 0607-0985, Approval Expires: 6/30/2018, and links for Accessibility, Privacy, and Security.

Figure 20. 2015 NCT PC screenshot with one variation of question/instruction text and including the additional MENA category

AN OFFICIAL WEBSITE OF THE UNITED STATES GOVERNMENT

United States[™]
Census Bureau

2015 National Content Test

Instructions FAQs Save and Log Out

Which categories describe John K Doe? (Help)
 Select one or more boxes **AND** enter details.
 Note, you may report more than one group.

- White
For example, German, Irish, English, Italian, Polish, French, and so on.
- Hispanic, Latino, or Spanish origin
For example, Mexican or Mexican American, Puerto Rican, Cuban, Dominican, Salvadoran, Colombian, and so on.
- Black or African Am.
For example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Ghanaian, and so on.
- Asian
For example, Chinese, Filipino, Asian Indian, Vietnamese, Korean, Japanese, and so on.
- Middle Eastern or North African
Enter, for example, Lebanese, Iranian, Egyptian, Syrian, Moroccan, Algerian, etc.
- American Indian or Alaska Native
For example, Navajo Nation, Blackfeet Tribe, Muscogee (Creek) Nation, Mayan, Doyon, Native Village of Barrow Inupiat Traditional Government, and so on.
- Native Hawaiian or Other Pacific Islander
For example, Native Hawaiian, Samoan, Guamanian or Chamorro, Tongan, Fijian, Marshallese, and so on.
- Some other race or origin

OMB No.: 0607-0985 Approval Expires: 6/30/2018 Accessibility Privacy Security

Figure 21. 2015 NCT PC screenshot with another variation of question/instruction text and including the additional MENA category

AN OFFICIAL WEBSITE OF THE UNITED STATES GOVERNMENT

United States[™]
Census
Bureau

2015 National Content Test

[Instructions](#) [FAQs](#) [Save and Log Out](#)

Which categories describe John K Doe? (Help)
Select one or more boxes **AND** enter details.

- White
For example, German, Irish, English, Italian, Polish, French, and so on.
- Hispanic, Latino, or Spanish origin
For example, Mexican or Mexican American, Puerto Rican, Cuban, Dominican, Salvadoran, Colombian, and so on.
- Black or African Am.
For example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Ghanaian, and so on.
- Asian
For example, Chinese, Filipino, Asian Indian, Vietnamese, Korean, Japanese, and so on.
- Middle Eastern or North African
Enter, for example, Lebanese, Iranian, Egyptian, Syrian, Moroccan, Algerian, etc.
- American Indian or Alaska Native
For example, Navajo Nation, Blackfeet Tribe, Muscogee (Creek) Nation, Mayan, Doyon, Native Village of Barrow Inupiat Traditional Government, and so on.
- Native Hawaiian or Other Pacific Islander
For example, Native Hawaiian, Samoan, Guamanian or Chamorro, Tongan, Fijian, Marshallese, and so on.
- Some other race or origin

OMB No.: 0607-0985 Approval Expires: 6/30/2018 [Accessibility](#) [Privacy](#) [Security](#)

Figure 22. 2015 NCT PC screenshot, similar to prior image but with a shorter instruction.

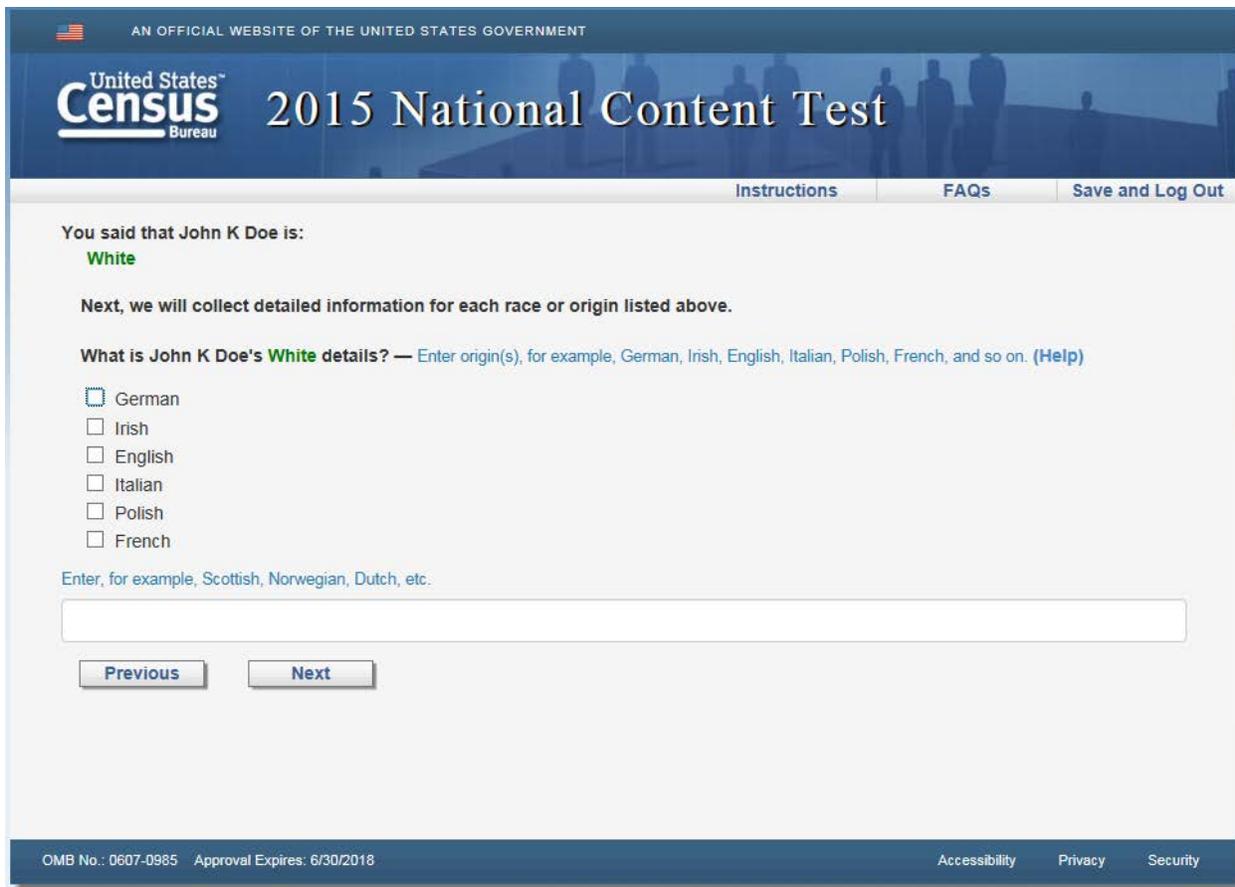


Figure 23. 2015 NCT PC screenshot of how respondents could enter detailed origins. Here there is a list of six detailed origins with a check box in front and a last option with an open-text field.

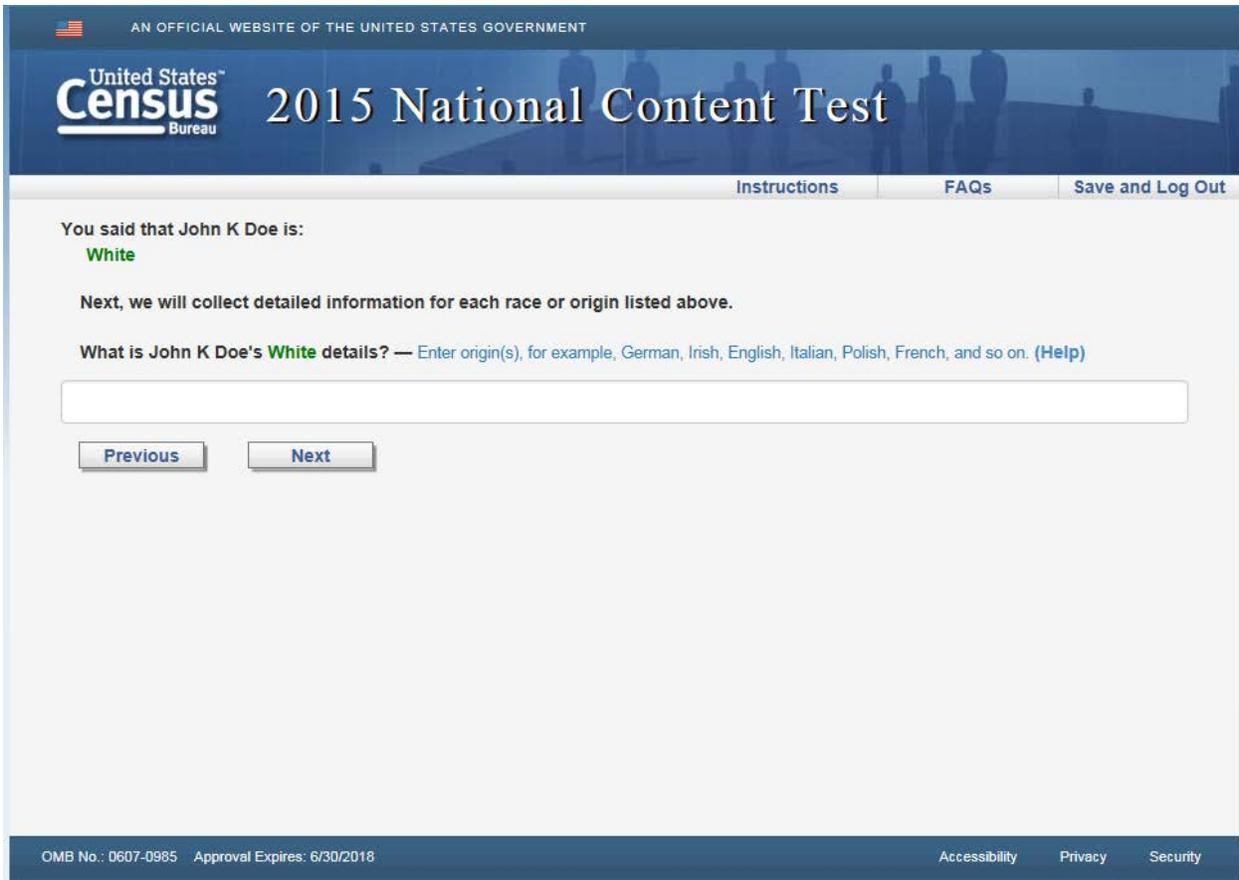


Figure 24. 2015 NCT PC screenshot of how respondents could enter detailed origins. Here it has examples listed out with an open-text field.

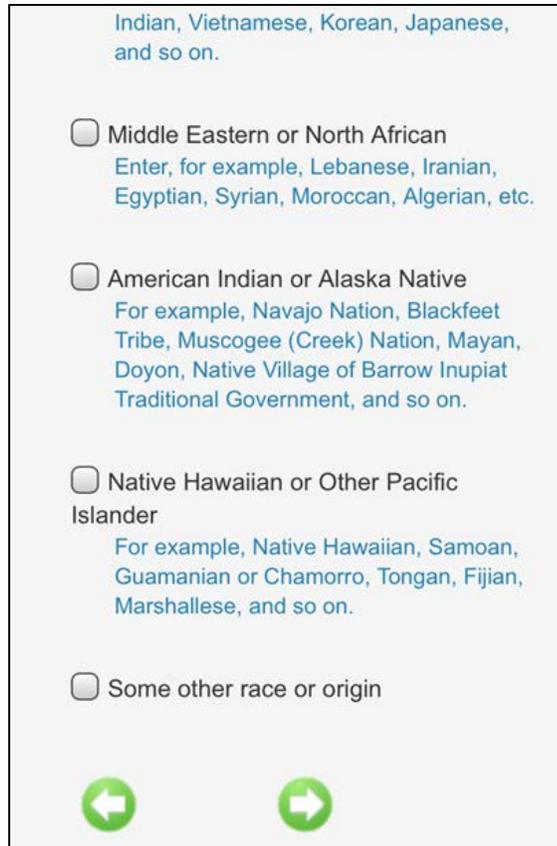
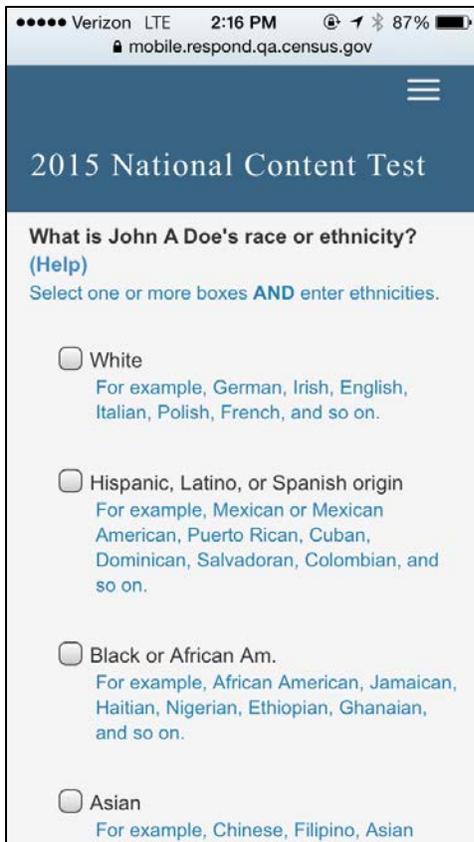


Figure 25. 2015 NCT mobile screenshot with one variation of question/instruction text and including the additional MENA category

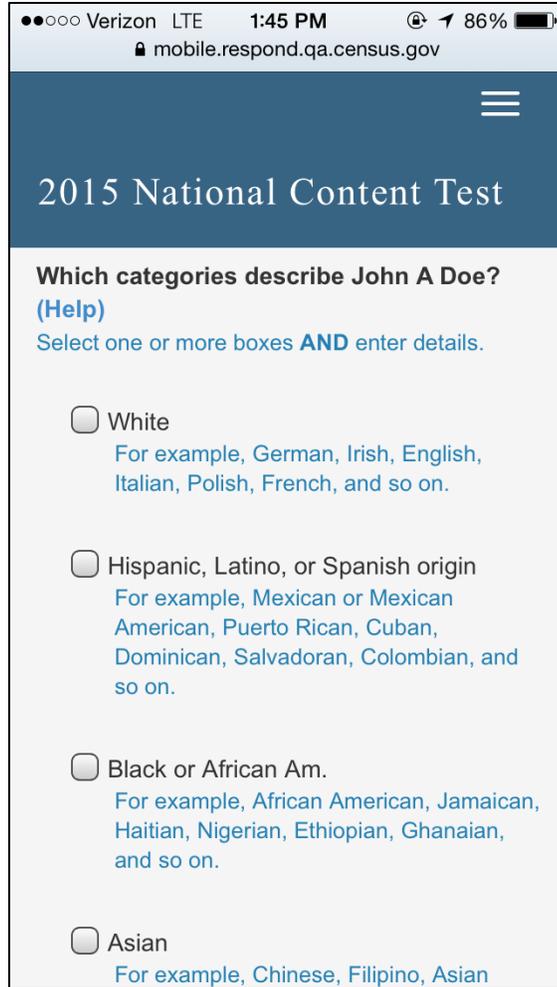
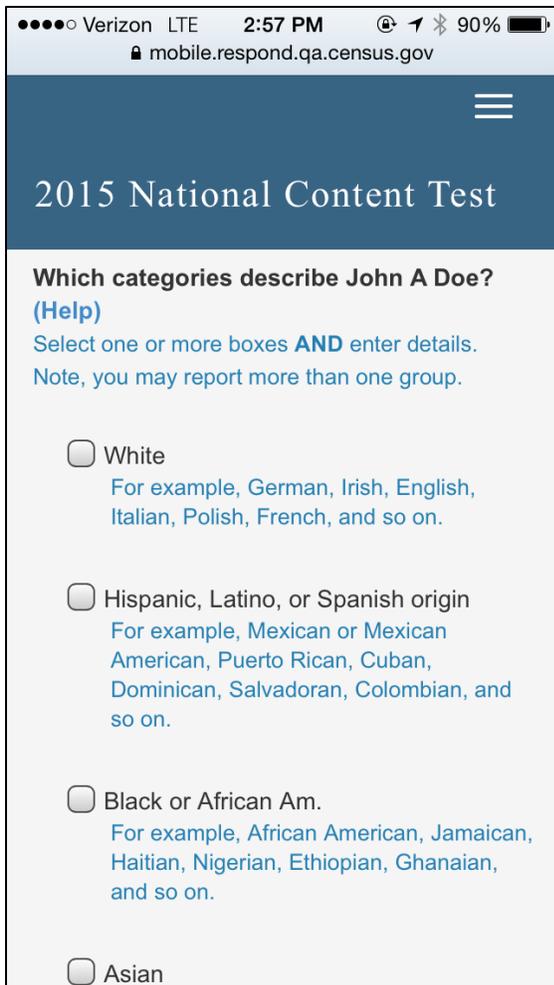


Figure 26. Two different 2015 NCT mobile screenshots, on left the screen has two instructions, on the right it has only one instruction.

