Optimizing the Decennial Census for Mobile – A Case Study

Elizabeth Nichols, Erica Olmsted Hawala, Rachel Horwitz, and Michael Bentley[1]

U.S. Census Bureau
4600 Silver Hill Road, Washington DC 20233

Abstract

The U.S. Census Bureau is committed to offering an Internet response option for the 2020 Census. We expect the majority of self-responses to come in through this medium. Decennial census field tests, such as the 2012 National Census Test and the 2014 Census Test, have used an online instrument to collect data in preparation for the 2020 Census. However, testing conducted through 2014 used online instruments that were designed for optimal view on a desktop or laptop. Although these surveys could be answered on tablets or smartphones, the design was not optimized for these smaller devices. On some mobile devices the screen display was very small and required the user to zoom or make other manipulations to enable the user to clearly read and answer the questions.

Mobile-ownership statistics show that, as of 2014, over half of adults owned a smartphone and some adults were dependent upon their smartphone for Internet access. These smartphone-dependent adults were more likely to be lower income, younger, and minority (Pew Internet Project, 2014). With the growth of mobile device ownership overall and the differences in device-dependent Internet access across subpopulations, the Census Bureau realized it must offer a responsive design for the online Census. That is, the questions and response categories must render optimally on the device, whether it is a large desktop computer or a small smartphone. A responsive design was developed for the 2015 Census Test. This means there was an optimized design for smaller devices such as smartphones and small tablets, and for larger devices such as large tablets, laptops, and desktops. Usability testing was conducted on different devices prior to fielding the survey.

This paper discusses the style rules we used to develop the mobile-optimized version of the 2015 Census Test instrument, and the issues that arose during usability testing. Additionally, we present device usage, completion time, and break-off data from the non-optimized 2014 and optimized 2015 Census Test online instruments. The two tests occurred in different geographic regions of the country with different population characteristics. So the typical usability metrics of time-on-task and task completion presented here to evaluate the effect of optimization are limited by sample confounds.

Introduction

The growth of Internet access in U.S. households and the widespread use of the Internet to accomplish tasks formerly done on paper have prompted the Census Bureau to offer the public the ability to complete their 2020 Census forms online. The Census Bureau expects the online form to be the primary data collection mode for this next population census.

Changing data collection modes for the U.S. population census is not new. Prior to the 1970 Census, data were collected in person by interviewers. Since 1970, the Census Bureau has collected census information by mailing a paper questionnaire to households and asking them to complete it and mail it back. An interviewer visited only those housing units that did not respond to this request. The current plans for the 2020 Census include mailing a letter to housing units asking the occupants to complete their census form online[2]. A paper form will later be sent to units that have not responded online or by telephone, and field interviewers will continue to play a part in collecting data for nonresponding households

With each data collection mode change, the survey questions and their display are optimized for the new mode. The underlying principle guiding any change is that the meaning and intent of the questions and response options are

[1] This report is released to inform interested parties of research and to encourage discussion of work in progress. Any views expressed on the methodological issues are those of the authors and not necessarily those of the U.S. Census Bureau.
[2] Approximately 20 percent of the country will also receive a paper questionnaire in the first mailing. This mailing will be targeted to areas with lower Internet access and usage rates.

consistent across modes and that the same respondent would give the same substantive answer to a question regardless of the mode of administration (Martin et al., 2007).

Many different types of devices, such as desktops, laptops, tablets, or smartphones, are used to access the Internet. Ownership of the smaller devices has grown over the past five years as shown in Figure 1. For the 2020 Census, the question and response option text and display will have to accommodate various devices, which differ by such things as the size of the screen, the keyboard configuration, and the input device (e.g., computer mouse, trackball, stylus, or finger) so that respondents would give the same answer regardless of device.

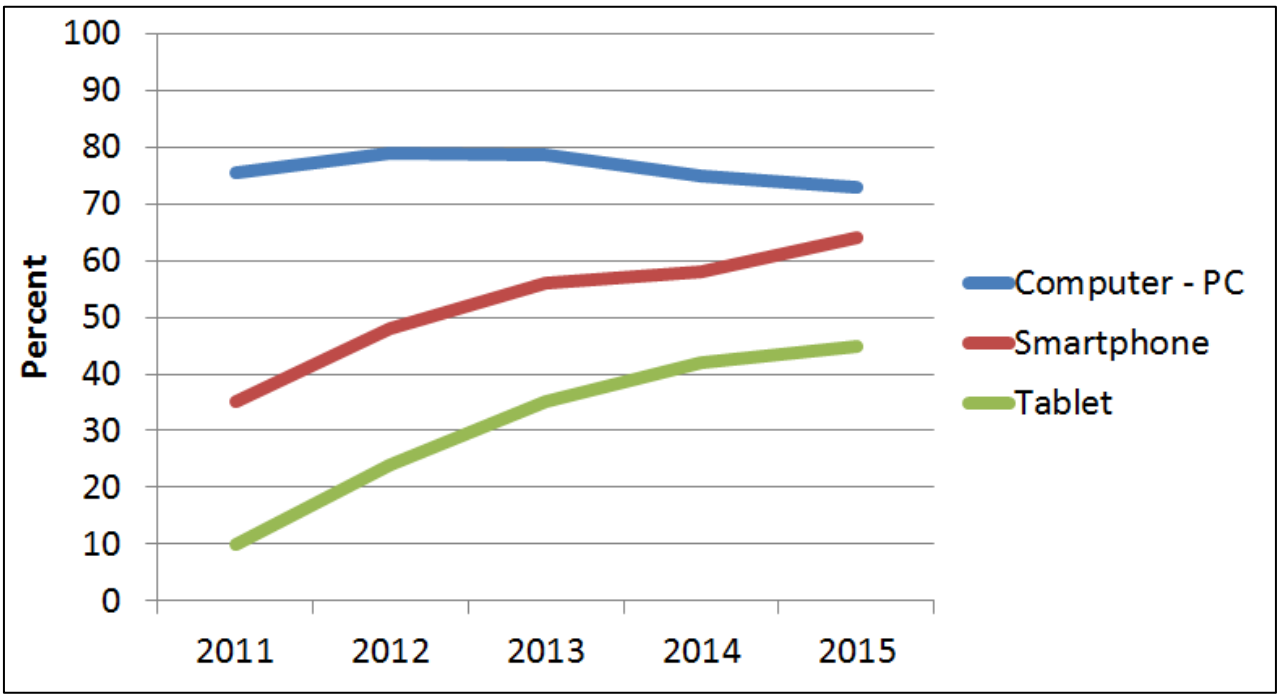

Figure 1: Device ownership growth
Source: U.S. Census Bureau Internet and Computer Usage Supplement (File, 2013) and Pew Research Center (Smith, 2013; Pew Internet project, 2014; Anderson, 2015)

Three mid-decade census tests have offered an online questionnaire. In the first two tests, the online questionnaire displayed optimally only on larger devices, such as desktops, laptops, and larger tablets. In the third test, a responsive design was used for the online questionnaire, meaning that it displayed optimally on smaller devices, such as smartphones and small tablets, as well as on larger ones. A style guide was created to guide the design decisions for the smaller device display (referred to as mobile-optimized display). Usability testing was then conducted prior to finalizing the mobile-optimized design decisions. Metrics, including task completion time and break-off rates by device type, were gathered during each of the three mid-decade census tests.

This paper focuses on the design decisions that were made when creating a mobile-optimized census questionnaire and the resulting effects of those decisions.

Brief history of three mid-decade U.S. census tests

In planning for the 2020 Census, minimizing costs was a definite priority, and a series of readiness tests conducted throughout the decade addressed this goal. Tests conducted in 2012, 2014 and 2015 each had slightly different objectives related to online reporting and increasing self-response as ways to contain costs. These tests experimented with the type and number of contacts with the household, the availability of the paper census form, and the design of the online questionnaire.

The 2012 National Census Test (NCT) was a nationwide sample of housing units drawn from the list of addresses which would have received a paper census form in the mail. The 2012 NCT had a Census Day of September 5, 2012 and the primary objective of the test was to obtain a baseline response rate for those who would report online. The control panel included four mailings: an advance letter; the initial mailing package with a letter and Internet

instruction card; a reminder postcard; and the final mailing package containing a paper questionnaire. Respondents accessed the browser-based Internet form using an identification number sent in the initial mailing package. The data collected in the online questionnaire included the names of everyone living or staying at the address and their demographic characteristics (sex, age and date of birth, relationship to householder, Hispanic origin, and race). The response rate for the control was 60.3 percent with a 38.1 percent Internet response rate.