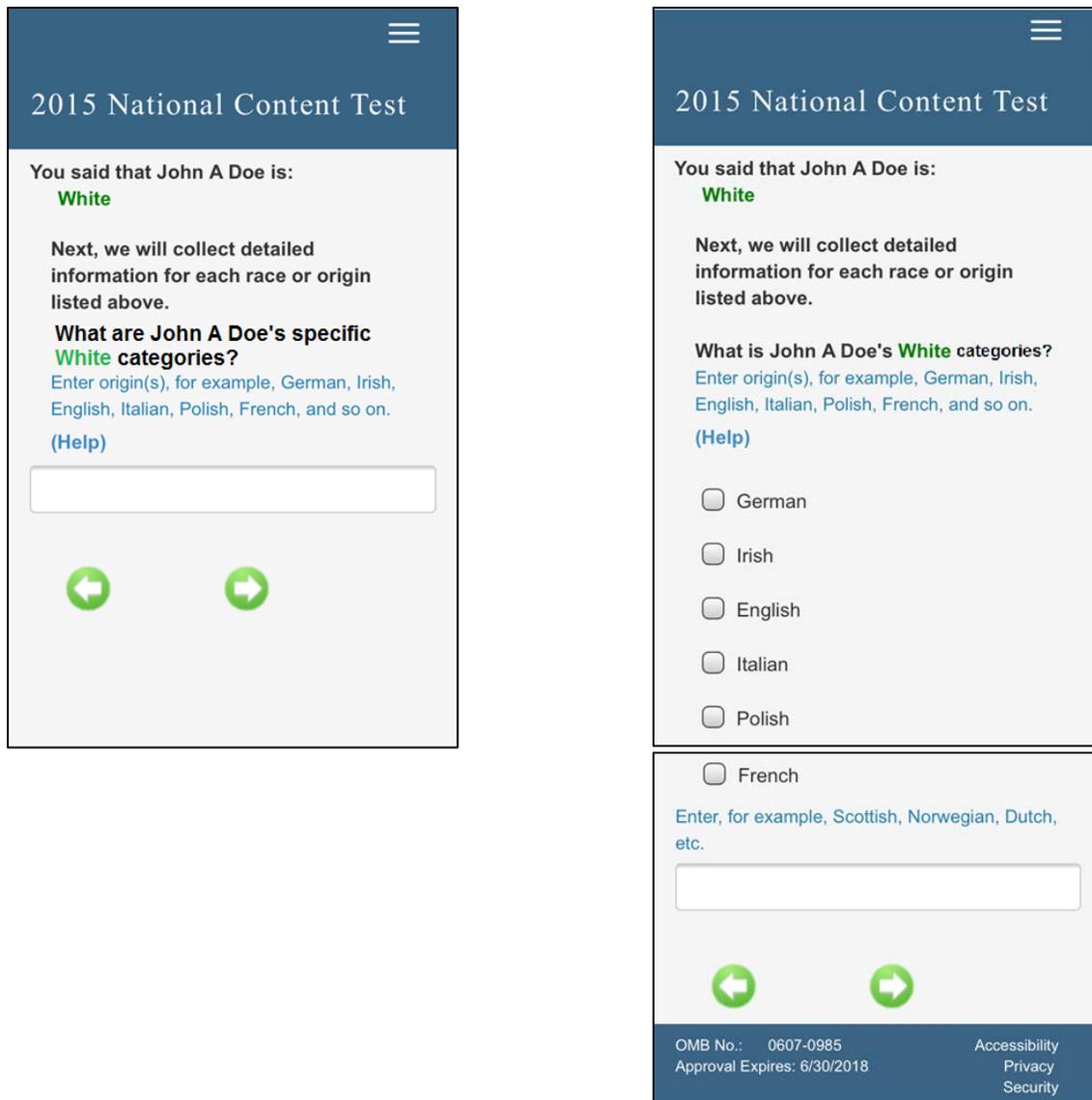


Figure 27. 2015 NCT mobile screenshot variations of how respondents could enter their detailed origins. On left, it is with examples listed out and an open-text field, on right it is lists top six detailed origins by population count with checkboxes and then an open-text field.

5.9 Usability Findings of the 2015 National Census Test Race Screens

In this test, there were a number of different designs of the combined race and ethnicity questions and detailed origins screens. During usability testing we did not test all the different versions but focused on new versions that had not yet been usability tested. We also actively recruited Middle Eastern and North African, and Puerto Rican participants. These two user groups gave us additional feedback on the new designs that had not yet been tested for usability.

5.9.1 Feedback on the MENA category

We did not hear any spontaneous participant comments about the additional MENA category unless the person answering (or someone in their household) was from the Middle East or North Africa or was African American. One participant who identified as African American verbalized that she could mark North African but she did not; she marked the Black or African American category. Another African American asked during the debriefing part of the session why we were separating North African from the rest of the continent, “Africa is Africa.”

One participant of Lebanese descent expressed positive sentiment that there was a MENA category. Saying, “Here we go, Middle Eastern, they finally labeled that, they put that onto something.” During debriefing she said, “It makes me feel good to see it. Makes me feel like I belong.” However, another participant of Lebanese descent who was surprised to see the MENA response option said she would be fine with including the category or not including it. She did not have a preference.

Another participant, of Iranian descent, said he was happy to see MENA, as he did not have to go through the “typical dilemma of if I should put White or something else.”

One participant who was of Amazigh (Berber) descent said, “Yeah I am White” but did not select that option. After scrolling down and choosing MENA, the participant said, “it’s White normally but they changed the categories.”

One participant of Turkish descent said, “I’m not White, I’m not Asian, I’m not Middle Eastern. Something else, some other race. I am a Turk.” The participant explained that Turkey is located between Asia and Europe, bordered with the Middle East. He also said that he was from the middle of Turkey, so he is “Some other race, just Turkish.”

One participant of Somali descent said, “Technically we are considered Middle Eastern, but I don’t know what to put. I could put either Black or Middle Eastern, but it really doesn’t matter. The U.S. sees me as Black. If I was walking down the street, other people would think I’m Black.” He marked his race as “Black or African American.”

See Table 5 for detailed information on what the MENA participant was initially screened as (e.g., by the recruiter over the phone), what category the participant selected when answering the online questionnaire and what the participant entered into the open-text field for his/her detailed origins.

Table 5. Participant Screened as a Potential MENA by What They Selected for Race and Detailed Race

Screened	Race/Category Chosen	Detailed Origins
Amazigh	MENA	Algerian
Amazigh	MENA	Algerian
Armenian	White	Armenian (wrote in)
Armenian	White	Armenian (wrote in)
Chaldean	MENA	Iraqi
Chaldean	MENA	Left blank (refused)
Iranian	MENA	Iranian (wrote in)
Israeli	White	Left blank

Israeli	White/Some other race	Left blank then typed Jewish/ Left blank was confused
Israeli	White/MENA	English (meaning as American)/Israeli
Israeli	White	American
Kurdish	MENA	Iraqi-Kurdish (wrote in)
Lebanese	MENA	Checked Lebanese
Lebanese	White/MENA	Wrote in Middle Eastern/Checked Lebanese
Lebanese	Black AA/Hispanic/MENA	Nigerian/Puerto Rican/Lebanese (wrote in)
Mauritanian	Black/African Am	Mauritanian (wrote in)
Mauritanian	MENA	Mauritania (wrote in)
Turkish	Some other race	Turkish
Turkish	MENA	Turkish
Turkish	White	Turkish
Somali	Black/African American	Somali-American (wrote in)
Somali	Black/African American	Somali
Sudanese	Black/African American	Sudanese (wrote in)
Sudanese	Black/African American	Sudanese (wrote in)

5.9.2 Confusion with when to use “American”

Similar to results in prior usability tests, we continued to observe parents who were first generation immigrants unsure how to identify their children’s detailed origins. Many said, “My children are American.” They were born here not in xx (e.g., where xx is the respondents’ country of origin, outside the U.S.). For one participant, of Chaldean descent, he said, “The children were born here, they have never been to Iraq. They are white.” He selected White and then saw the examples for White (e.g., German, English, Irish.) Seeing these examples, he changed his mind and he went back to the previous screen and selected the MENA response category and wrote in “Iraqi” for his children. Another participant commented that she had to think about how to list her son since he was born in the U.S. She said, “We [the parents] are from Somalia but my children have never been there.” Ultimately, she decided to list her children as Somali-American.

One participant of Israeli origin chose “White” and wrote “American” saying that he is now a U.S. citizen. However, he was also looking for a specific religion (e.g., Jewish) to be listed when asked for his detailed origins.

Usability staff recommended that the help text could address this question of first generation parents for what to enter for their American-born children. This recommendation was not accepted, and we continued to see this response pattern among first-generation immigrants.

5.9.3 Feedback on question stem “race or ethnicity” versus “categories”

It appeared that participants were not reading the question stem too closely as there were no spontaneous comments on either the question stem with “race or ethnicity” or the question stem

that used the word “categories.” Instead, participants deduced what was being asked from the response choices. Even people who were not born in this country knew it was asking about race. During debriefing, when probed on the question stems, participants offered additional feedback. Comments received about the use of the term “categories” included, “This is how we divide people” and, “what categories are we talking about?” This participant continued that it made her feel like we want information on “what she looks like.” Another participant said the use of the term “categories” was more politically correct than the word “race.” Another participant rephrased the question to say, “What category race would I put.”

A participant when probed about the term “race or ethnicity” said, “Race is how you are read by other people, that is how it matters the most.” When we asked the participants if they had a preference between using “race or ethnicity” or “categories” there seemed to be a pretty even split.

5.9.4 “Enter details” instruction on Race screen is confusing

In Figure 26, the instruction, “Select one or more boxes AND enter details” generated some feedback from a few participants. They said that there was no space to add details on this screen, so it was confusing to have that instruction on this screen. The details could be added on the next screen. As this next screen has its own instruction of “AND enter details” we recommended that same instruction was not needed on the initial race screen and could be removed to clear up any confusion. This change was made prior to the 2016 CT.

5.9.5 Inconsistencies with “And so on” and “etc.”

Figure 26 showed some inconsistencies with the use of terms “and so on” and “etc.” While no participants spontaneously commented on the inconsistency, we recommended only using one term or the other term, not both. This change was implemented and by 2016 CT usability testing the screens all used “etc.”

5.9.6 Feedback on instructions for selecting more than one race

We did not observe any difference in reporting multiple races based on the different instructional text below the question. Some screens had two instructions (e.g., Select one or more boxes AND enter details. Note: you may report more than one group) while other screens tested just had the one mention of reporting multiple races (e.g., Select one or more boxes AND enter details, see Figure 25 and Figure 26). Not many participants appeared to be reading the instruction verbatim but most of them reported only a single race, see Table 5. During debriefing when participants were asked if they thought the two instructions were different, all reported that they basically meant the same thing. Based on this and the fact that less clutter on the screen is considered better, the usability team advised to only keep one instruction, in a regular black font (not blue as people may think they are links) that read something along the lines of “Select all that apply.” This recommendation was not adopted. In a recent study on instructions, Horwitz, Nichols & Coombs (2018), did not find that having the instruction “*check all that apply*” on the screen or having it behind an additional “help” link impacted the number of items that were checked. They did recommend additional testing on this topic.

5.9.7 Open-text field on the detailed origins screen is not immediately apparent

We noticed that when the detailed origins screen used checkboxes, one participant, who did not see his country listed in either the checkboxes or in the examples in the unnamed other field, navigated back a screen and modified his initial selection to “Some Other Race.” Consequently, it could be that if participants looking for their detailed origin to be listed and do not see it, they may think their “detailed origin” is not part of that race and ethnicity group. A fix that may address this problem would be to have the open-text field on the same level as the other checkboxes. Label the checkbox “Other” or “Other ethnicities” and have a specify box to the right or below the examples. See the original screen in Figure 28 and a mockup of the modification in Figure 29. This change was not implemented for the next round of testing.



The screenshot shows the United States Census Bureau's 2015 National Content Test interface. The header includes the Census Bureau logo and the title "2015 National Content Test". Navigation links for "Instructions", "FAQs", and "Save and Log Out" are visible. The main content area displays a user's selection of "White" and prompts for detailed ethnicity information. A list of checkboxes includes German, Irish, English, Italian, Polish, and French. Below this list is an open-text field with a placeholder text: "Enter, for example, Scottish, Norwegian, Dutch, etc.". "Previous" and "Next" buttons are located at the bottom of the form.

Figure 28. Screenshot of how 2015 PC version looked with open-text field not in-line with the other ethnicities.

The screenshot shows the top header of the United States Census Bureau's 2015 National Content Test. Below the header, there are navigation links for 'Instructions', 'FAQs', and 'Save and Log Out'. The main content area displays the text 'You said that John K Doe is: White'. Below this, it states 'Next, we will collect detailed information for each race or origin listed above.' The question is 'What are John K Doe's White ethnicities? Enter origin(s), for example, German, Irish, English, Italian, Polish, French, and so on. (Help)'. There is a list of checkboxes for 'German', 'Irish', 'English', 'Italian', 'Polish', 'French', and 'Other'. The 'Other' option has a text input field next to it with the prompt 'Enter, for example, Scottish, Norwegian, Dutch, etc.'. At the bottom of the form, there are 'Previous' and 'Next' buttons.

Figure 29. Mock up with slight modification for the open-text field to have a check box in front of it an “other” option with examples listed out.

5.9.8 Gathering detailed origins on multiple screens caused additional respondent burden

We noticed that when the open-text field design asked for race, if a participant had marked more than one race, that participant did not anticipate that they would be asked to add in both ethnicities one screen after the other. For example, one participant of Israeli descent who checked “White” and “Some Other Race” got confused because there were two detailed origins screens. He received the first open-text field and did not want to enter a White detailed origin because the examples were from Europe and he was born in Israel from Russian and Argentinian ancestry. He got the error message because he left it blank, he then typed “Jewish” and selected “next.” This took him to the second “Some other race” detailed origins screen and he felt like he was stuck in a loop. He typed “None” for the open-text field on the “Some other race” detailed origin screen. A potential fix for this design could be to combine the two (or more) detailed origins onto one screen if multiple races were chosen. We have created a rough mock up in Figure 30 with two detailed origins being collected on one screen. This recommendation was not adopted.

United States™
Census Bureau

2015 National Content Test

AN OFFICIAL WEBSITE OF THE UNITED STATES GOVERNMENT

Instructions | FAQs | Save and Log Out

Next, we will collect detailed information for each race or ethnicity selected.

You said that Jane Doe is:
Black or African Am.

What are Jane Doe's **Black or African Am.** ethnicities? — Enter, for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, etc. Note, you may report more than one group. [\(Help\)](#)

You said that John K Doe is:
White

What are John K Doe's specific **White** categories? Enter origin(s), for example, German, Irish, English, Italian, Polish, French, and so on. [\(Help\)](#)

[Previous](#) [Next](#)

OMB No.: 0607-0985 Approval Expires: 6/30/2018 [Accessibility](#) [Privacy](#) [Security](#)

Figure 30. Mock up with slight modification of collecting multiple additional ethnic details on one screen.

5.9.9 Preference on checkboxes vs open-text design (detailed information for race)

During the debriefing, TAs showed the participant both designs for gathering detailed origins, the one with the checkboxes (Figure 23) and the one with the open-text field (Figure 24). In general, we noticed that if the participant could see their detailed origin listed, they preferred the checkbox design. If the participant did not see their detailed origin listed, they preferred the open-text field design.

5.9.10 Combined race and ethnicity question worked for participants – but a potential of fewer detailed origins for Puerto Rican participants

There were no negative comments on the combined race and ethnicity question. In this round of testing, we had the opportunity to work with three Puerto Rican participants. Two out of three Puerto Rican participants selected only one category (i.e., “Hispanic”) when they had selected both a race (Black/African American) and a Hispanic origin (Puerto Rican) on the paper questionnaire that was given to them at the beginning of the study. The other participant selected White and Hispanic and then in the White detailed origins box typed Puerto Rican and in the Hispanic detailed origins checked the Puerto Rican checkbox. This is not a usability issue; however, it is possible that there will be fewer detailed origins on respondents who identify as Puerto Rican when using

a combined race and ethnicity question. All three participants described above had no negative comments with having Hispanic origin included as a response option for the combined race and ethnicity question.

5.10 Design of the 2016 Census Test Race Screens

The 2016 CT used a combined question on race and ethnicity, however the question stem did not use the words “race” or “origin.” The question was, “Which categories describe <NAME>?” The instructions read, “Select all boxes that apply. Note, you may report more than one group.” The instructions for this test and onward were no longer in blue font but in black font. The response choices combined the race and Hispanic origin response options, and included the category of MENA tested the prior year. There was one minor change made to the response choice of “Black or African American.” In one panel, the word “American” was spelled out. In the other panel, the word American was abbreviated to Am so it was listed as “Black or African Am.” This abbreviation had been used on prior census tests and was a carry-over from the paper questionnaire, which has space limitations compared to the online questionnaire. See Figure 31 through Figure 33.

The screenshot shows the top navigation bar of the United States Census Bureau website with links for FAQ, Instructions, and Burden, and a language/account dropdown. Below is a breadcrumb trail: Home > Person > John C Doe > Race. The main heading is "Which categories describe John C Doe?". Instructions state: "Select all boxes that apply. Note, you may report more than one group." The list of categories includes: White (German, Irish, English, Italian, Polish, French, etc.); Hispanic, Latino, or Spanish (Mexican or Mexican American, Puerto Rican, Cuban, Salvadoran, Dominican, Colombian, etc.); Black or African Am. (African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, etc.); Asian (Chinese, Filipino, Asian Indian, Vietnamese, Korean, Japanese, etc.); American Indian or Alaska Native (Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Native Village of Barrow Inupiat Traditional Government, Nome Eskimo Community, etc.); Middle Eastern or North African (Lebanese, Iranian, Egyptian, Syrian, Moroccan, Algerian, etc.); Native Hawaiian or Other Pacific Islander (Native Hawaiian, Samoan, Chamorro, Tongan, Fijian, Marshallese, etc.); and Some other race, ethnicity, or origin. Navigation buttons for "Previous" and "Next" are at the bottom. Footer text includes: Build: 53b4379 | OMB No.: 0607-0979 | Approval Expires: 5/31/2015 and Accessibility | Privacy | Security.

Figure 31. 2016 CT PC screen shot of race panel where American is not spelled out in “Black or African Am.”

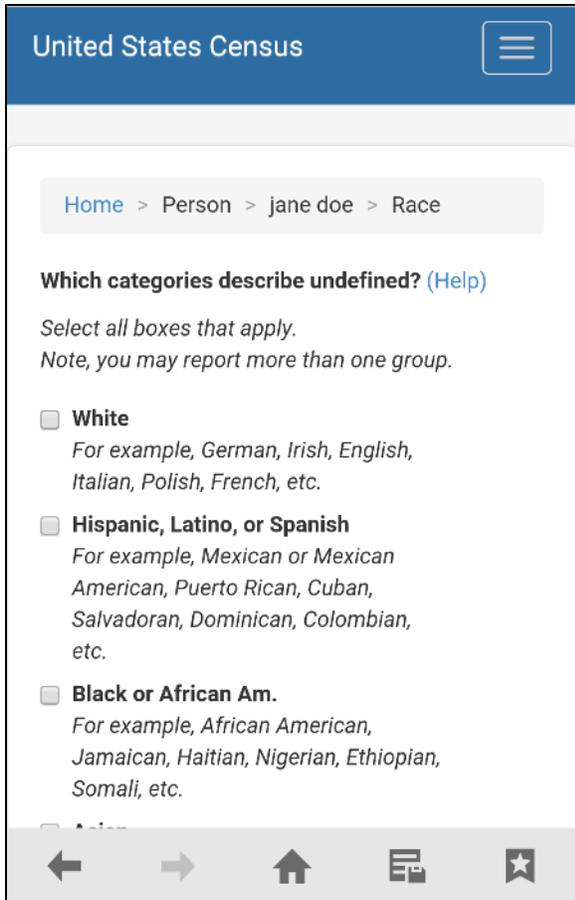


Figure 32. 2016 CT mobile screen shot of race panel where American is not spelled out in “Black or African Am.” Here “undefined” would be the respondent’s name.

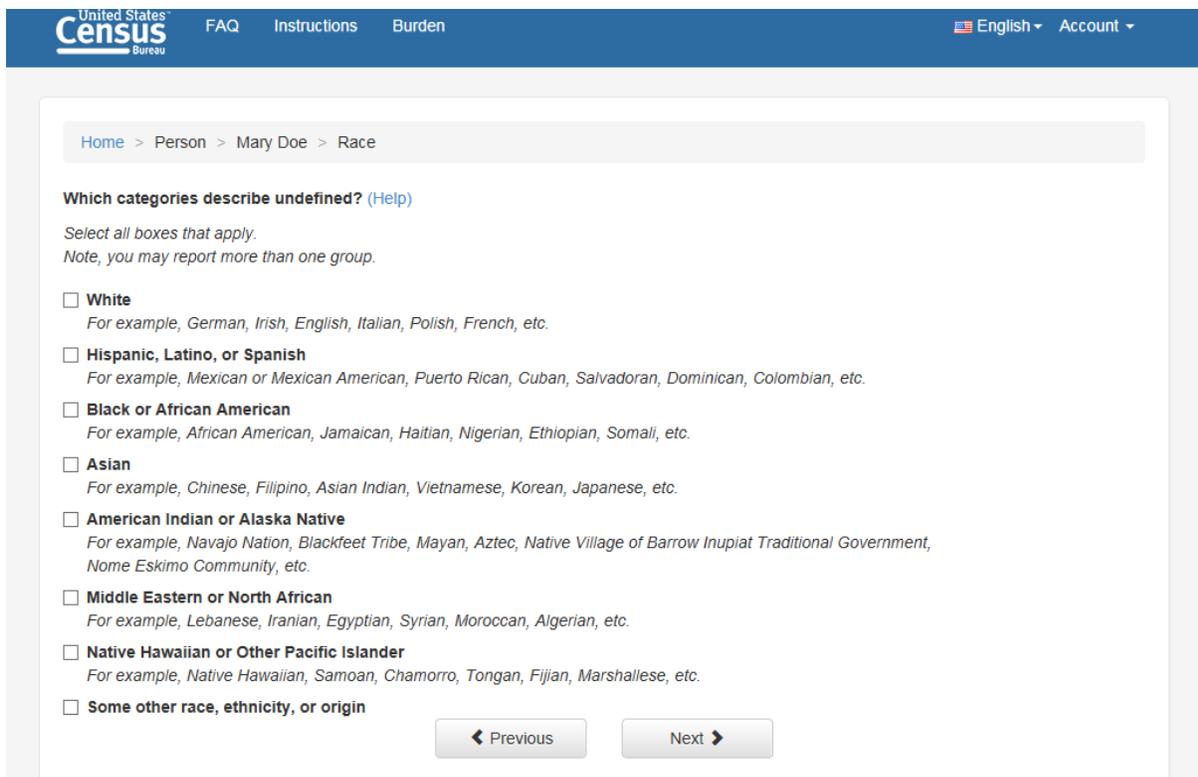


Figure 33. 2016 CT PC screenshot of race panel where American is spelled out in “Black or African American.” Here “undefined” would be the respondent’s name.

Like the previous studies, there were follow-up questions on the participants’ detailed origins. If White, Hispanic, Black or African American, Asian, Middle Eastern or North African, Native Hawaiian or Pacific Islander, or Some Other Race was selected, the next screen contained a question about the detailed origins for the selected category that the respondent had chosen (it would be on multiple screens if multiple races were selected). That question’s response option design used a list of checkboxes for the six most frequently occurring details for that race group in the U.S. based on American Community Survey data. There was also an open-text field if none of those choices applied to the respondent, similar to one of the panels in the 2015 NCT. If the American Indian or Alaska Native (AIAN) race response option was selected, the next screen did not have six checkboxes. Instead, the screen contained examples and an open-text field. If the Some Other Race response option was selected, the next screen had no checkboxes but did include an open-text field to enter detailed origins. Screenshot examples are shown for the Asian category, in PC (Figure 34) and mobile (Figure 35) and AIAN (Figure 36).

[Home](#) > [Person](#) > [John X Doe](#) > [Asian Categories](#)

What are John X Doe's specific ASIAN categories? [\(Help\)](#)

Select all boxes that apply and/or enter details in the space below.
Note, you may report more than one group.

- Chinese
- Filipino
- Asian Indian
- Vietnamese
- Korean
- Japanese

Enter, for example, Pakistani, Cambodian, Hmong, etc.

[◀ Previous](#)

[Next ▶](#)

Figure 34. 2016 CT PC screenshot for the detailed origins for Asian.

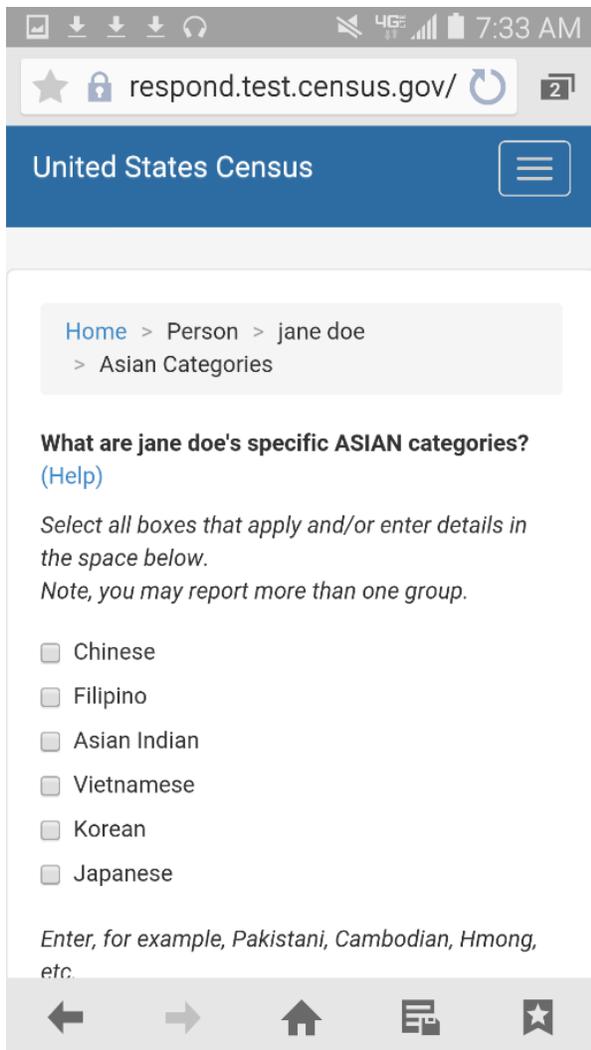


Figure 35. 2016CT mobile screenshot for the detailed origins for Asian.

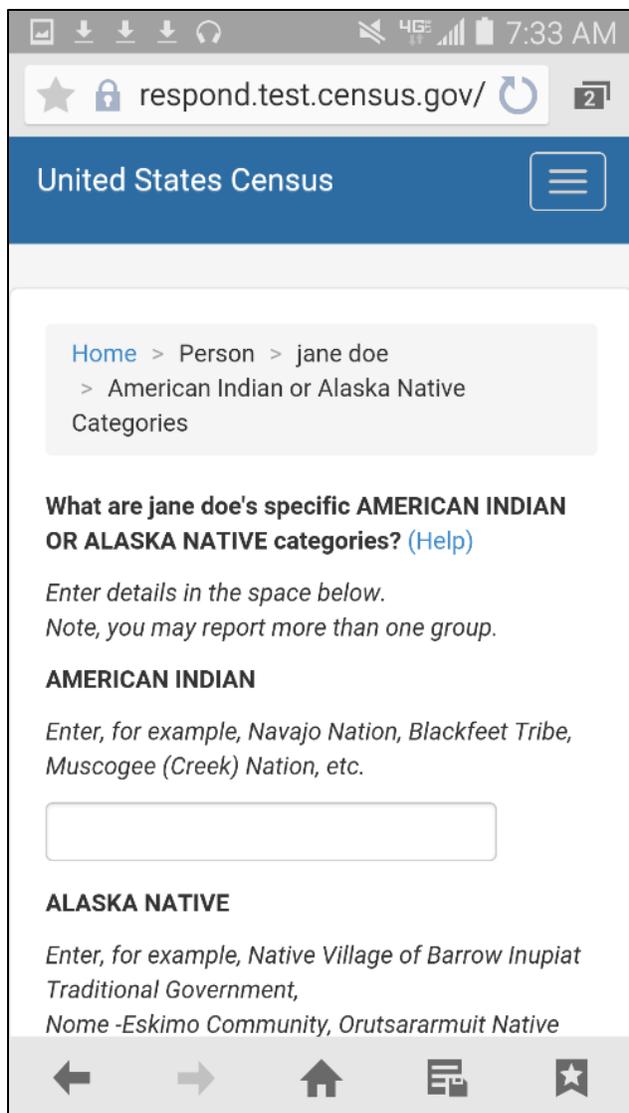


Figure 36. 2016 CT mobile screenshot for the detailed origins for the American Indian or Alaska Native

5.11 Usability Findings of the 2016 Census Test Race Screens

In this test, there was one race and one detailed origin design. The combined question on race and ethnicity was used, including the new MENA category with detailed origins collected on the next screen in the form of checkboxes and an open-text field for detailed origins beyond the six listed. There was a slight difference in the way the response option for “Black or African American” appeared. In one “American” was spelled out, in the other it was abbreviated.

5.11.1 Participants appear to know what the race question is by looking at the response choices. The question stem without the word “race” was not what most expected but also did not cause issues.