The look-and-feel of the 2012 NCT online census form was designed to best display on a desktop or laptop rather than a smaller mobile device as smartphone penetration had not yet reached 50 percent (File, 2013). Respondents could complete the 2012 NCT on a smaller mobile device, but the display shrank to fit on the smaller screen and the respondent had to manipulate the screen by zooming and panning to see and answer the questions.

Across all the experimental panels, respondents used computers to access the online questionnaire over 90 percent of the time. Only 2.4 percent of the questionnaires were accessed by a smartphone and only 4.7 percent were accessed by a tablet. Completion time and break-offs were higher, on average, for questionnaires completed on smartphones (U.S. Census Bureau, 2014).

Given the overall success of the online response option in the 2012 NCT, the basic design of that online census form was re-used in the 2014 Census Test (see Figure 2). Meanwhile, mobile-device ownership continued to grow as shown in Figure 1 above. In 2013, the planning year for the 2014 Census Test, smartphone ownership grew to 56 percent of adults 18 years old or older (Smith, 2013). By 2014, 64 percent of adults owned a smartphone and seven percent of adults depended upon their smartphone for Internet access. These smartphone-dependent adults were more likely to be lower income, younger, and minority (Pew, 2014). Creating a responsive design, where the look-and-feel of the survey was optimized based on the dimensions of the device accessing the URL, was discussed for the 2014 Census Test, but had to be put on hold due to competing budgetary priorities.

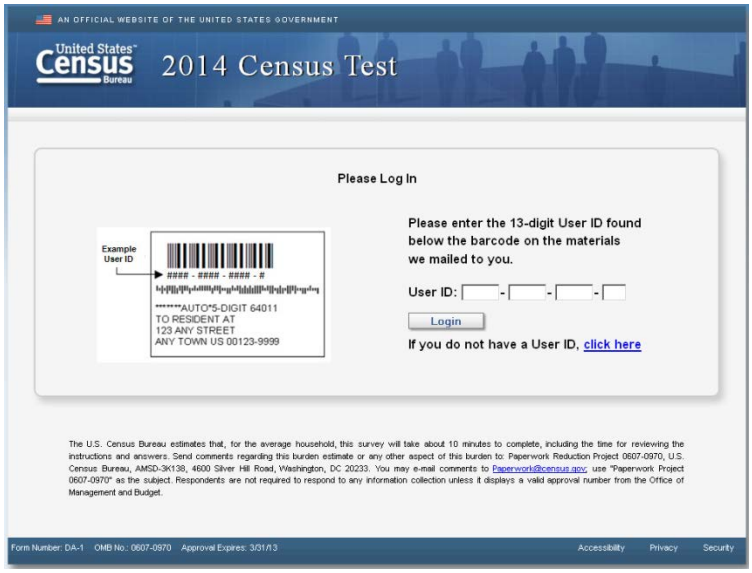

Figure 2: Login screen for 2014 Census Test, accessed via a desktop

Census Day for the 2014 Census Test was June 1, 2014 and the test was conducted in areas of Washington DC and a neighboring county. The control panel for this test first used postal mail to notify the sample of the online option and then sent the paper form to nonresponding households. This panel used an identification number for entry to the online census questionnaire. With an overall self-response rate of 65 percent and 50 percent online response rate, this panel was even more successful at moving responses online than in 2012.

Across all the experimental panels, respondents with an identification number used a desktop or laptop to access the online questionnaire 87.2 percent of the time, with 10.2 percent accessing it from a tablet and 2.6 percent accessing it from a smartphone. Thus, even though there was a growth in smartphone ownership, we did not observe an uptick in smartphone usage to access the census test compared to 2012. The increase in tablet use to access the census test online form could be attributed to the high educational level of the people living in DC and the nearby county as tablet ownership is correlated with higher education (Horwitz, 2015; Anderson, 2015).

The 2015 Census Test was conducted in two mid-range media markets: the Savannah, Georgia area and Maricopa County, Arizona. Census Day was April 1, 2015. A responsive design for the online form was developed for this test. Users who accessed the survey from larger devices were offered the existing look-and-feel for desktops, laptops, and larger tablets used in 2012 and 2014 (see Figure 3) but screens changed design when displayed on a smaller-device (see Figure 4). Those users saw a mobile-optimized version of the survey. In Savannah, the overall response rate for the control panel was 47.5 percent and 33.4 percent online response rate. In Maricopa, the overall response rate for the control panel was 54.9 percent and 39.7 percent online response rate.

Across both geographic areas and all panels, respondents with an identification number used a desktop or laptop to access the online questionnaire 80.5 percent of the time, with 16.2 percent accessing it from a tablet and 6.4 percent accessing it on a smartphone. We are unable to untangle whether the increased use of smaller devices to access the survey was because of increased ownership of the devices, demographic and socio-economic differences between the locations, or the fact that the survey was mobile optimized.

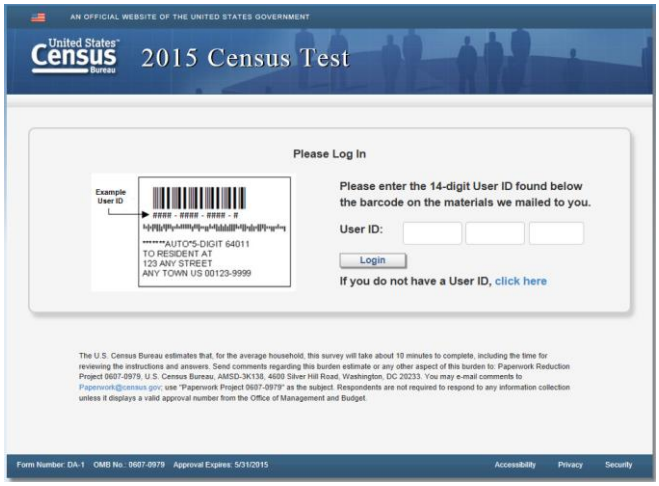

Figure 3: Login screen for the 2015 Census Test optimized for desktop, laptop and larger tablets

Figure 4: Login screen for the 2015 Census Test optimized for a smartphone or smaller tablets

The remaining sections of this paper discuss the limitations of comparisons across the two census tests, a description of the desktop and mobile questionnaire design, usability testing highlights, and metrics by device type for the 2014 and 2015 Census Tests.

Limitations comparing across census tests

As mentioned in the previous section, the 2014 Census Test and the 2015 Census Test occurred in different geographic areas and at different times. The location difference confounds any comparisons between the two tests. The population in those areas differs by demographics and Internet usage. The average household size in Washington, DC (the 2014 Census Test area) was approximately 2.1 and 88.4 percent of the adults 18 years old or older had at least a high school degree (Community Facts, American Fact Finder, 2013). In the 2015 Census Test areas, the average household size is a little larger (2.4 for Savannah and 2.7 for Maricopa County) and the percent of high school educated adults is a little lower (85.6 percent and 86.4 percent, respectively) than the 2014 location (Community Facts, American Fact Finder, 2013). We would expect the online questionnaire to take a little longer to complete on average in the 2015 Census Test because the household sizes were larger and the respondent has to answer questions about each person in the household. Additionally, educational level of the respondent, which differs between areas, has been shown to affect survey completion time, with people with lower levels of education taking longer due to a variety of factors (Gummer and Rossmann, 2015; Horwitz, 2015; Couper and Kreuter,2013; Yan and Tourangeau, 2008). The field period differences confound device usage rates as device ownership changed over time.

Desktop online questionnaire design for the 2014 and 2015 Census Tests

As shown in Figures 2 and 3 above and in Figures 5 and 6 below, the basic design of the desktop online version of the census test instrument included a banner, some instructional links in the upper and lower right corners, one question per screen, and Previous and Next buttons left-justified below the last response option. Question text was bolded while the response options were not. Item-specific help was offered as a link (Help) on all pages. Fills (such as the name of the resident embedded in the question) and edit flags (i.e., when a participant failed to enter information, a notice appeared on the screen) were included (see Figure 5).