Just like the 2015 NCT, there were no spontaneous comments on the question stem for race, “Which categories describe <name>?” Eye-tracking results on the four participants where we have data show that all participants on seeing the screen for the first time, look first at the response options and then look at the question stem, and only after that look at the instruction text. For the second time and all future times seeing the screen (for the next person/people in their household) participants look at the response choices but only two out of four participants look to the question stem and instruction text. Spontaneous words uttered on this screen included: “the race list,” “race,” “ethnicity” and “nationality.” These results suggest that for the self-response census questionnaire the race question stem is not as important as the answer choices.

During debriefing, we showed participants the combined question on race and ethnicity and we asked each of them what the question meant. One participant said that the question (e.g., what categories describe <name>” was asking what your genetic background was. Another participant said that the question was about races, but the categories were not all races. She said that Hispanic is an ethnicity and that ethnicities and races are usually separate questions. She had never seen Middle Eastern before as a race category. Another participant said of the word ‘category’ that, “it doesn’t bring in ethnicity.” She continued, “Category makes it feel like an object and not a sense of being.” She said that the word “category, threw her off,” but we did not observe any noticeable sign of that during the session. One participant said that she saw the “races” [meaning the response options] and did not see the question. When we pointed out the question stem in particular, she said the question sounded confusing to her. She would have reworded it to, “What race do you identify with the most or what is your ethnicity?” She said, “I don’t like the words ‘categories describe.’” Another participant said that the question means, “How people identify themselves, their physical background.” “Categories” means choices or groups. She said that the choices are standard, White, Black, Hispanic, etc. Her question was how far back in her history to go. She said that she is 1/32 American Indian. Although she did not select American Indian, she was not sure how specific we want. Another participant did not comment on the word “categories” but said the choices were “races.” He said that some people get annoyed about using the word race compared to ethnicity, but that he did not care. Another participant also did not comment on the question stem, but when asked how she would write a question for those choices, she said, “Please state your race.” Still another participant said it was not unusual that we did not use the term ‘race’ even though she said that was what we were looking for. She also said she could have glanced at the question, but went right to her response choice and moved on. Based on this feedback we recommended using the question, “Which race(s) or ethnicities describe <name>?” A modification of this wording was incorporated for the 2017 CT.

5.11.2 Feedback on response choices to the combined question on race and ethnicity

The response choices generated some comments, but everyone appeared to answer the combined question on race and ethnicity consistent with how they had answered previous demographic questions in the background section of the usability evaluation.

The new MENA category did not generate any negative comments and the only comment during the debriefing was from the participant who said she had never seen the response choice available before on other surveys.

The Black or African American category elicited spontaneous comments from two participants who self-identified with that race. One person said she did not like that Black and African American were combined, because she said she had never been to Africa and “there are people from Africa, like Charlize Theron, who would not be called African American.” Another participant said that when she was younger, Black and African American were synonymous, but now African American means African immigrants.

The other spontaneous comment on the response choices was about the lack of guidance on how Brazilians should answer. Brazilian is not provided as an example for any of the response choices. The participant who made this comment had family friends who were Brazilians and said that in his experience Brazilians do not know where to classify themselves as they do not always consider themselves Hispanic. He commented that there were some choices, like Chamorro which he had never heard of before, compared to Brazil, which he said was a big country. It was recommended that there be some additional text to the help section on what to do in these situations. This recommendation was not implemented.

5.11.3 Preference for African American to be spelled out, but abbreviation did not cause any measurement error

Five participants were presented with the race list with the category with “African American” spelled out and six participants received the version with “African Am.” (American abbreviated). No one spontaneously commented on the abbreviated form. Thus, it did not appear that anyone noticed it during the sessions. However, when asked about it during debriefing, almost everyone said that they preferred American spelled out. One participant said that the abbreviation is fine but it should be used elsewhere. She said that it couldn’t be that we were just trying to save space. “Don’t know what the point would be to abbreviate it.” One participant said that she like the word American spelled out and another said that she wouldn’t care if it were abbreviated, but noted that the other race groups were not abbreviated. One participant said she preferred it spelled out because everything else is spelled out. Another participant said she wouldn’t have noticed it and said that everyone was texting these days implying that you might type an abbreviation if you were texting. Another participant said we should spell out American, and asked what Am. was an abbreviation for. Jokingly he said, “‘Amateur’ as compared to ‘I’m a professional African.’” The last two participants also said they would like Am. to be spelled out. Though not the focus of this paper, screen readers that read off the word do not say American but instead say “A.M.” as though it is referring to the time of day. Based on the almost unanimous feedback from participants as well as for 508-compliance we recommended spelling out the word American and that recommendation was accepted.

5.11.4 Feedback on instructions for selecting more than one race

In the 2016 CT, there were two instructions that appeared just under the response question stem. The first read, “Select all boxes that apply.” The next line read, “Note, you may report more than one group.” As the participants were filling out their census questionnaire, we received no spontaneous comments about the instructions. The four participants’ eye-tracking data that we captured, (one was multi-racial and one lived with someone who was multi-racial,) shows that the first time on the race screen all saw the instruction and on average spent 5 seconds looking at the

instruction. This was approximately 5 percent of the time spent on the screen. When the participant revisits this screen for other household members, the time spent on reading the instructions dropped by a second to 4 seconds which was about 4 percent of the time spent on the screen.

During the live session, one participant selected multiple categories for her race. During debriefing, when asked whether the participant could select more than one race, a participant that had marked only one race said, “I have a cousin who is white and black, and there is no joint category.” Everyone else said that it was possible to select more than one race, but not all said they saw the instruction. Once the TA pointed out the two instructions, six out of 10 participants said they meant the same thing and it did not matter that there were two instructions. A few participants said you could simply leave it as, “Select all that apply.” However, one participant said that the instruction, “Note, you may report more than one group” made it seem like she was supposed to report for her entire household on that screen. Perhaps the word “group” made her think that way. During her actual session, she said that she did not read those instructions. Another participant said the instructions seemed the same, but that was okay. She said, “It was just letting people know that they are allowed to select more than one race group.”

The recommendation from the usability team was to keep only one instruction as it appeared sufficient for the participants and would eliminate some clutter on the screen. To reiterate the concept, the second instruction could be placed in help. This recommendation was not implemented.

5.11.5 Wide variability in reporting detailed origins

There was confusion and conceptual differences when reporting ones’ detailed origins about how far back to go in a person’s family ancestry. For example, one participant (a first generation American) did not report the same detailed origins for herself that she did for her parents while others went back three or four generations. One participant explicitly stated during debriefing that she was not sure how far back she should go to report her “ancestry” information.

5.11.6 African American as race and as detailed origin

One participant said in debriefing that if African American was not a choice she would have written “NA” in the other box, since the word “African American” was already captured on the initial race screen. With another participant there was a sense that having African American on the initial race screen and again in the detailed origins screen was redundant. One participant was not sure whether to check African American for one of her roommates who was Haitian and Jamaican. She said in life, this person would not say she was African American, but the respondent selected it anyway. She said “African American is all inclusive.” The participant who was upset by the combination of Black with African American on the race screen also selected African American on the detailed origins screen. During the debriefing, she said she considered just writing in Black in the open-text field. Another participant read the question aloud and selected African American and Jamaican for herself, and then selected African American for her partner saying, “He doesn’t know if he is any of these things so I will just put African American.” One participant said that she was not born in Cameroon and so chose “African Am” from the check box to distinguish

herself and her sibling from her parents who were Cameroon. She had written in “Cameroon” into the open-text field for her parents.

5.11.7 Use of checkboxes and open-text fields on the detailed race screen

Checkboxes on the detailed race screens could produce measurement error by exaggerating counts of the detailed origins of those displayed with checkboxes. Participants seemed quick to check the boxes that were available. One participant selected Irish, English and then said, “A little bit of German” and selected that choice. She also wrote in Swedish. One participant checked Irish, German and English, but during the debriefing said he would not have typed in those choices if it had only been an open-text field. He was only 1/8th English. He asked, “Does it become a leading question?” with the choices as selections. Another participant selected two boxes and wrote in something else. It is unclear whether participants would be as willing to type out the detailed origins (or write - if on paper) if there was only an open-text field. This census test was the first time we observed this behavior with the checkboxes. One reason may be that we had fewer people working with the checkboxes in earlier tests. Usually there were two or three different race questions being tested so not all participants interacted with the checkbox version of the question which may account for why this behavior was not noted in earlier rounds of testing.

5.11.8 Open-text field on the detailed origins screen is not immediately apparent

The layout of the “other” open-text field on the detailed origins screen was not parallel to the other response options. In the design with the checkboxes for the most populous detailed origins, there is no checkbox in front of the open-text field. One participant commented spontaneously that she did not see her category and did not think the list was exhaustive. She later realized that she could enter her detailed origin in the open-text field.

However, other participants used the open-text field without commenting. Three participants used the open-text field, even with checkboxes present. One of the three did not check any other response choice but immediately entered her father’s country of origin saying, “It doesn’t have it there, but I will enter ‘AFRICAN NATION.’” Another participant selected a few checkboxes in front of the response choices and also used the open-text field to enter an detailed origin. Another participant did not select any checkboxes and instead typed Lithuanian and Scottish in the open-text field for herself. She said about the open-text field that “it was for putting other ancestries that were not already included in the list of checkboxes” which indicated she clearly understood the purpose of the open-text field.

Still since at least one participant did not initially see the open-text field, it makes sense to add a checkbox in front of the open-text field at the same level of importance as the listed examples, for a parallel design. See the original design in Figure 37 and a mockup of the recommendation in Figure 38. Note: This was recommended after the 2015 NCT as the same design then led to similar participant findings. This recommendation was not implemented.

Figure 37. PC 2016 CT detailed origin screen

Figure 38. Example of recommendation with a check box in front of the open-text field.

5.11.9 Feedback on the detailed origins question

More participants seemed to read the detailed origins question stem than they did the combined race and ethnicity question stem. Still, participants looked at the response choices on the detailed origins screen to know what we wanted. One participant spontaneously said “ancestry” when answering the detailed categories question. Another participant made a face when reading the detailed categories question.

Eye-tracking data for the Black or African American detailed race screen for three participants showed that they first looked to the response options within 1.7 seconds of getting to the screen, then they looked at the instruction, within 11 seconds of the screen, and then they looked to the question stem within the first 52 seconds on the screen. For the White detailed race screen, two participants first looked to the response options within the first 3 seconds of getting to the screen, then they looked at the instruction, within the first 11 seconds of the screen, and then they looked to the question stem after 44 seconds on the screen. For the American Indian or Alaska Native detailed race screen, the one participant who saw the screen never looked at the question stem and looked at the response options within 5 seconds of the screen opening, spending about 48 seconds on that area of the screen. This participant did read the post question instruction within 8 seconds of the screen loading but spent only 1.6 seconds looking at that area. All these examples highlight that participants first looked at the response choices before looking at the question stem.

5.12 Design of the 2017 Census Test Race Screens

In the 2017 CT usability sessions, we tested a similar race question to what had been tested in the prior years with a combined question on race and ethnicity with the additional MENA category added. However, the question stem wording asked for the respondent's "race or ethnicity" instead of asking, "What categories describe..." the respondent, as had been used in the 2016 CT. In addition, the layout looked slightly different with the race examples set off to the right, rather than directly below the race. The design also implemented a "swim lane" design with shading of every other row. The subsequent detailed origins question used the checkbox design as seen in the 2016 CT. See Figure 39 and Figure 40.

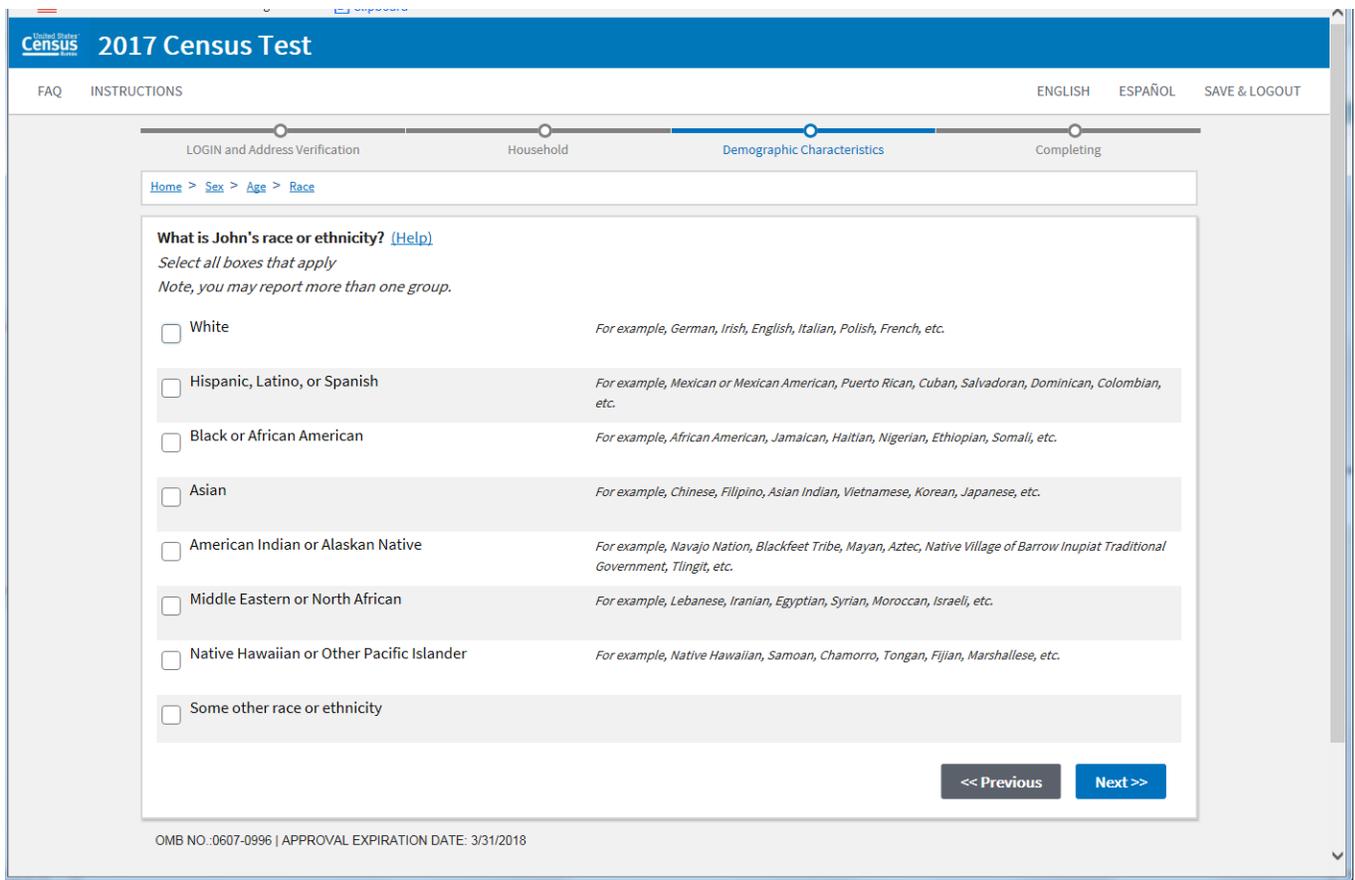


Figure 39. 2017 CT PC race screenshot where race examples are set off to the right and every other row is shaded.