Both the 2014 and 2015 instruments used the same basic design for the questions. For questions where the respondent had to make a selection from pre-specified categories, the response option layout was vertical. The radio button or checkbox was to the left of the response option. The race question (Figure 5) and the relationship question (Figure 6) had the largest number of response categories, but under standard screen size and display settings, no vertical or horizontal scrolling was required.

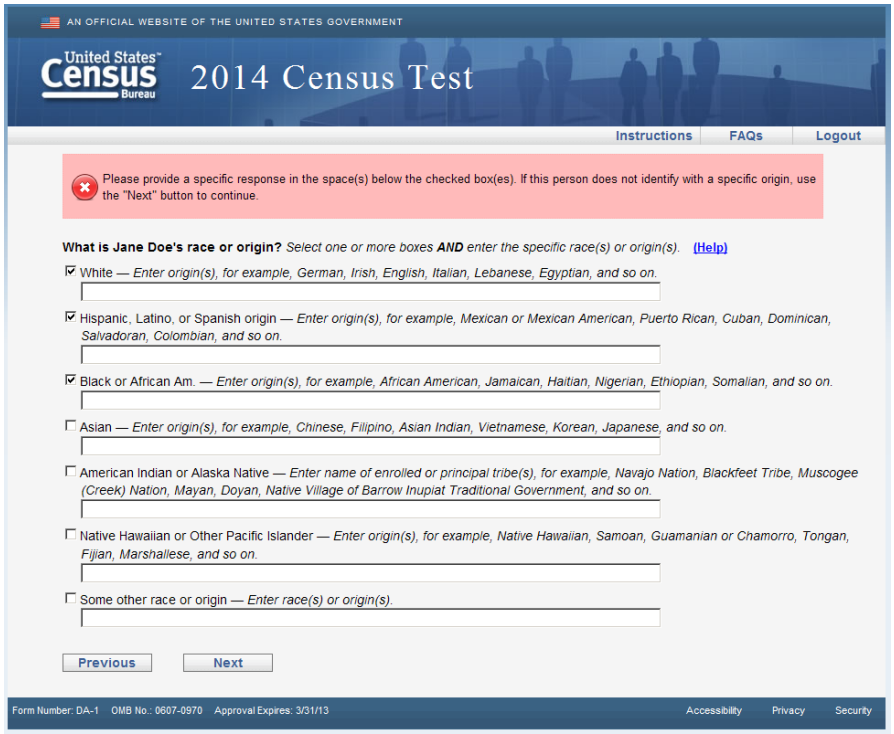

Figure 5: A combined race and Hispanic origin question from the 2014 Census Test with an edit flag optimized for desktop, laptop and larger tablets

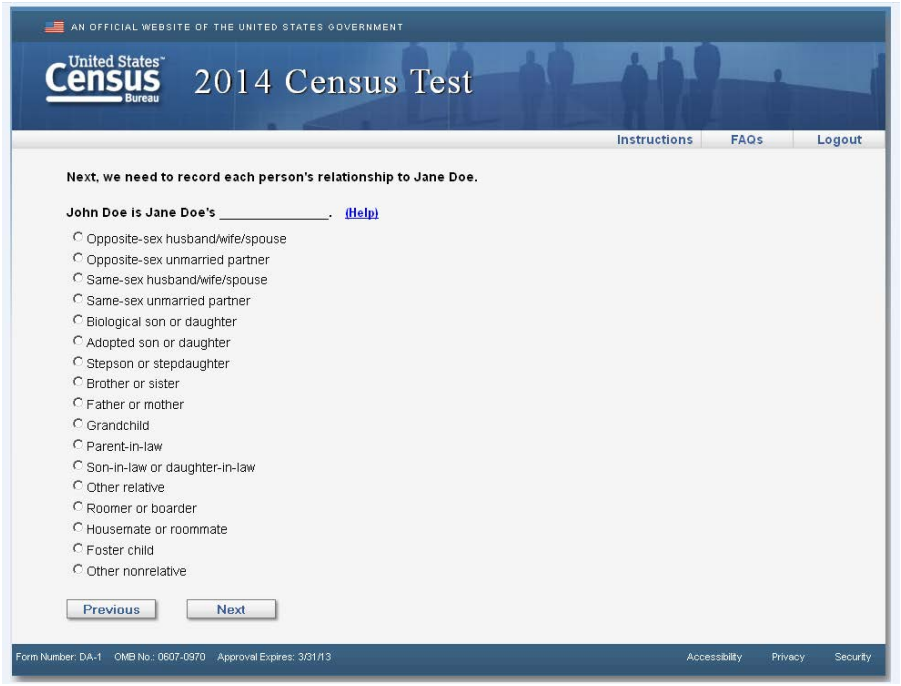

Figure 6: The relationship question in the 2014 Census Test optimized for desktop, laptop and larger tablets

The response fields for questions where the respondent entered information rather than chose an answer, such as name (Figure 7), date of birth (Figure 8), and address (Figure 9) were arranged using a layout typical for writing or displaying those data items. For example, first name, middle name, and last name fields were displayed horizontally on one line, as were month, day and year of birth fields. Address fields included street number, street name, apartment number fields on one line with city, state and ZIP code fields on the line below it.

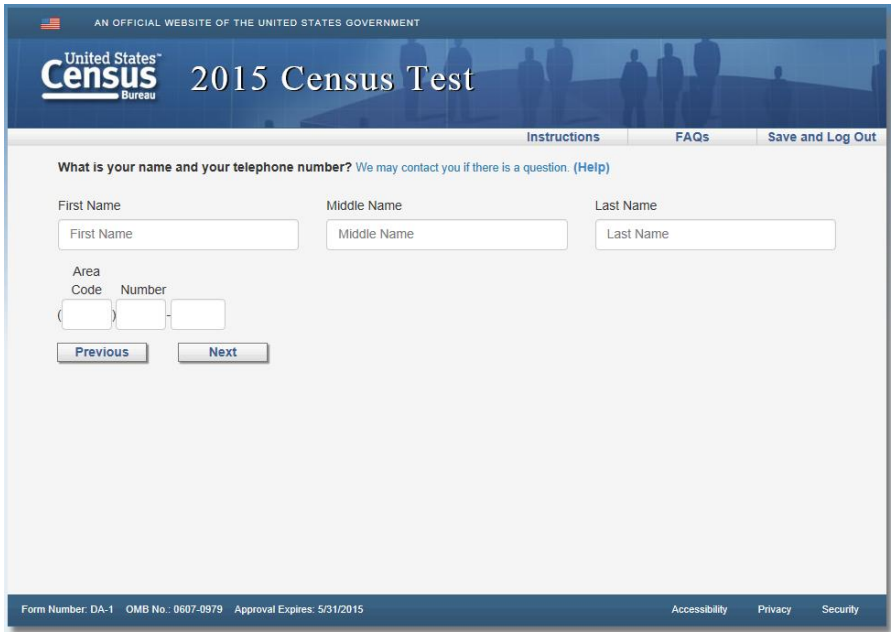

Figure 7: Respondent name screen in 2015 Census Test optimized for desktop, laptop and larger tablets