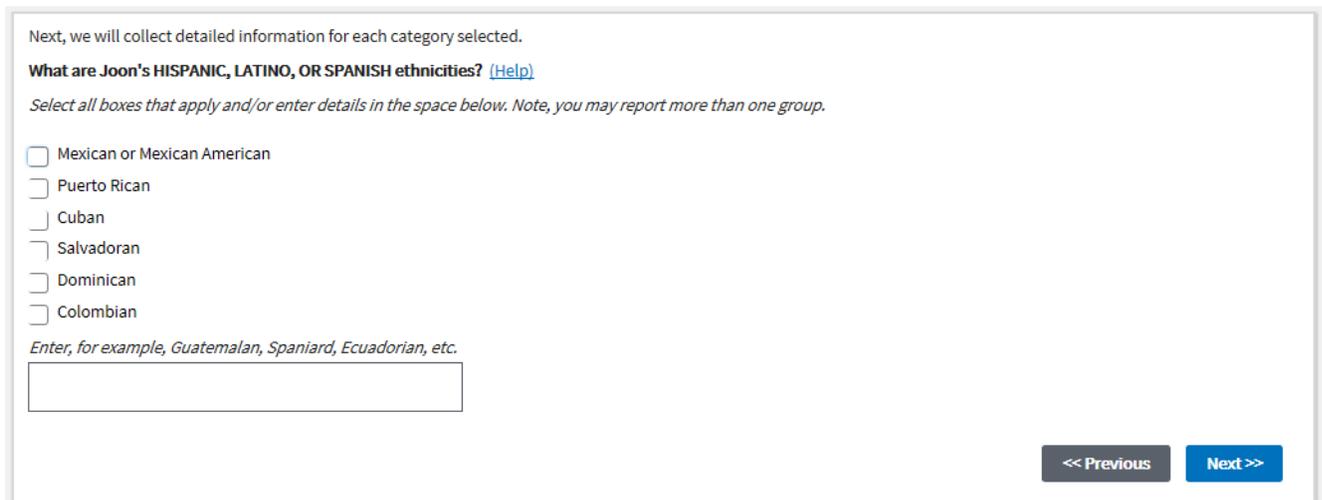


Figure 40. 2017 CT PC screenshot for the detailed origins for Hispanic origin.

5.13 Usability Findings of the 2017 Census Test Race Screens

There was only one design in 2017 tested: that of the combined race and ethnicity. As well there was only one design for the detailed origins screen. The combined race and ethnicity question

came first on one screen. The next screen had the detailed origins screen showing the six checkboxes and one open-text field based off of what had response option had been selected on the prior screen. On the combined race and ethnicity screen, the examples were in black font but off to the right-hand side of the screen on the PC design. In general, the question stem and the response choices for the race question worked well. We observed no problems with the new design where the examples were located to the far right. However, this was a small sample of users and may benefit from additional testing. One participant commented that people know what to choose and don't need the examples.

5.13.1 Feedback on instructions for selecting more than one race

Although one participant commented that we did not have a multi-racial category while she was answering the question, we did have another participant who selected more than one race during the session and understood that this was an available option to her. During debriefing the former participant said, "I really liked how you can do more than one. I didn't really see [initially] that you can report more than one. I was going fast." So during the session she did not see that reporting more than one race was allowed, (though she herself only had one race to report.) Then during debriefing when the instruction was pointed out to her, she was satisfied. Another participant said during debriefing about selecting her race, "I already know what my ethnicity is. There is no question about it. For someone who has multiple... oh it says you can choose more than one, so that's not a problem..." This user feedback (e.g., where the user initially did not need the instruction and then as she thought of the situation of answering more than one race and noticed the instruction) may indicate that when looking for information on what to do if you need to report more than one race, the instruction helps.

During debriefing, when asked about the post-question instructions, participants said they understood both instructions to mean the same thing (e.g., "Select all boxes that apply." And "Note you may report more than one group"). One participant said the instructions, "make it very clear that one can select multiple [boxes]."

5.13.2 Political climate, including Executive Order 13769, potentially influences race reporting

Two participants living in unrelated household situations mentioned that they would not feel comfortable answering the race of the unrelated people in their household based on current (2017) political climate. One participant explicitly mentioned the Muslim ban⁹, the other participant mentioned the president (Trump) as a rationale for why they did not feel comfortable to report on their roommates' race.

5.13.3 Feedback on the detailed race question

We heard from participants that the choices listed on the detailed race screens were about nationality and not ethnicity, which was similar to feedback in earlier rounds of testing. One participant commented on whether census was interested in his "heritage" or his "ethnicity." He

⁹ Executive Order 13769, titled Protecting the Nation from Foreign Terrorist Entry into the United States, also known as the Muslim ban or travel ban,

answered the question about his “heritage.” In addition, there were a few comments from participants about “how far back to go” with their “ancestry.” These came mostly from people who had been in America for at least a few generations and may not know specific detailed origins. As recommended earlier, this could be addressed in the help text. The recommendation was not accepted.

5.13.4 Use of checkboxes and open-text fields on the detailed race screen

Participants used both the checkboxes and open-text fields. We received feedback from participants asking why these choices were there (as the checkboxes). One participant who checked a few of the boxes said she would not have typed in all those responses if she had to type each one out, but many participants said they would have taken the time to type out each detailed origin that they had checked. As this was a small sample, it would be useful to test this further with additional empirical studies to determine if this design leads to measurement error.

5.13.5 Error messages encourage participants to enter a specific detailed origin, even when they do not have strong ties to that detailed origin, or do not know the detailed origin

Participants living with unrelated roommates were unsure what to enter for their roommate. One participant, after triggering the error message, did not know what to enter for her roommate. She clicked on the help link but did not find any useful information. The help text did not appear to address this issue. It is unclear what to enter if the respondent does not know the answer. Adding something to help on what to do in such a situation could be useful for respondents with unrelated roommates. This recommendation was not implemented.

5.13.6 Open-text field on the detailed origins screen is not immediately apparent

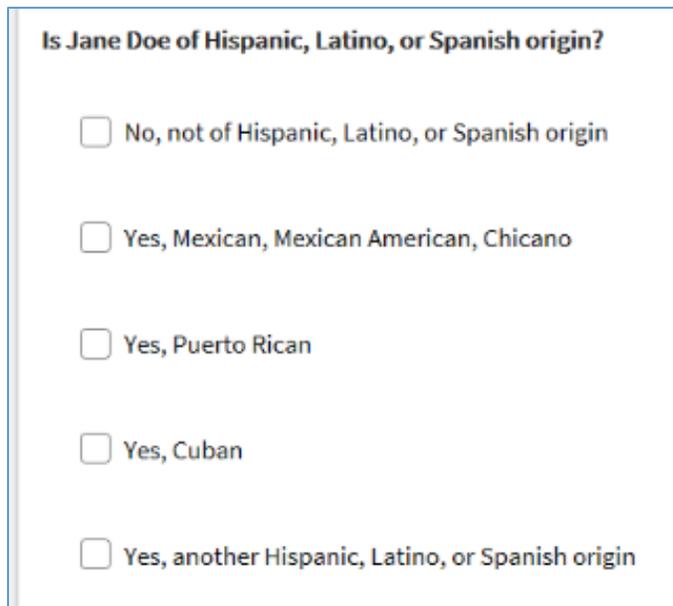
The layout of the “other” open-text field on the detailed origin screen was not parallel to the other response options. As in 2016 CT, the design of the last option on the detailed origins screen was an open-text field that lacked a check box in front of it. This is a non-standard way of including open-text fields. At least two participants said there should be an “other” check box before the open-text field. This feedback came up in earlier rounds of user testing. We recommended that the last field include a checkbox at the front so that the design is parallel to the prior options in the list. See the mock up from the last round of testing in Figure 38. This recommendation was not implemented.

5.14 Design of the 2020 Census Race Screens

At this writing, the 2020 Census has not occurred. However, the race question and the Hispanic origin question will return to two questions, similar to the 2010 Census. The Hispanic origin question comes first, on its own screen, in the online questionnaire as it did in the prior tests where the two-question version was tested. The Hispanic origin question comes first because the OMB standards state that ethnicity must come before race. Three separate rounds of usability testing occurred prior to the fielding of the 2020 Census and the screens tested had slight variations in each round which are described below.

5.14.1 Design of the 2020 Census Usability Test Round 1 Hispanic origin and Race Screens

In the first round of usability testing conducted for the 2020 Census, the Hispanic origin came first (see Figure 41) and the race question came next. The race response options were what had been used in the 2010 Census, and had been tested again in the 2014 CT and 2015 NCT. However, new with this design was that the race question displayed in an “unfolding” design. That is where initially all the participant could see was the list of races as the response options (Figure 42 and Figure 43). After the participant clicked on the race category, the examples would appear along with an open-text field asking the user to enter more specific details (Figure 44). There was a slight delay to the opening of the examples and the open-text field. If a participant attempted to go forward without filling out the open-text field, the screen would not move forward but instead remained on the same screen displaying an error message and the open-text field would be outlined in red (Figure 45). These screens are also shown in the mobile design (Figure 46 and Figure 47).



The screenshot shows a question titled "Is Jane Doe of Hispanic, Latino, or Spanish origin?". Below the title are five radio button options:

- No, not of Hispanic, Latino, or Spanish origin
- Yes, Mexican, Mexican American, Chicano
- Yes, Puerto Rican
- Yes, Cuban
- Yes, another Hispanic, Latino, or Spanish origin

Figure 41. 2020 Census R1 PC screenshot Hispanic origin question. Note: “Help” link not shown, as it was located far to left of screen.

What is Jane Doe's race?

Select one or more boxes **AND** enter origins. For this census, Hispanic origins are not races.

- White
- Black or African American
- American Indian or Alaska Native
- Chinese
- Filipino
- Asian Indian
- Vietnamese

Figure 42. 2020 Census R1 PC screenshot race question. Note: “Help” link not shown, as it was located far to left of screen.

Korean

Japanese

Other Asian

Native Hawaiian

Samoan

Chamorro

Other Pacific Islander

Some other race

Figure 43. 2020 Census R1 PC screenshot, race question scrolled down.

Black or African American

Enter, for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, etc.

Figure 44. 2020 R1 Census PC screenshot the “unfolding” design opened after a participant had clicked on the race. In this case the participant has clicked on the response “Black or African American.”

 Please provide a response. If this person does not have an answer, continue to the next page

What is Jane Doe's race?

Select one or more boxes **AND** enter origins. For this census, Hispanic origins are not races.

White

Enter, for example, German, Irish, English, Italian, Lebanese, Egyptian, etc.

Black or African American

American Indian or Alaska Native

Chinese

Filipino

Figure 45. 2020 Census R1 PC screenshot; display of race question after participant has selected a race but did not add in detailed origins before clicking on the next button. Note: “Help” link not shown, as it was located far to left of screen.

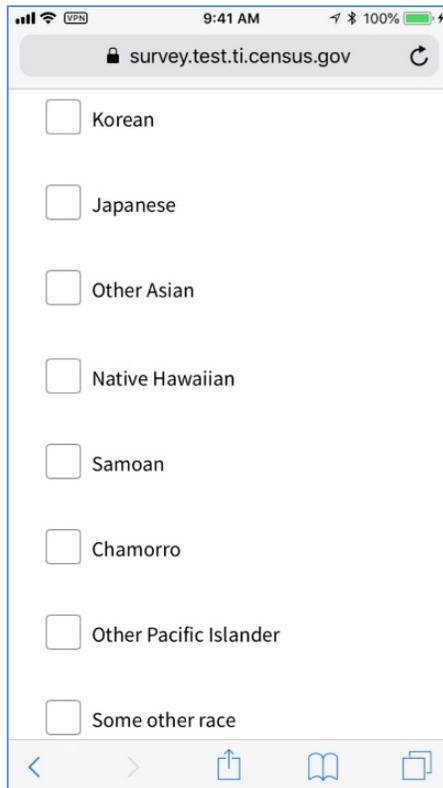
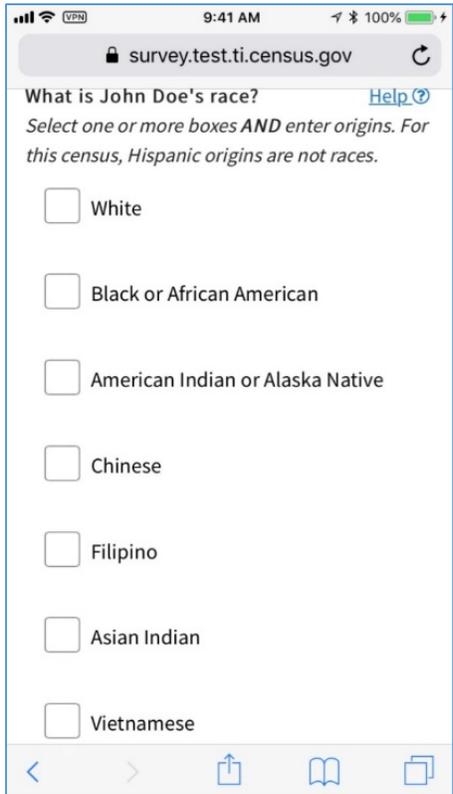
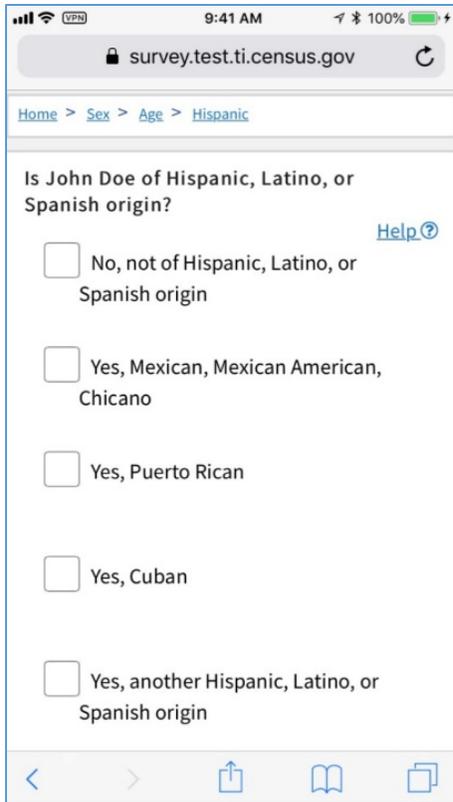


Figure 46. 2020 Census R1 Mobile screenshot; display of Hispanic origin and race screens

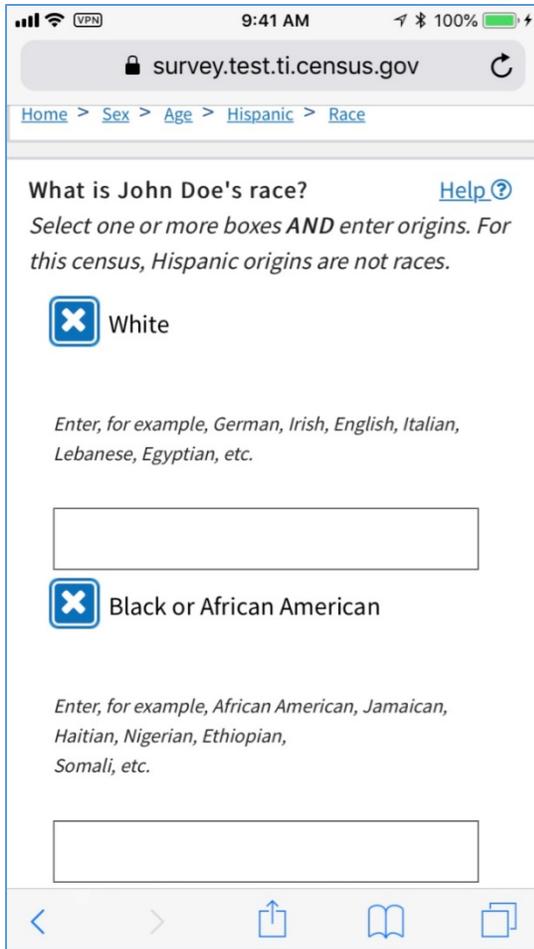


Figure 47. 2020 Census R1 Mobile screenshot; display of race screen after the open-text field has been activated.