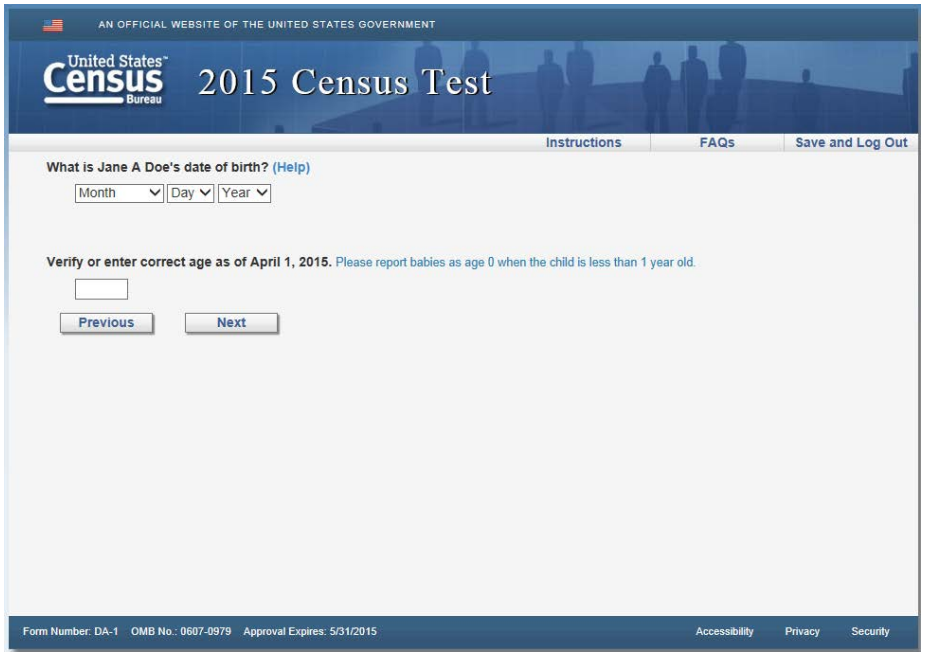

Figure 8: Date of birth and age questions in the 2015 Census Test optimized for desktop, laptop and larger tablets

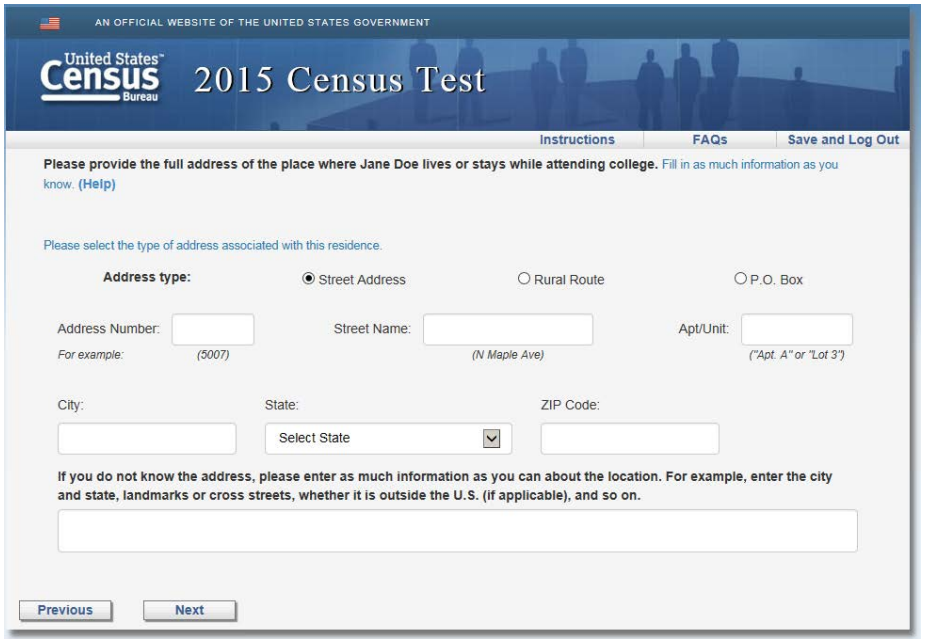

Figure 9: An address collection screen in the 2015 Census Test optimized for desktop, laptop and larger tablets

Navigation buttons guided the respondent through the online form in a linear manner. There was no opportunity to skip sections; however, the respondent could logout and resume later with the correct authentication. In both tests, the online forms first collected the roster of people living at the address. Then, the demographics were asked using a topic-based design—they were grouped together so that, for example, the form collected the age and date of birth for everyone before going to the next topic.

The data collected and online features were very similar between tests. There were minor differences between the 2014 and 2015 instruments, including the addition in 2015 of a review screen that allowed respondents to edit the list of people living at the address and a streamlined address collection for people living at other residences (Nichols et al., 2015). However, the major difference between the tests was that the 2015 Census Test offered a responsive instrument, which would display the questions optimally when accessed on smaller devices or larger devices.

Mobile online questionnaire design for the 2015 Census Test

In preparation for the conversion to a responsive design, staff at the Census Bureau created a style guide for the display on smaller devices. Topics covered by the guide were influenced by the usability testing experience in late 2014 and early 2015 of the un-optimized American Community Survey (ACS) on smartphones (Olmsted Hawala et al., forthcoming) and by a review of commercial mobile optimized websites (including Target and Heathcare.gov) and the literature on best practices when optimizing for mobile (Baker, 2014; Geisen and Olmsted, 2011; Jue and Luck, 2014; Link et al., 2014; Peytchev and Hill, 2010; Rempel and Bridges, 2013). This internal style guide was provided to the people who programmed the mobile device software.

The style guide addressed display configurations, the touch gestures allowed, the text display, edit messaging, help screens, the banner, the menu, and navigation. Display configurations included making the questions readable and response categories selectable on the mobile device without having to zoom; maximizing use of the real estate on the smaller device and eliminating horizontal scrolling. The following examples highlight the guidelines provided.

Eliminate need for zooming

Although we did not test the un-optimized 2014 Census Test instrument on a smaller device, we had tested a very similar instrument, the ACS, and thus use that instrument to illustrate the need for zooming that a 2014 Census Test respondent would have experienced on a smartphone. In Figure 10 the ACS login screen was unreadable when it first displayed on the smartphone; the 2014 Census Test instrument would have looked similarly. The participants had to use their fingers to zoom and pan around the screen until they found the desired place to enter their identification number fields (see Figure 11). Figure 11 also shows what typically happened when participants zoomed in -- they could no longer see the question, instructions, or any other useful content. The 2015 Census Test login page was designed to eliminate the need for zooming as shown in Figure 4 above. For screens with response options such as radio buttons or checkboxes (Figures 5 and 6), more space was needed between the response options so that the respondent did not misclick or have to zoom first to select the desired response with his or her finger.

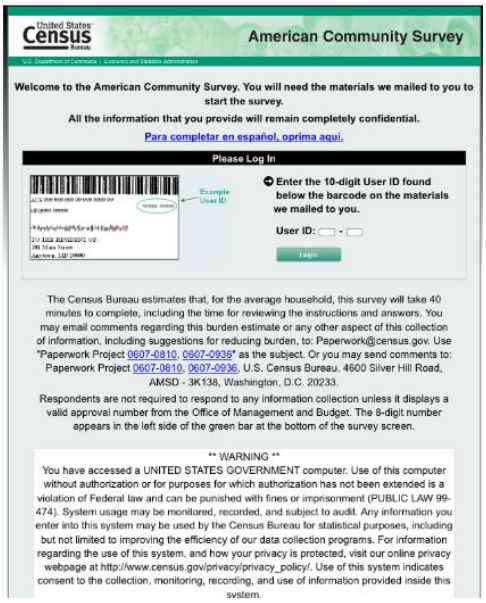

Figure 10: Login screen for the ACS – un-optimized online design accessed via a smartphone.

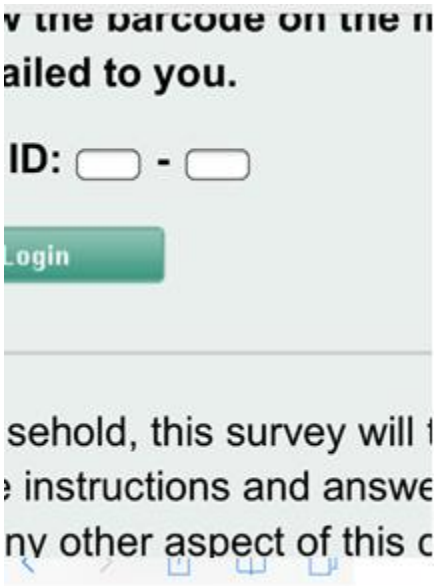

Figure 11: Identification fields in Figure 10 magnified by zooming

Use the entire screen and eliminate white space

Figure 12 is an example of how a typical data collection screen appeared in the un-optimized version. The question takes up less than half of the screen real estate, with the rest of the screen filled with empty white space. Figure 13 shows how that screen was changed so that it "fit" on the full smartphone screen and allowed the respondent to read the text without zooming.