5.14.2 Design of the 2020 Census Usability Test Round 2 Hispanic origin and Race Screens

During the second round of usability testing of the 2020 Census questionnaire, the Hispanic origin and race questions were identical to the first round with the exception that there was no unfolding design on the race screen. The detailed origins open-text field and examples were present on the screen from the initial loading of the screen. See Figure 48 and Figure 49.

5.14.3 Design of the 2020 Census Usability Test Round 3 Hispanic origin and Race Screens

During the third round of usability testing of the 2020 Census questionnaire, the Hispanic origin (Figure 48) and race questions (Figure 49) were identical to the second round with the detailed origins open-text fields and examples on the screen, but if no detailed origin was entered for a selected race and the next or previous button was selected, the open-text field was not outlined in red (Figure 50). Additionally, when a selection was made, no longer was the indicator a blue “x”; it changed instead to a solid blue fill.

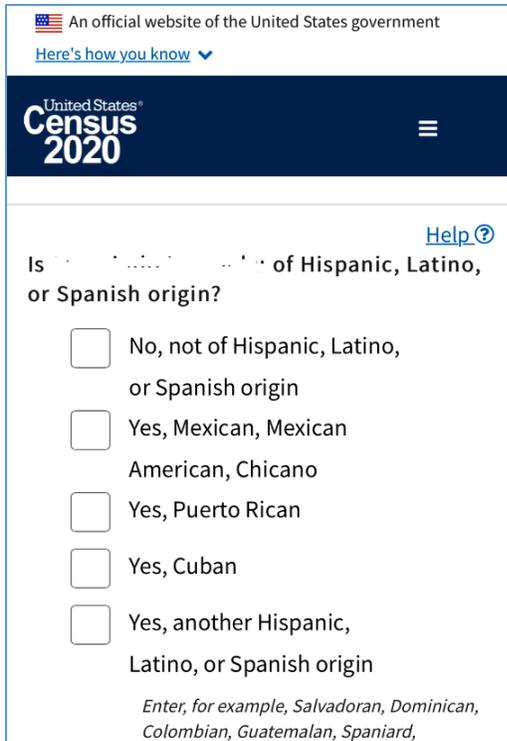


Figure 48. 2020 Census R2 Mobile screenshot; display of Hispanic origin screen on an iOS device

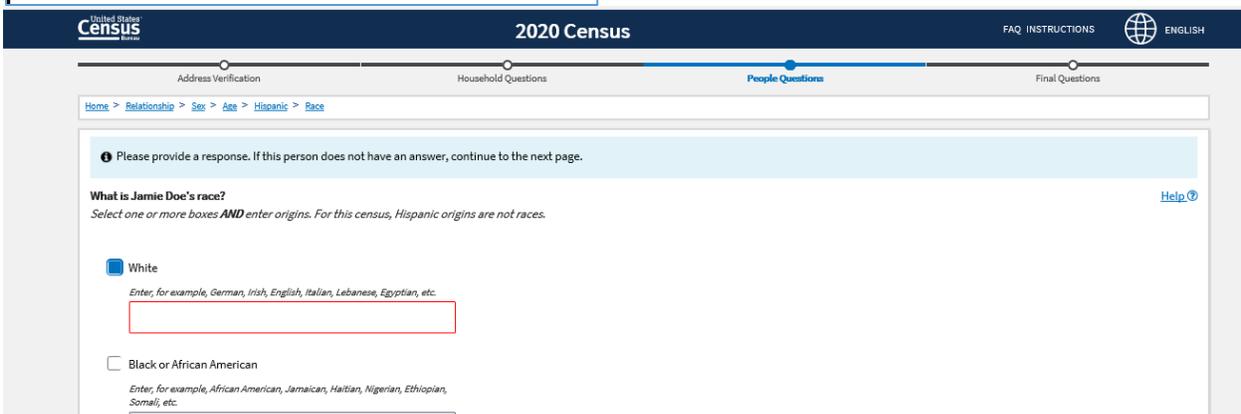
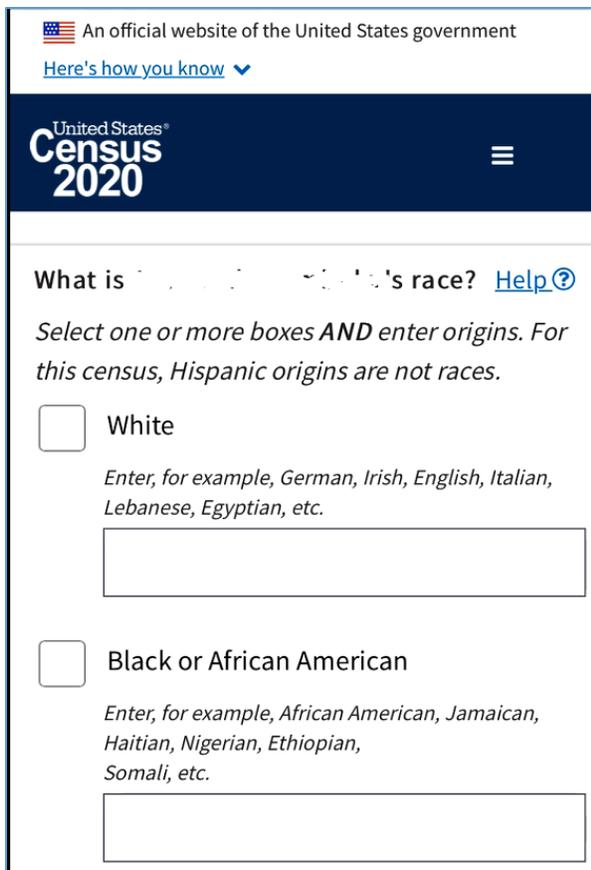


Figure 49. 2020 Census R2: Upper image of mobile screenshot display of race screen on an iOS device; lower image PC display of race screen with error message showing, note error message at top and open-text field highlighted in red.

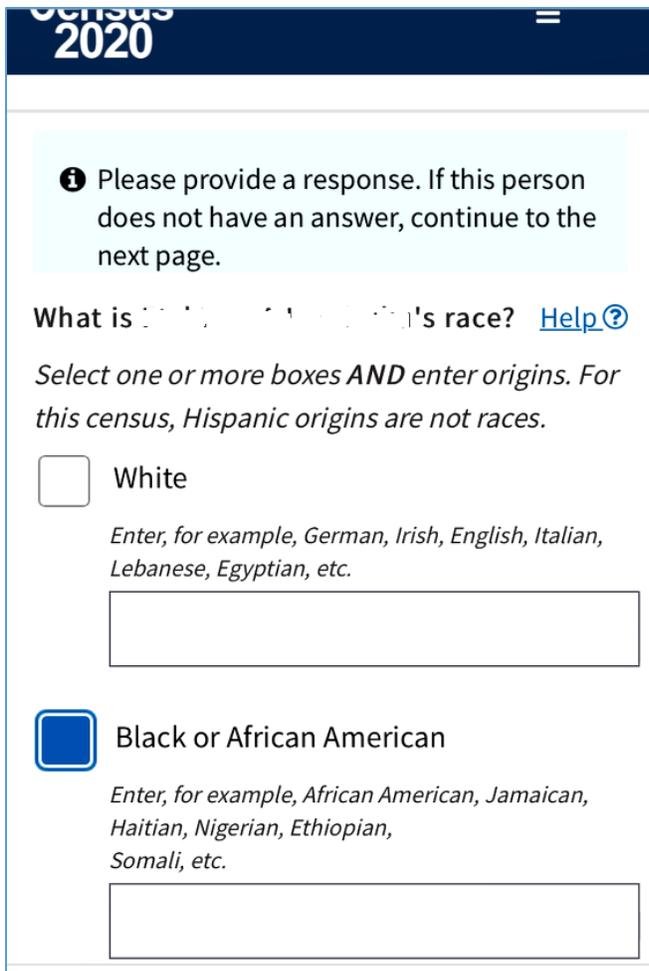


Figure 50. 2020 Census R3 Mobile screenshot; display of race screen with an error message because no detailed origin was entered (no red box around that field) prior to clicking on “next.”

5.14.4 Final Design of 2020 Hispanic origin and Race Screens

The 2020 Census will use separate Hispanic origin and race questions. The indicator of a selected response will be a solid blue fill. The detailed origin open-text fields and examples will always be present on the race screen (that is, no unfolding design will be used). If respondents leave a detailed field blank and try to move to the next screen, they will receive an error message and a red outline of the open-text field will appear. Respondents can ignore the error message, leave the field blank, and select the “next” navigation button again, which will move them to the next field. There will be no swim lane design, and the color will be black font on a white background, with the exception of the error message, which will be black font on a blue background.

5.15 Usability Findings of R1 2020 Census

There were many comments on the race screen this round of testing, likely because the race screen reverted to a version that had been used in the 2010 Census. This version has the Hispanic origin question first and then on the next screen, the race question without Hispanic or MENA as a

response option. This design had been usability tested as Version 3 in the 2014 CT as well as field tested in the 2015 NCT.¹⁰ We had heard similar comments in the TQA debriefings (Nichols & Olmsted-Hawala, 2018; Nichols, Olmsted-Hawala, & Katz, 2016) that we heard in this round of testing and from those participants that received Version 3 in the 2014 CT, mainly that having the Hispanic origin separate from the race question was confusing and, to some participants, insulting.

5.15.1 Separating Hispanic origin and race is confusing

One participant with a Hispanic roommate spontaneously commented that it was strange that Hispanic origin was not listed on the race screen. She said she had never heard of Chamorro “yet Hispanic is not on there.” Another participant read the instruction on the race screen, that said “Hispanic origin is not a race” and asked, “I’m just wondering is it always that way, it is not a race, or is it just for the census?” During debriefing this user said she thought it could be offensive for “Hispanic” not to be a race. Another participant said she did not understand why Hispanic origin and race were not just one question all together. Based on how problematic this question can be for participants and how participants do not have issues when race and Hispanic origin questions are combined into one question on race and ethnicity, the team recommended reverting to a combined question when possible. This recommendation was not implemented.

5.15.2 Open-text field on the detailed origins screen is not immediately apparent

In Round 1, the detailed origins open-text field did not appear until after a race was selected. Once a participant selected a race the open-text field dynamically appeared on the screen, beneath the selected race choice. Participants frequently missed the dynamic open-text field. Instead participants selected ‘next’ as they attempted to move forward. This action triggered an error message. Seven out of 10 participants received at least one error message on this screen because they did not complete the detailed race open-text field. Some participants missed this field repeatedly with subsequent household members (e.g., one participant fired this error 3 times - for each of her family members). On the laptop version of the question, the dynamic open-text field was slow and participants often scrolled down right after making their selection to the “next” button, and thus missed even seeing the dynamic open-text field.

This design changed towards the very end of our testing window so that one participant (the last participant in the study) saw a new design of the question. In the new design, the open-text fields were always available on the screen (e.g., no dynamic opening). See Figure 51. However, this participant also failed to fill out the open-text field and triggered an error message. When asked during debriefing about his experience he said, "When I clicked on White, I didn't realize I had to enter detailed information."

¹⁰ For the 2015 usability test of the NCT, we did not do any usability sessions using this version (e.g., Hispanic origin and race as two separate questions) due to it being the control that had been used in the prior census. As well, there were so many new versions to test and a relatively limited number of participants compared to all the different variants of the race screen. Thus in the 2015 NCT usability test we only had participants interact with the newer screen designs.

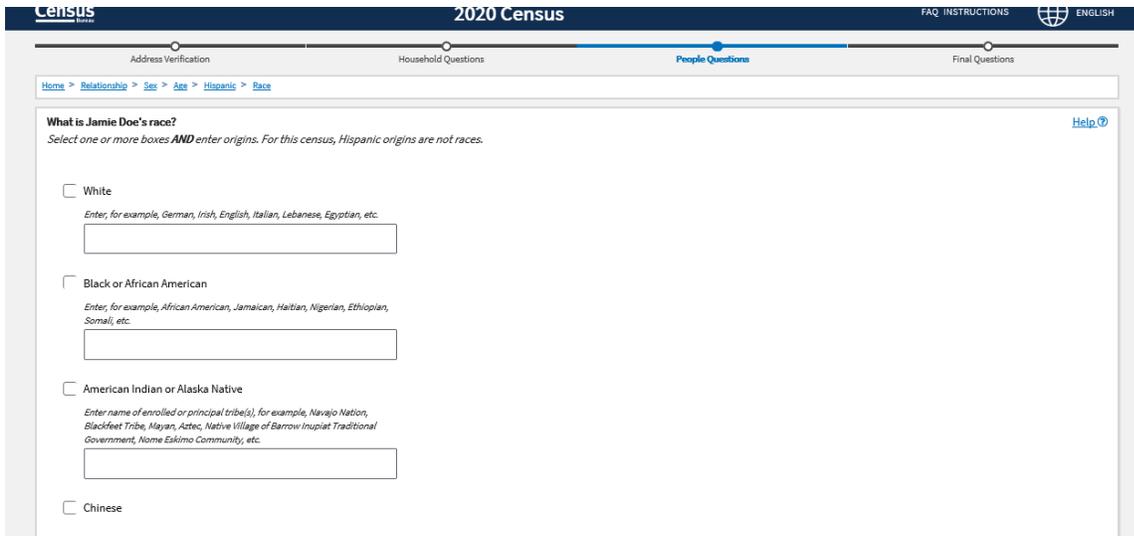


Figure 51. PC Screenshot of the 2020 Round 1 version that was implemented towards end of user testing cycle. Open-text field was available on screen from the screen loading.

We had noticed in earlier rounds of usability testing that participants who had initially left a field blank, upon getting an error message, filled out the field. So in our knowledge check administered in 2020 R1, R2 and R3, we asked the yes/no question of “If you don’t know the answer to a question on the Census, do you think you can leave it blank and move to the next question?” Six participants answered “yes” that they could have left it blank if they did not know the answer. Ten participants answered “no” and nine participants answered that they “didn’t know.” See Figure 52 This is a strong indicator that at least 19 out of 25 people were unaware that if they triggered an error message on the screen that they could leave the screen blank and move forward if they did not know the answer to the census question.

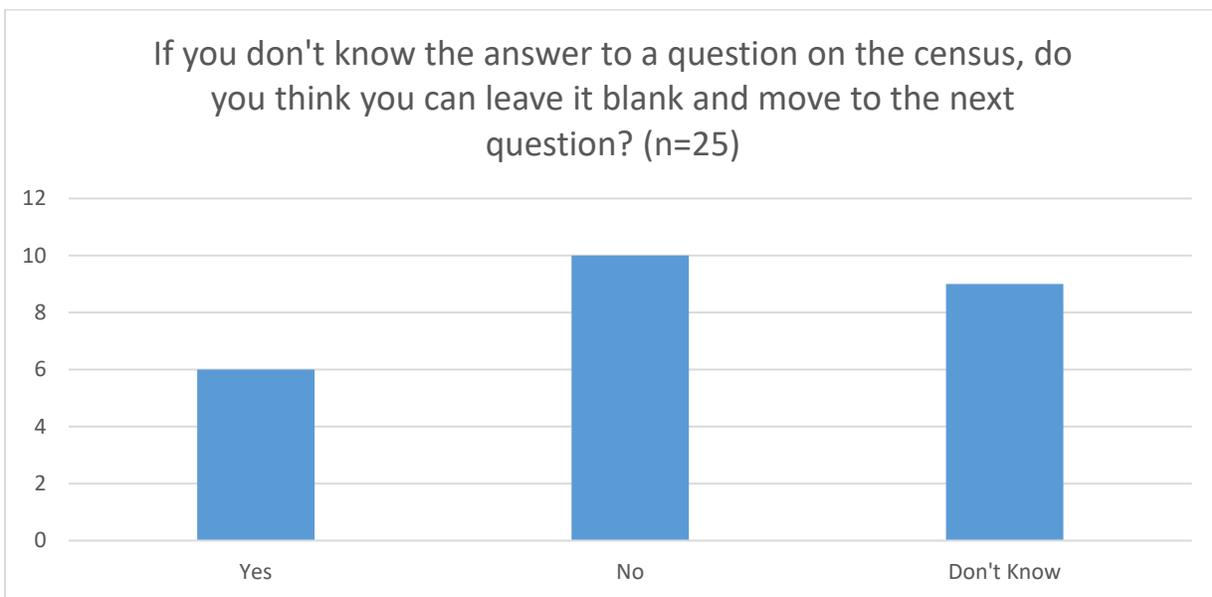


Figure 52. Chart of one question on the knowledge check asking if participants think they can leave a question blank if they don’t know the answer.