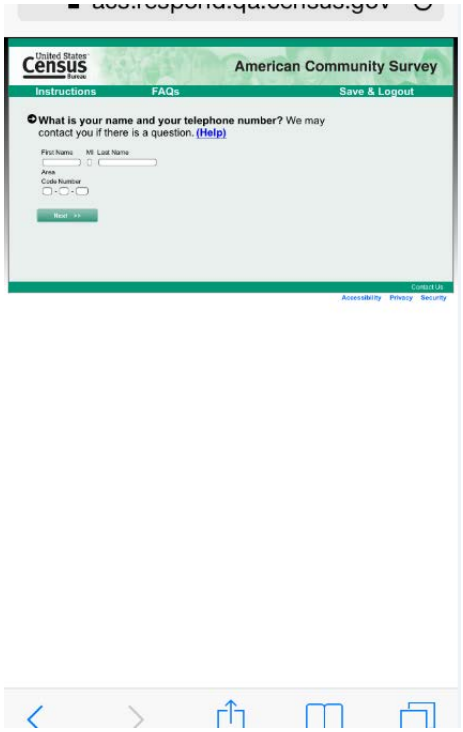

Figure 12: ACS name and telephone number question- un-optimized online design accessed via a smartphone

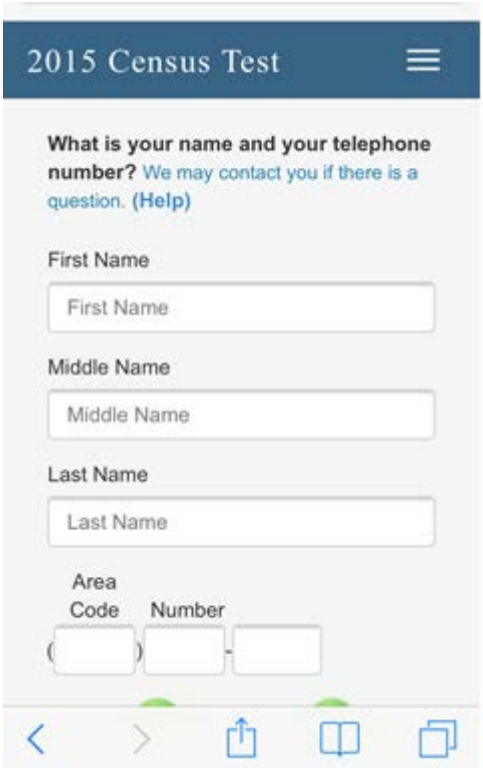

Figure 13: Respondent name screen in Round 1 of the 2015 Census Test optimized for a smartphone or smaller tablets

Allow vertical scrolling for long lists

For questions such as relationship and race, the style guide suggested allowing vertical scrolling as long as the respondent could read the questions, scroll to the correct response, enter the correct response and navigate to the next page without having to pinch to zoom or horizontally scroll as shown in Figures 14 and 15 of the optimized version of the 2015 Census Test survey.

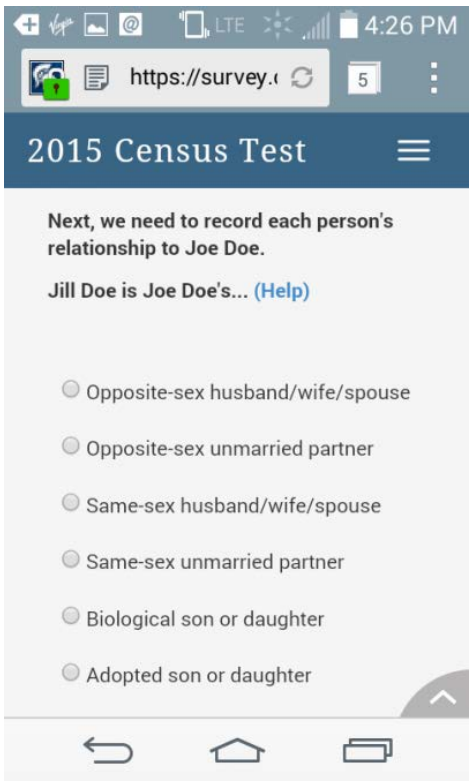

Figure 14: Relationship question in the 2015 Census Test optimized for a smartphone or smaller tablets.

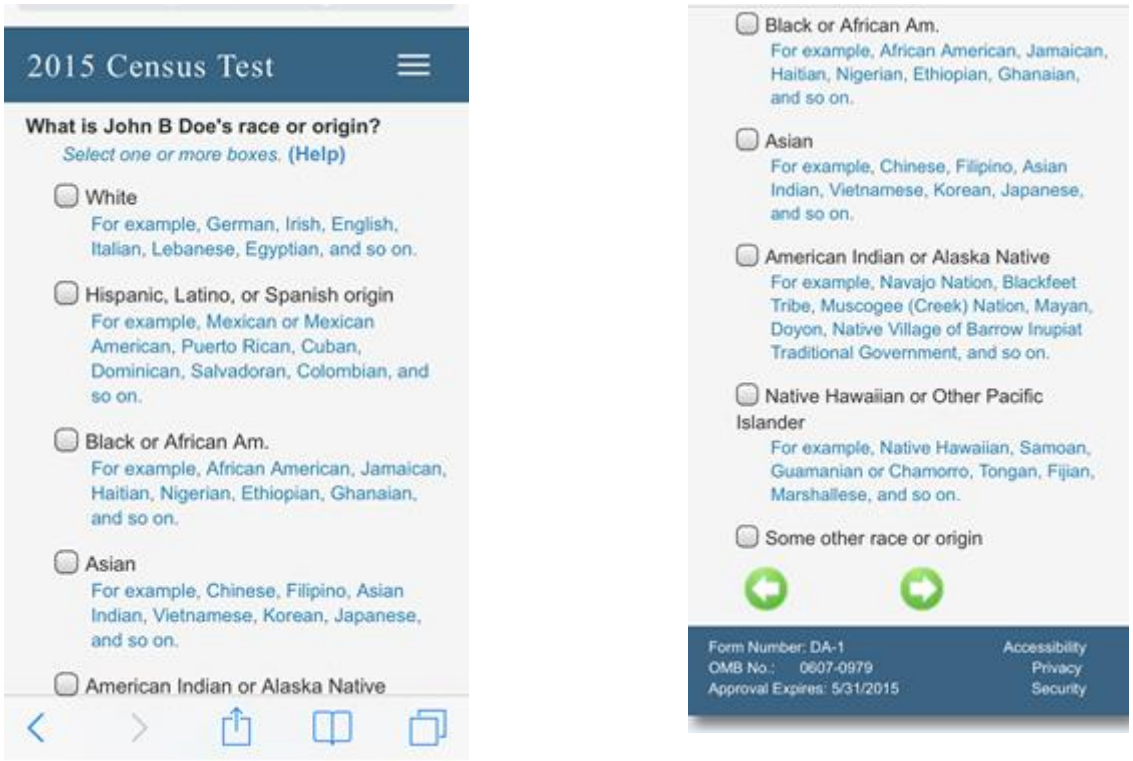

Figure 15: Combined race and Hispanic origin question in the 2015 Census Test optimized for a smartphone or smaller tablets.

Eliminate horizontal scrolling by redesigning horizontal response fields to vertical

The screens requiring major reconfigurations, given the style guide suggestions, were the login screen (Figures 2 and 3) and the other screens that collected data horizontally, such as name (Figure 7), date of birth (Figure 8) and address (Figure 9). For these screens, we suggested a vertical design for the data entry, with field labels above the data entry field, and additional information, such as the example for the login, below the response entry field. The revised screens for login (Figure 4), name (Figure 13) and address (Figure 16) show the screens after implementing the style guide suggestions (but prior to usability testing).