5.15.3 Race Error May Cause Measurement Error

An error message was triggered when respondents left the detailed origins open-text field blank. Participants who received the error message (See Figure 45) may have been under the impression that they had to put something in the open-text field in order to move forward. (While the error message is not a “hard edit” e.g., it can be bypassed if a person leaves the field blank for a second time and clicks “next” for a second time, it is rare for a participant to receive the error message, leave the field blank and attempt to navigate forward a second time.)

During testing, after triggering the error message, one woman said she did not understand why she had to enter African American again when she had just clicked on the black or African American race category. “It says for example African-American. I’ve already clicked on that box.” Another woman said she did not know what to put and did not know how far back the census wanted her to go. “It’s just plain white... I don’t identify any other way other than just American. It’s pushing me. It wants an answer. I could make something up.” After triggering the error message in our testing, no one left the box blank. One person, answering for an unrelated roommate, understood that she had to fill out the box and so wrote “not sure” into the open-text field in order to move forward. Another participant who triggered the race error message thought she had inadvertently marked an incorrect race and attempted to navigate backwards to fix her “error.” She used the browser navigation buttons, which corrupted her user session, and started her back at the beginning of the survey. She then blamed herself for doing something wrong. Because we cannot know if the open-text fields were intentionally left blank or not, we recommended removing the error messages on the race screen and simply allowing the participant to navigate forward the first time, even if the open-text field was left blank. This error message is a mode difference with paper and could potentially inflate the detailed origins from which people may not closely identify. This recommendation was not implemented.

5.15.4 Challenges when Reporting Race as a Proxy

In a proxy situation where a person answers the census questions for someone with whom they are typically unrelated (e.g., a roommate) the person may not know the correct answers. For example, we had a participant answer “Some other race” for his middle-eastern roommate. He then wrote in “Middle Eastern.” In another example, when marking for himself one participant marked “Asian Indian.” For his roommate he was unsure what to do. He said his roommate was also of Asian Indian descent, but unlike the participant, his roommate was born in America. He first selected American Indian. He saw the dynamic open-text field which displayed the examples and read the tribal names and realized it was not the correct race for his roommate. He went back to mark “Asian Indian” for his roommate. For situations where participants don’t know the answer to the questions, there appears to be some desire to have a “don’t know” option. One participant even wrote into the box “don’t know.” Other participants have asked for a way to save their information, log out, and then either they would have their roommates log back in to complete their own section or they say they would ask their roommates and then finish it up (by logging back in themselves) after getting the answer. The functionality for save and logout is not possible for the 2020 Census, however we report it to document it as a request that some participants made on how they would solve the issue of not knowing the roommate’s race information.

5.15.5 Spacing and formatting issues appearing on race screens

The spacing between race response options was very spread out. The numerous response options on the combined race and ethnic origin question required extensive scrolling by participants in order to find the correct response. The spacing between the response option and the open-text field made it seem like the open-text field corresponded to the subsequent race response option as it was closer to that race response option than the prior one. See Figure 53. The screens need to be re-formatted so that the race checkboxes are closer to their appropriate open-text field. In addition, the team should consider testing of a slightly more compact design with slightly less spacing between the response options. The second round of the 2015 CT (optimizing for mobile) had identified the appropriate amount of space between the response options. They were not too close together to trigger “fat finger” issues (that is where one inadvertently touches an incorrect response option due to not enough spacing or padding around a link or text field). And they were not too far away, as is occurring in the design below, to make one think the open-text field is somehow related to an inaccurate response option. See the image on the right in Figure 43 for a view of spacing that worked for participants.

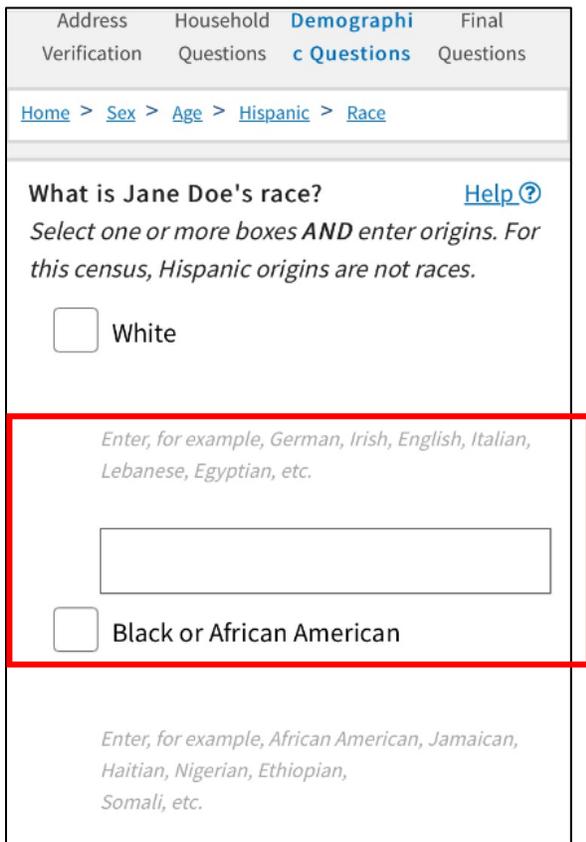


Figure 53. Mobile screenshot of white instruction text and open-text field appearing closer to Black or African American race checkbox than to the White race checkbox.

5.16 Usability Findings of R2 2020 Census

The design of the screen was the same as in Round 1 although this time the examples for the race categories showed from the time the screen loaded. There was still some delay that was occurring on the screen and this may have drawn users' eyes to the examples and the open-text field that accompanied some of the response choices for race. Many of the findings we heard in this round had been voiced in the previous round of testing.

5.16.1 Challenges when Reporting Race as a Proxy

Some of the roommates who were answering for people they did not know too well mentioned that they wanted a way to indicate that they were not completely sure about the answer they had given. Participants said, after marking the race of their roommate, "I'm not sure" or "I don't know the answer to this."

One participant who was answering for her roommates couldn't find one of the roommates race in the list provided. She was looking for "Indian" and did not see it as "Asian Indian." She was herself Chinese and did not consider Asian Indian to be appropriate for her roommate as she said, "he always goes home to India in the summer so I think he's Indian." She said, "I feel confused when I see 'Pacific Islander' and 'Some other race' cause I don't know where to put in 'Indian.' There is no 'Indian' here."

5.16.2 African American as race and as detailed origin

Two people who identified as African American said it was redundant to have to check a box and then type in the exact same words (e.g., African American) as well. One participant said, "That part is redundant. A little too much. To click and write in, it should just be one or the other."

5.16.3 Separating Hispanic origin and race is confusing

On seeing the race question after the Hispanic origin screen one participant said, "It feels like you would have already answered it on the previous page." This participant reiterated what we had heard in prior rounds about the two question format of Hispanic origin and race was confusing.

5.17 Usability Findings of R3 2020 Census

This final round of usability testing in anticipation of the 2020 Census was to see how users interacted with the mobile design. At the time of testing, the iOS design was more fully developed than the Android design. It is possible that during testing when the design was not performing as it supposed to be (e.g., as it was specified) the participants may have had more issues than they would otherwise have experienced.

5.17.1 Separating Hispanic origin and race is confusing

Again there was confusion with the Hispanic origin and race being on two different screens where two out of five participants had problems. One participant, a non-native English speaker, saw the Hispanic origins screen and said, as she looked at the response options, "no none of these." Then in the open-text field connected to the response choice: "Yes, another Hispanic, Latino, or Spanish

origin” the participant typed in “African American from Ethiopia.” On the next screen she selected a race but did not write in any detailed origin. This interaction occurred on the Android device and at the time of testing, the open-text fields were in the wrong location (see Figure 54). It is unclear if the screens had looked normal whether she would have entered the same information into the open-text field or not. She did this for each family member, never realizing that the first question was not about her race.



Figure 54. Screenshot of an Android smartphone where open-text field was overlaid on the next race response option.

Another example that demonstrates how separating the Hispanic origin and Race screens were confusing to users occurred when a participant said we were emphasizing Hispanic over everything else. She said that “being Hispanic or not is what the Census Bureau is most interested in.”

5.17.2 Spacing and formatting issues appearing on race screens

This was an issue in the mobile design noted earlier in Round 1. It occurred throughout all three rounds of testing of the 2020 mobile designs. In June of 2019 the first time we tested the screens

on Android phones, the spacing was off and not conforming to specifications on the race screen (see Figure 54). Formatting issues, such as these can lead to inaccurate and incomplete data.

5.17.3 Feedback on instructions for selecting more than one race

We had noticed in earlier rounds of testing that not all participants read the race instruction to check one or more boxes for race. This could be because the issue did not apply to them or because they did not see the instruction. In the knowledge check questions that were administered for the first time in the 2020 usability testing sessions, one of the questions asked, “Do you think someone can select more than one race on the census questionnaire?” Out of 25 participants across the three rounds of 2020 usability testing, eighteen said yes which indicates either they already knew this information or they read the instructions on the race screen. Six participants marked that they did not know and one participant marked “no.” See Figure 55. This chart indicates that seven out of 25 participants did not attend to the instructions on the race screen.

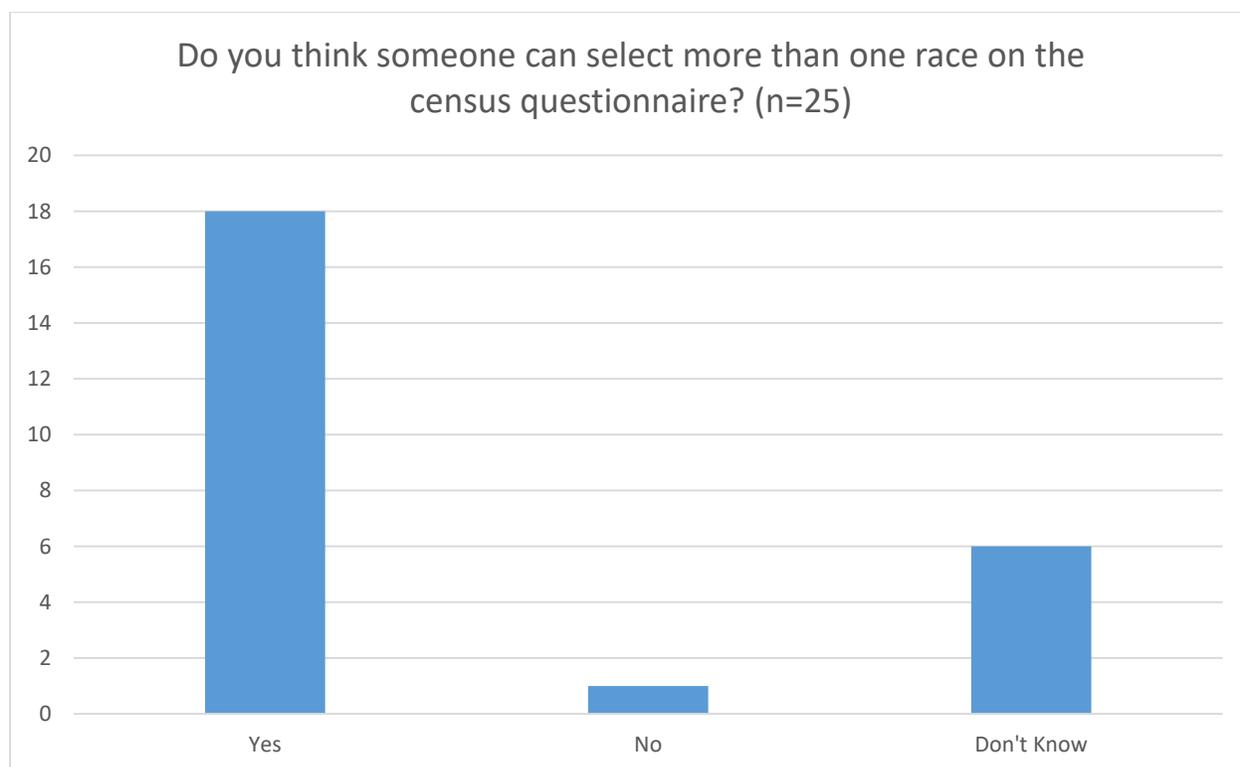


Figure 55. Chart of one question on the knowledge check asking if participants think they can select more than one race.

6 Summary and Discussion

Content organization of Hispanic origin and race: There have been a number of different designs on the Hispanic origin and race screens. For most of the decade, usability testing occurred on the combined question on race and ethnicity. The combined question appeared to work better for participants than when it was separated into two questions on two separate screens.

When it was separated into two questions (e.g., Hispanic origin & race), we heard comments either spontaneously or during debriefing that the participant did not know why the census separated Hispanics out from other races. People consider Hispanic origin to be a race and when it is separated on its own screen, there is no context to why the census appears only interested in Hispanics. The instruction on the race screen that “Hispanic origins are not races” adds additional confusion to Hispanic respondents who have interpreted that sentence to mean that “Hispanic” is not “equal” to the other races.

This is not only a problem in the self-administered mode. During interview debriefings with the telephone agent staff, we also heard that when the two questions were separated it caused comments and confusions for respondents that did not occur when the two questions were combined into one (Nichols & Olmsted-Hawala, 2018; Nichols, Olmsted-Hawala & Katz, 2016). Thus, for both modes, when it is possible to incorporate the two questions into one we suspect there will be less user confusion and more user satisfaction on that question. While the Census Bureau must follow the OMB standards for this question, we recommend further research into using the combined question design and presenting that research to OMB as the combined question was much more acceptable to participants and appeared to produce more accurate data than when the questions were separated. If the questions must remain as two separate questions, we think it would make sense to test having them appear on the same screen for the self-administered mode so that respondents may know that there are two related questions on the topic of Hispanic origin and race, which could reduce some of the confusion. See for an example of this in Figure 56. This will be something to consider testing in the future.

United States[®]
Census
Bureau

National Survey of Psychiatrists

Frequently Asked Questions Survey Information

Are you of Hispanic, Latino, or Spanish origin?

Yes

No

What is your race? Select all that apply.

White

Black or African American

Asian

American Indian or Alaska Native

Native Hawaiian or Other Pacific Islander

OMB No.: xxxx-xxxx
OMB Expiration Date: mm/dd/yyyy

Figure 56. Screenshot from another survey that asks Hispanic origin and race on the same screen.