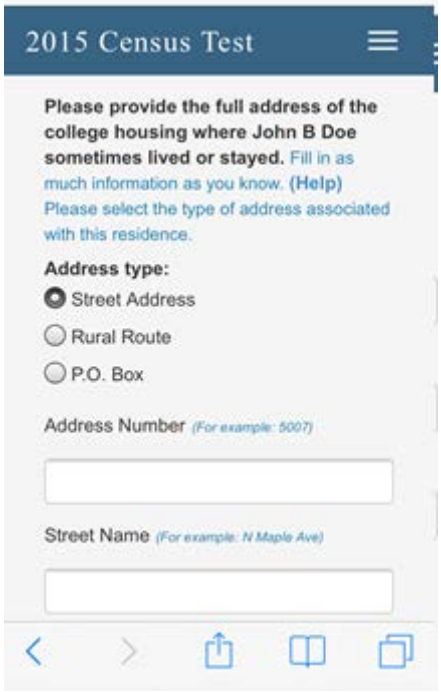

Figure 16: An address collection screen in the 2015 Census Test optimized for a smartphone or smaller tablets

Maximize use of the smaller screen real estate

The style guide suggested minimizing the banner and using "a hamburger menu" for FAQs and logout features (Figures 17 and 18). To trigger the edits and to navigate forward to the next page or backward to the previous page, the style guide suggested keeping "buttons" at the bottom of the page, but reducing the size of those buttons (Figures 19 and 20). The buttons for navigation also forced respondents to scroll down through all the response choices before moving onto the next page.

Figure 17: Banner and menu items on a desktop

Figure 18: Banner and "hamburger" menu on a smartphone

Figure 19: Navigational buttons on a desktop

Figure 20: Navigational buttons on a smartphone

Eliminate large blocks of text

In the 2014 Census Test, the race examples were italicized and on the same line as the response choice, which created a big block of text on smaller devices because of wrapping (Figure 21). The style guide discouraged big blocks of text that would display when examples wrapped. In addition, the use of italics was discouraged because in the smaller screen it was more difficult to read. To accommodate these suggestions, examples were moved to a new line, and were not italicized as shown in the design used for the Race screen in Figure 15 above.

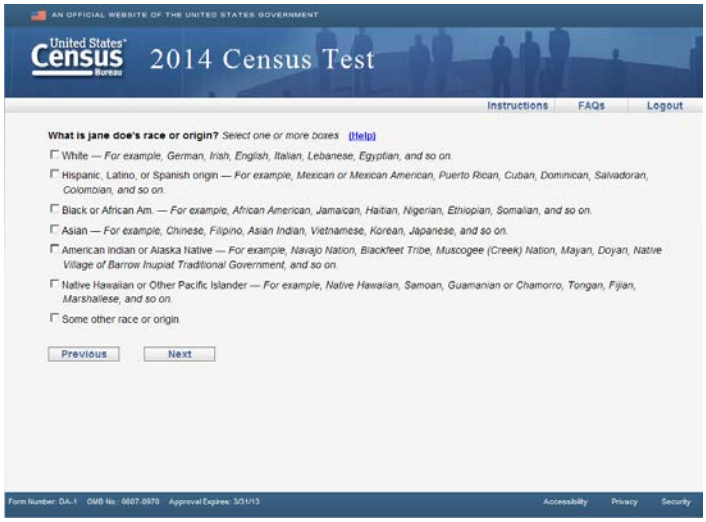

Figure 21: A race screen in the 2014 Census Test on a desktop

Usability testing highlights

To improve the census online instruments, usability evaluations were conducted prior to fielding both the 2014 and 2015 Census Tests using functional versions of the online surveys (Nichols et al., 2015; Nichols et al., forthcoming). A similar protocol was used in each evaluation. Participants from the general public were recruited to be a part of a one-and-a-half hour research study. In the studies, each participant completed the survey while thinking aloud. Each session was recorded and observed by a researcher. Debriefing questions were asked after the respondent completed the online form. The researchers took notes on issues where the respondent had difficulty completing the question, answered inaccurately, or made oral comments concerning confusion or respondent burden.

For the 2015 Census Test usability evaluation which focused on the mobile optimization, 30 English-speaking participants took part. Two rounds of testing were conducted, and changes were made between rounds based on findings from the previous round. All but three participants used their own mobile device to complete the online form. Six participants used a tablet and 20 participants used a smartphone. A variety of smartphones and tablets were used, including Apple products, e.g., iPhone versions 3 through 5, iPads (mini and 10 inch), and Android products including Samsung, Galaxy Note, Prism, LG phones, a Nook, a Galaxy Tablet and a Microsoft Surface Pro tablet.

Zooming and Vertical Scrolling

During the first round of testing on mobile devices, we found that the spacing between the response choices on a smartphone was not large enough for respondents to make accurate selections. "Fat finger" is a phrase used to describe the situation when the user's finger touches the wrong choice by mistake. For example, the relationship question radio buttons were too close together in Round 1, causing users to initially select the response option above or below the one they wanted (Figure 22). Some users would zoom to make the correct selection while others would simply tap again. The issue was resolved once more space was added for Round 2 (Figure 23). The respondent had to scroll in both designs to see all the choices and use the navigation arrows. However, the revised page required even more scrolling as the response choices were separated more. The additional scrolling on the relationship screen was not a problem for participants. Scrolling was not a problem on the race or address screens, which were the other two long screens in the instrument.

Because the responsive design detected screen size and displayed a version appropriate for the size, small tablets (e.g., an iPad mini, vertically held) showed the mobile-optimized version while larger tablets (e.g., a 10-inch tablet) displayed a desktop version. We observed some *fat finger* issues with the tablets showing the desktop design. The style guide did not address increases in the space between the response choices for the desktop display, and no changes were made to the spacing before the 2015 Census Test.

Figure 22: Relationship question in Round 1

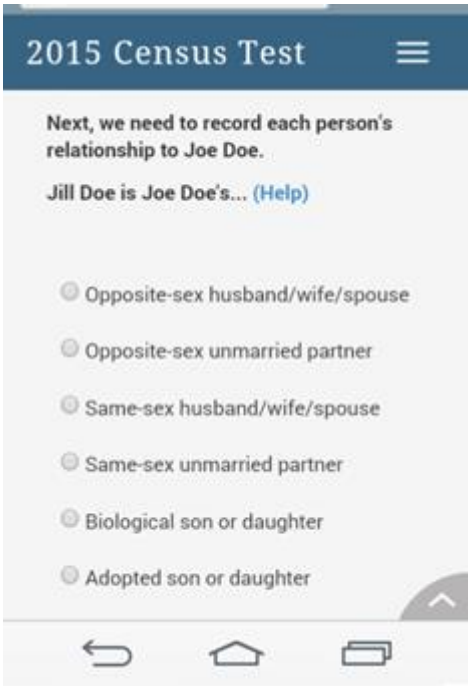

Figure 23: Relationship question in Round 2

Eliminating white-space, eliminating large blocks of text, and maximizing use of the smaller screen real estate

All white space was eliminated as shown in the mobile-optimized figures and the big block of text for the race examples was eliminated (Figure 15). Users had no comments on the layout of the pages and selected race response options without problems using the design. Most users were familiar with the hamburger menu design and all users used the navigation arrows to move between pages.

Redesigning horizontal response fields to vertical

For the name and address questions, we recommended a vertical design rather than a horizontal design in the style guide. Because there was so much space between the label and the field on the address screen we tested, it was ambiguous which label went with which field (see Figure 24). Some participants assumed the labels were associated with the field above the label and entered the information into the incorrect field. This usability issue was easily corrected by moving the label closer to the field below it prior to the instrument going live in the 2015 Census Test (see Figure 25). On the name screen, the labels were closer to the field in the initial testing (see Figures 28 and 29 below) and we did not observe the same usability problem on those screens.

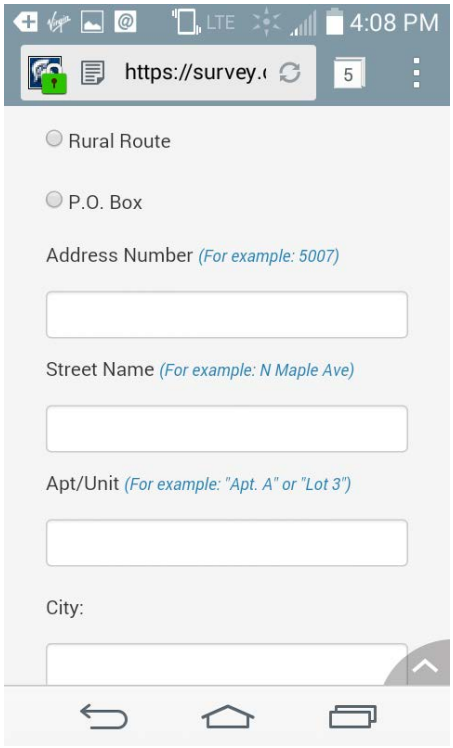

Figure 24: Address fields in the 2015 Census Test Round 1 testing instrument

Figure 25: Revised address fields used in the 2015 Census Test

Another usability issue occurred on the address screen (Figure 24). The design required participants to separate their address number and street name into two fields, which is not a typical address design. Many participants typed both pieces of information into the first field since that field was long enough to accommodate all this information. The field size indicates to respondents the desired length of their answer (Seckler et al., 2014). The desktop design for that field (see Figure 9) was shorter, but during usability testing, a longer field was inadvertently used with the design when displayed on smaller devices. Figure 25 shows the corrected field length for address number.

No additional issues were uncovered for the name or address questions where we switched the desktop's horizontal screen orientation to a vertical orientation for smaller devices.

Spacing between response choices was an issue with the vertically designed date-of-birth response options in Round 1. Some respondents mistakenly selected day when they wanted to select month (Figure 26). The horizontally designed response options in Round 2 (Figure 27) eliminated this usability issue. Thus, not all response choices had to be changed to a vertical display to be readable.

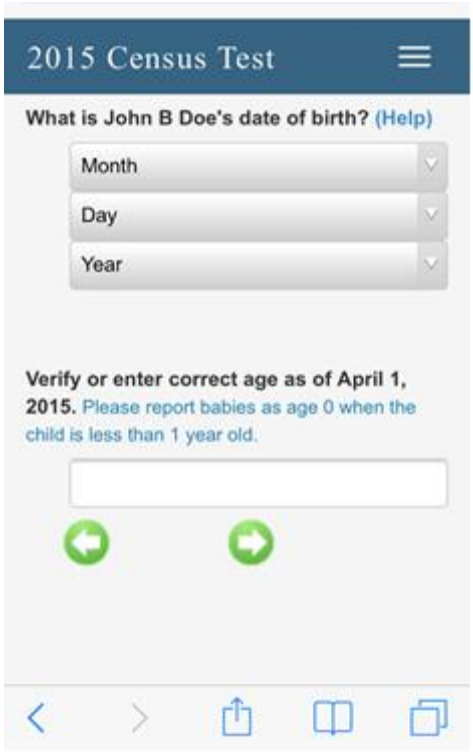

Figure 26: Round 1 date of birth response option design

Figure 27: Round 2 date of birth response option design

Symbolism

We observed a usability issue with the relationship question during the 2015 testing that we had not observed in earlier desktop testing, and therefore was not in the style guide. Several participants clicked on the underline in the statement "James C Doe is Jane A Doe's _____" (Figure 22). During debriefing, these participants indicated that they thought the underline was a drop down list. We changed the design to use ellipses in both the smartphone and desktop design (Figure 23). Not only did this design change eliminate the usability problem, it also made more sense when read aloud by the screen reader JAWS (Job Access with Speech).

Labeling

Labels, placement of labels, and field length affect the usability of the web forms on all devices (Seckler et al., 2014), but were not mentioned specifically in the style guide. A few examples of problems with position of labels and field length occurred when moving from a horizontal design to a vertical design. We noticed one other problem with the labels with the optimized design when displayed on smaller devices. On the name screen, participants in Round 1 had difficulty with the telephone number label (Figure 28). Several respondents tried to type their ZIP code instead of their telephone number. It seems they were interpreting "Code Number" as the field label instead of "Area Code Number." The label was changed to "Telephone Number" in Round 2 (Figure 29) which seemed to clear up the confusion as no one tried to type in their ZIP code. This label change was also made to the computer design.

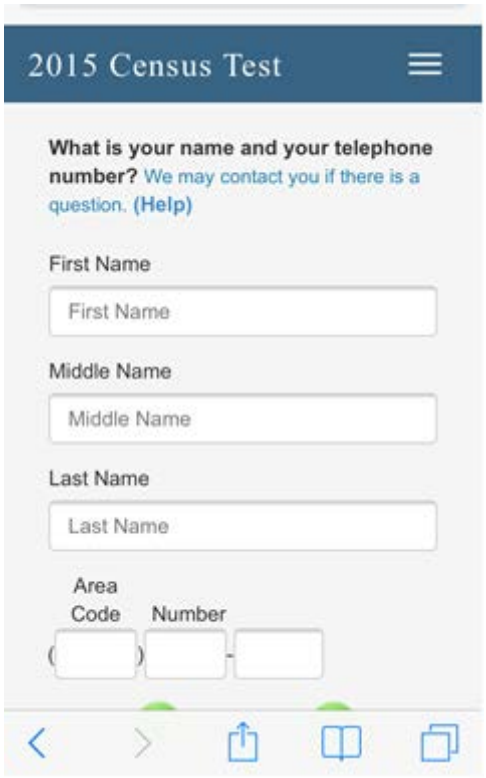

Figure 28: Respondent name screen in Round 1 optimized for a smartphone or smaller tablets

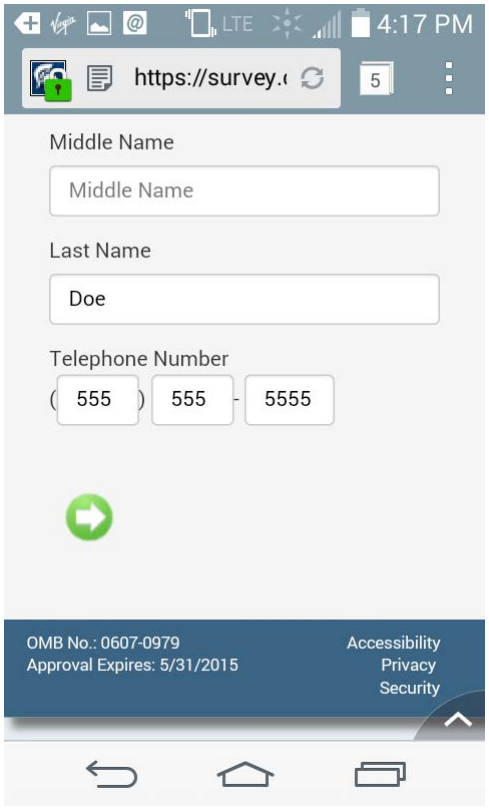

Figure 29: Respondent name screen in Round 2 optimized for a smartphone or smaller tablets

Other issues not identified in the style guide

System errors are always uncovered in any usability evaluation. Because our participants owned and brought in so many different types of devices, we were able to observe display issues specific to some smartphone browsers and hardware. These issues made the form display sub-optimally on some devices. These were not necessarily usability problems we could resolve without further testing on those devices and it is not clear that the subsequent programming resolved issues with all devices. For the most common device producers, including Apple and Samsung, the display was correct.

Another notable usability issue we observed dealt with the speed of the system when loading pages, especially when participants chose to use a public Wi-Fi instead of a data plan. The system was so slow that often users selected the navigation buttons over and over because they were unfamiliar with the spinning wheel that indicated a page was loading.

Finally, participants had difficulty with the drop-down design on Apple devices. For questions with drop-down pick lists (e.g., the date-of-birth and the census security verification question), Apple devices displayed the response choices below the screen as a spin wheel or "picker" (Figure 30 shows the security questions pick list). Some participants had never seen this design and searched a long time before seeing the response choices. We did not observe the same problem with Android products as the pick lists display more similar to a laptop/desktop design with a child window (Figure 31).