Response Categories: The 2015 NCT introduced MENA as a new category. This was tested from the 2015 NCT through the 2017 CT. Per the decision memo of 2018, it was removed for 2020. See the Memorandum for more on why this category was not fielded (2018). In general, the MENA category in the combined race and ethnicity question tested well with participants. A number of people who identified as MENA selected that as their response. We did not hear comments from other people who do not identify as MENA, aside from one African American who said she could select either North African or Black or African American. (She did select African American in the end.) We should note that the primary testing for MENA occurred in 2015 prior to Executive Order 13769 which may have impacted people’s willingness to report if they were MENA. This seemed to be the case for some testing on the MENA category that was done by NCHS later in the decade after Executive Order 13769. (See Willson, Dunston, and Rios, 2018). We recommend more testing on this category.

Many participants, if not most, do not make a distinction between Hispanic origin and race. When the questions are separated, they look for a Hispanic category in the race response choices and when the response choices are together, they do not question why Hispanic is listed as a response choice.

As for missing response choices, one participant commented that Brazilians are difficult to categorize. For unexpected response choices, several participants questioned why “Chamorro” was listed. We recommend further testing.

Proxy reporting: Feedback came from participants regarding reporting the race of a household member that was unrelated to them. Some participants felt uncomfortable with this task and wanted a way for that person to report for him or herself. Participants suggested that they be allowed to save the information already entered and have the roommate log back in to finish their own section, the ability to leave it blank or the ability to indicate they did not know or were not sure about their roommate’s race. The ability to allow uncertainty for proxy reports is something that could be investigated in the future.

Overview of designs tested of the detailed origins screens: Over the past decade, we usability tested different ways to ask for detailed origins in the online self-response mode. In early and late rounds of testing, the open-text field asking for detailed origins appeared directly next to the race option or below all race response options, but all on the same screen. In some rounds of testing, the open-text field unfolded when the main race was selected; and sometimes it was more similar to the paper questionnaire and the open-text field was present on the screen at all times and could be completed even before the main race was selected. Mid-decade the detailed origins questions were asked most often on a separate screen. One design had a list of checkboxes and a final open-text field if none of the other checkboxes fit. The other design was an open-text field without any checkboxes. When the open-text field was on the same screen as the main race categories, we saw many participants initially leave the open-text field blank and attempt to move forward only to trigger an error message. We did not observe this behavior when the detailed origins open-text field was on a separate screen (s). And, when the detailed origins were on the same screen, we observed that participants spent more time on the screen when the open-text field was at the bottom of the screen, far from the associated race, than when it was closer to the race response choice.

Measurement error with the detailed origins: With all of the designs tested during the decade, participants seemed to feel compelled to enter a detailed origin, whether they felt closely associated with that detailed origin or not. We heard many participants remarking that they did not know how far back to go; did not feel particularly attached to the detailed origin, or frankly did not know the detailed origin. This was especially true when participants answered for unrelated people in their household.

Some participants, particularly White and African American participants that had been in the United States for many generations, made comments that indicated they did not identify strongly with anything other than what they had already marked (e.g., White or African American). Some of the African American participants mentioned how it was redundant have to type in “African American” since they had just checked a box that had those exact words. Some of the participants that selected “White” as their response option said that their ancestors came from a mix of many different countries and they did not know what exactly they were made up of, other than just plain American. Having these participants identify with detailed origins by triggering an error message when the field was left blank, or by using a separate screen for the item without a “don’t know” response choice may have led to some amount of error in the data.

Measurement error due to error messages: Although numerous types of error messages were tested in the effort to communicate to respondents that they did not have to enter a detail, none were effective in communicating that ability. (Participants did not leave the detailed origin fields blank, even when the error message text told them they could.) What *is* understood by the error message is that something must be typed into the field in order to move forward in the questionnaire. It does not seem to matter what the error message says, rather participants do not think they can move forward in the survey when they receive one.

In some of the tests we investigated how usability was impacted when the error messages appeared in different colors (e.g., green or red) and with different associated icons as well (an “x” or an exclamation point “!”). Feedback from users was mixed. Participants had different opinions about the colors (e.g., some preferred the green color; some said the red was more noticeable, etc.) Moreover, while all participants gave a comment on color that was asked during the 2014 CT debriefing, the few who triggered a race error message live during the session did not appear to give any specific meaning to the color. Consequently, our take away was that the color and error message wording really did not affect users’ behavior with the screen; it was simply that the participant stayed on the screen and a message appeared.

The error message is a good solution for the person who has detailed origins to enter and inadvertently left the field blank. But it is not a good solution for the person who intended to leave the field blank because he or she did not identify with any detailed origins. During testing, it was sometimes difficult to determine whether leaving the field blank was because they simply missed noticing the open-text field, or because they did not have any detailed origins to add. Results from the knowledge check indicate that more people than not assume they were required to provide an answer rather than leave an answer blank (see Figure 52).

Consequently, based on the way the error messages indicated to participants that they could not move on unless they answered the question, our team recommended the error messages should not

be triggered on the screen. This is a mode difference with paper and may inflate the detailed origins count with which people may actually not closely identify.

Additional usability issues with detailed origins screen designs In addition to leaving the open-text field blank, other usability issues came up with the detailed origins designs:

- **Checkbox design:** Usability issues with the checkbox design was that some participants marked more checkboxes than they perhaps closely identified with. If they had been required to type out the detailed origin (or write it in if on paper) it is possible that the same participants may not have written in/typed out as many of the boxes as they checked. When asked about this during debriefing some said this was accurate, they would only have written the detailed origin that they were “mostly” while others disagreed and said they would have written them all in. We noticed that in at least one case, a family member marked different checkboxes for her siblings even though she said they were biological siblings. So there may be some amount of error with the checkbox design, perhaps attributed with the ease in which one can check a box. With respect to preference, it occurred generally that if a person saw their detailed origin listed they were satisfied, and if not they were less satisfied, sometimes making a negative comment about why the checkboxes did not include what they were looking for.
- **Checkbox design missing an “other” checkbox:** The design of the open-text field on the checkbox design screen (e.g., the other field) was not standard and was mentioned by participants as missing a check box in front of the response choice. This issue occurred across multiple rounds of testing. This led to our recommendation of placing a checkbox in front of the open-text field to be parallel with the other response options. In addition, it may make sense to add the word “other.” So for example, on the Hispanic origin detailed information screen it would have the open-text field with a checkbox at the front and the words: “*Other, for example, Guatemalan, Spaniard, Ecuadorian, etc.*” We would recommend further testing of this design.
- **Unfolding design for detailed origins:** This was the design tested in the first round of 2020 Census testing. The box for detailed origins appeared under the race, once the race was selected. The issue was that participants were not expecting this box and there was a delay in it appearing on the screen. Participants almost always selected “next” instead of answering in the box; this would trigger the error message. This design was changed for the next round of testing so that the open-text fields was on the screen when the screen first loaded. We do not recommend the unfolding design for detailed origins.
- **Outlining the open-text field with red when not answered:** The red around the field did put focus on it; but that also led participants to believe that they had to enter an answer. We also observed that with the red outlining the box, one participant actually deleted races she had reported because she did not understand the error message. We do not recommend triggering error messages with the red box around the empty open-text field.
- **Separate detailed origin screen(s) for each race selected:** While the separate screen for details reduced the number of error messages, when participants marked more than one race they frequently entered all the detailed origin on the first screen. They did not anticipate that they were going to have the opportunity to enter detailed origins for each race on screens one after the other. This could be addressed by asking for all detailed origins for every race to be added on one screen (rather than separating it out into two or more screens). We continue to recommend this design.

Selecting more than one race and the instruction: Another of the issues we noted throughout the decade was whether participants understood that they could select more than one race category.

This was difficult to test because most of the participants in the studies identified with a single race. We did ask during debriefing whether people noticed the instruction and received feedback on the wording of the instruction at that point. The instruction currently reads, “*Select one or more boxes AND enter origins. For this census, Hispanic origins are not races.*” While we had some participants mention that the second sentence was confusing, (see description of issue with Hispanic origin not a race above) participants in the most recent round of testing did not comment on the first sentence in the instruction. In prior rounds of testing the instruction varied but included such wording as “*Select one or more boxes **And** enter origins.*” Or, “*Select one or more boxes and enter origins. Note you may report more than one group.*” Or, “*Select all boxes that apply. Note, you may report more than one group.*” We have yet to understand whether two instruction statements are necessary, and if they draw more or less attention to that area of the screen. Our suggestions for additional testing on this topic include a targeted approach of recruiting participants that identify with more than one race combined with eye-tracking analysis to determine whether having two similar sentences draws more attention to the area or whether it has the opposite effect. This is something that could be tested in the future. Still, from what we could tell during the decade of testing, the layout and design seem to have more of an impact on what is selected or entered compared to the instruction.

7 Acknowledgements

The authors thank members of the Internet Self Response team, and our reviewers Marisa Hotchkiss, Rachel Marks, and Joanne Pascale.

8 References

- Ashenfelter, K., Olmsted-Hawala, E., Lakhe, S., and Malakhoff, L. (2013). Final Report of Usability Testing of the 2013 National Census Test (NCT) Online Survey: Focus on Coverage, Race Questions, Date Format. Center for Survey Measurement Study Series (Survey Methodology #2013-01). U.S. Census Bureau. Available online at: <https://www.census.gov/srd/papers/pdf/ssm2013-01.pdf>
- Compton, E., Bentley, M., Ennis, S., & Rastogi, S. (2012.) 2010 Census Race and Hispanic Origin Alternative Questionnaire Experiment. DSSD CPEX Memorandum Series #B-05-R2. Available online at: https://www.census.gov/2010census/pdf/2010_Census_Race_HO_AQE.pdf
- Ericsson, K. A., & Simon, H. A. (1996). Protocol Analysis: Using Verbal Reports as Data. (Revised ed.) Cambridge, MA: The MIT Press.
- Horwitz, R., Nichols, E., & Coombs, J. (2018). The Impact of Removing Instructions on Web Surveys. 2020 Census Program Internal Memorandum Series 2018.36.i, U.S. Census Bureau. October 22, 2018. Available online at: <https://www.census.gov/srd/papers/pdf/ssm2013-01.pdf>
- Mathews, K., Phelan, J., Jones, N., Konya, S., Marks, R., Pratt, B., Coombs, J., Bentley, M. (2017). 2015 National Content Test Race and Ethnicity Analysis Report: A new design for the 21st century. Available online at:

<https://www2.census.gov/programs-surveys/decennial/2020/program-management/final-analysis-reports/2015nct-race-ethnicity-analysis.pdf>

Memorandum 2018.02. (2018). Using Two Separate Questions for Race and Ethnicity in 2018 End-to-End Census Test and 2020 Census. Available online at: https://www.census.gov/programs-surveys/decennial-census/2020-census/planning-management/memo-series/2020-memo-2018_02.html

Nichols, E. M., Olmsted-Hawala, E.L. (2018). Evaluating the Decennial Call-in Option: Results from an Interviewer Debriefing of the 2016 Census Test. U.S. Census Bureau. (Survey Methodology #2018-04). Available online at:

<https://www.census.gov/content/dam/Census/library/working-papers/2018/adrm/rsm2018-04.pdf>

Nichols, E., Olmsted-Hawala, E. & Katz, J. (2016). Evaluating the Decennial Census Call-In Option: Results from an Interviewer Debriefing of the 2015 National Content Test. *Center for Survey Measurement Research Report Series (Survey Methodology #2016-02)*. U.S. Census Bureau. Available online at: <http://www.census.gov/srd/papers/pdf/rsm2016-02.pdf>

Office of Management and Budget. (1997). “Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity.” Available online at: <https://www.whitehouse.gov/wp-content/uploads/2017/11/Revisions-to-the-Standards-for-the-Classification-of-Federal-Data-on-Race-and-Ethnicity-October30-1997.pdf>

Willson, S., Dunston, S., and Rios, M. (2018). Results of a Cognitive Interview Evaluation of the Revised Race Question, with Special Emphasis on the Newly Proposed Middle Eastern/North African Response Option.” Presented at the FCSM Research and Policy Conference, March 8, 2018. Available online at: <http://www.copafs.org/UserFiles/file/FCSM2018FinalProgram.pdf>

U.S. Census Bureau. (2012). “Census Bureau Releases Results From the 2010 Census Race and Hispanic Origin Alternative Questionnaire Research.” Census Bureau Newsroom Archive. Available online at: https://www.census.gov/newsroom/releases/archives/2010_census/cb12-146.html

Appendix A. General Device and Equipment used in Usability Tests 2012-2020

	Devices used in testing	Recording	Browser	Eye tracking
2012 NCT	Census Provided Desktop PC with Windows 7 operating system	Camtasia	IE Version 10	TOBII X120 or T120
2014 CT	Census Provided Desktop PC with Windows 7 operating system	Camtasia	IE Version 10	TOBII X120 or T120
2015 CT R1	BYOD smartphone / tablet; Census provided 13.5 inch Laptop & Desktop PC Windows 7	Camtasia	Phone - users' default browser; Computer: IE Version 11	TOBII T120 or X2-60
2015 CT R2	BYOD smartphone / tablet	Camtasia	Phone - users' default browser	NA
2015 NCT	BYOD smartphones / tablet & Census provided 13.5 inch Laptop PC Windows 7	Camtasia	IE Version 11	TOBII ET X2-60 (mobile/tablet)
2016 CT	BYOD smartphone / tablet & Census provided 13.5 inch Laptop Dell Precision M4800 or Dell Latitude E6540 PC Windows 7	Camtasia	IE Version 11	SMI Red 250 (Laptop) TOBII ET X2-60 (mobile/tablet)
2017 CT	Census Bureau provided Dell Precision M4800 13.5 inch Laptop PC Windows 7	Camtasia	Firefox	NA
2020 R1	Census Bureau provided iPhone 8 or 13.5 inch Dell Latitude E6430 PC Windows 7	Camtasia & QuickTime	iPhone = Safari; Laptop = Firefox	TOBII X2-60
2020 R2	Census Bureau provided iPhone 8 or 13.5 inch Dell Latitude E6430 PC Windows 7	Camtasia & QuickTime	iPhone = Safari; Laptop = Firefox	TOBII X2-60
2020 R3	Census Bureau provided iPhone 8 or Android (Moto Z3)	Camtasia & QuickTime	iPhone = Safari; Android = Chrome	NA

Appendix B. Details of Smartphone, Tablet and Personal Computers used during testing 2012 – 2020

Device	2012	2014	2015	2015	2015	2016	2017	2020	2020	2020
	NCT	CT	CT R1	CT R2	NC T	CT	CT	R1	R2	R3
Smartphone										
iPhone 3			1	1						
iPhone 4				1	1					
iPhone 5 S and C			4	2						
iPhone 6					2	2				
iPhone 8								5	3	3
Android (included Samsung 4, Samsung Galaxy Note 3 and Note 4, Android Galaxy S3, Prism 2, LG, LG2, ZTE Zmax, Metro PC, LG P659, Kyocera)			3	8	4					
Nokia Windows Smartphone			1							
Android Moto Z3										2
Tablet										
iPad 2 (10 inch)			2		1	1				
iPad 2 mini				1	2	0				
Nook (7 inch)			1							
Samsung Galaxy Tablet 10 inch				1	1	1				
Samsung Galaxy Note 4 with stylus					1					
Android Polaroid 7"					1					
Microsoft Windows surface pro				1	1					
Personal Computer										

Census provided Laptop PC		1	1		9	7	10	5	7
Census Bureau Desktop PC	20	10	2						

*20 contractor participants used smartphones: (iPhone 5s, 6, Samsung Galaxy 3, 4, S5, Samsung Mega-Galaxy) and Tablets (Samsung Tab 2, Tab 10, iPad 2, Toshiba tablet) but exact number breakdowns unavailable.