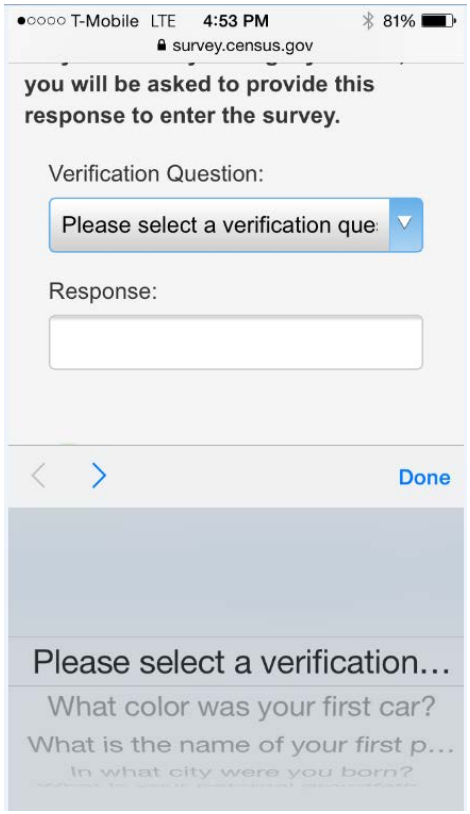

Figure 30: Pick list display on an iPhone

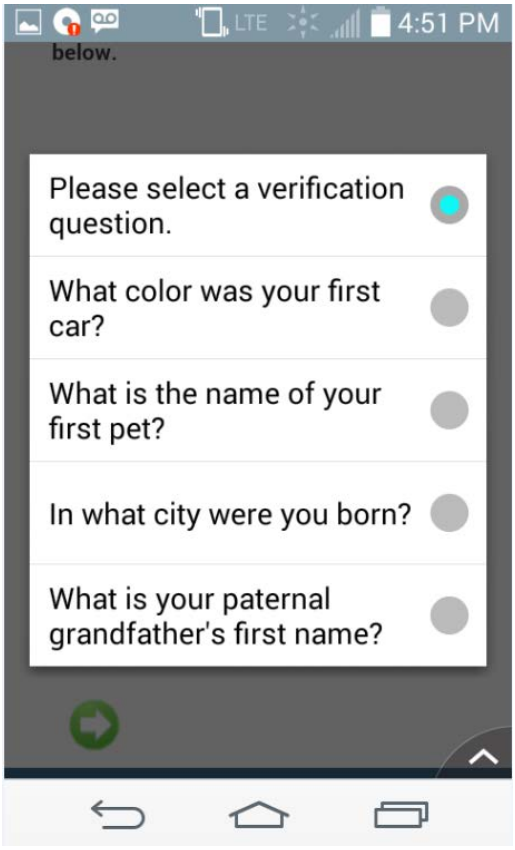

Figure 31: Pick list display on an Android phone

Usability testing results confirmed many of the style guide suggestions made; identified correctable spacing and labeling issues, and highlighted design limitations such as connection speed and unfamiliarity of device displays when completing the census online questionnaire on a mobile device.

Usability performance metrics in the 2014 Census Test and 2015 Census Test

Efficiency (as measured by completion time) has long been an important metric of measuring the usability of a product. The two other typical measures are whether the user accomplishes his or her goals effectively (i.e., accurately) and if the user is satisfied with the experience (ISO standard 9421-11). During usability testing of the 2014 and 2015 Census Test instruments, we did not measure completion time due to the protocol, but in the actual tests, completion time was measured.

In a meta-analysis, Couper and Peterson (2015) show that completing a survey on a smaller device takes anywhere from 1.1 times to 3.1 times as long as it does on a PC when the survey is not optimized for the smaller device. They also discovered that even when the survey is optimized, it still takes longer on a smartphone in the majority of case studies. Not only is actual completion time longer when taking a survey on a smartphone, but perceived time is longer as well (de Bruijne and Wijnant, 2013). We found similar results in both the 2014 and 2015 Census Tests.

The median completion times presented in Table 1 show that it took longer to complete the 2015 Census Test online form than the 2014 Census Test online form for each type of device.[3] The minor changes to the questions between the two tests would not explain why it took longer in one of the tests compared to the other. Differences in the test locations could help explain these numbers. As mentioned earlier, the larger household sizes and lower educational level of respondents expected in the 2015 Census Test could increase the time needed to complete the online survey regardless of device.

[3] We used median instead of mean completion time to lessen the effect of outlier cases which had very long completion times perhaps because they left the survey "running" while accomplishing other tasks.

Table 1: Median completion time in minutes for the 2014 Census Test and the 2015 Census Test online instruments by device type

	2014 Census Test (minutes)	2015 Census Test (minutes)
Computers	7.3	8.4
Tablets	8.8	10.2
Smartphones	10.3	11.1

Source: 2014 and 2015 Census Tests: all completed or partial cases with an identification number to authenticate to the survey

Even with the differences in the locations, these results confirm what other researchers have found with smartphone completion time (Couper and Peterson, 2015). The median completion time is higher for smartphones than for computers or tablets whether the design is optimized for smaller mobile devices (2015 Census Test) or not (2014 Census Test). The ratio of median completion time of smartphones to computers is also consistent with prior research. It took respondents 1.4 times as long to complete the device on a smartphone as it did on a computer with an un-optimized form in 2014 and 1.3 times as long when optimized in 2015. We did not conduct significance testing due to the confounding factors between the locations.

Couper, Antoun, and Mavetova (2015) speculate that scrolling accounts for a lot of the increased time needed to complete a survey on a smartphone compared to a computer, with Internet connection speed responsible for a smaller portion. Our usability testing experience suggests Internet connection speed could have been a major reason for increased completion time on smartphones. However, based on observation alone, we are unable to rule out scrolling as a cause of the time increase. Other reasons could include multi-tasking or taking longer to read the text on a smartphone because the text takes up more of the screen real estate when optimized. Based on our usability experience, we wondered if the unique way drop-downs were displayed on iPhones could have attributed to the increased interview length. However, this explanation was not the case as the median completion time for iPhones in the 2015 Census Test was 10.18 minutes compared to 11.82 minutes for other smartphones.

In both tests, completing the survey on a tablet took approximately 1.2 times as long to complete as on computers. We were unable to detect what size tablet respondents used and whether the tablet was held at a vertical or horizontal orientation. Thus, we do not know which display (the desktop-look or the smartphone-look) the respondent saw.

The second usability performance metric available was the percent of break-offs, which is related to accuracy and satisfaction. Table 2 contains the break-off rate by test by device type. A break-off occurs when a respondent starts the online questionnaire but leaves without getting to the last screen and does not return to complete it. Dissatisfaction with the user experience could be one reason for a break-off, as could other reasons such as difficulty with the survey questions or outside distractions.

Table 2: Percentage of break-offs for the 2014 Census Test and the 2015 Census Test online instruments by device type

	2014 Census Test	2015 Census Test
Computers	4.1%	3.1%
Tablets	3.4%	3.6%
Smartphones	6.6%	5.4%

Source: 2014 and 2015: all completed or partial cases with an identification number to authenticate to the survey

Tablets had a very stable break-off rate across tests. This could be due to the population who uses tablets to complete surveys. The break-off rates are less stable across tests for computer- and smartphone-users. They are lower in 2015 than in 2014. Respondents who used smartphones were 1.7 times as likely as those using computers to break-off in 2015 when the survey was optimized and 1.6 times as likely to break off in 2014 when the form was not optimized.

Discussion

Between the 2014 and 2015 Census Tests, smartphone respondents increased from 2.6 to 6.4 percent and tablet users increased from 10 to 16 percent of all respondents. Optimizing the online census form for smartphones is one possible reason why the percent of respondents who answered via a smartphone or tablet increased between the two census

tests. Perhaps in 2014 more people initially accessed the instrument on their smartphone or tablet than actually completed it using that device. These people could have been turned off immediately by the un-optimized design on the login page. We speculate that they either did not respond at all or opted to answer on a larger device while those in 2015 who accessed the online form continued answering on their smaller device because the form was optimized. The increase in mobile device usage to answer the 2015 Census Test also could be due to population differences in the areas. In both the 2015 Census Test areas, the education level was lower than in the 2014 Census Test area. Lower education has been tied to an increase in smartphone device usage in other Census Bureau surveys (Horwitz, 2015).

Although in general the optimized version worked well during usability testing, we did observe that our style guide suggestions were not inclusive enough, as we found usability issues relating to response field spacing, labeling, and lengths. These issues were all corrected prior to the production fielding of the instrument. We also observed usability issues that we did not have control over, such as the hardware or browser used, Internet connection speed, and unfamiliarity with Apple device drop-down display.

The performance metrics of completion time and break-offs were not sufficient to measure improvements in usability between an un-optimized mobile design and the mobile-optimized design. Like prior studies, we observed no relative improvement in response time between a smartphone and a larger device when the design was optimized for the smartphone device compared to when it was not optimized. Break-off ratios also remained relatively constant across tests. Both of these could be due to factors other than usability – such as population differences in the test locations, Internet access speed differences, multi-tasking, etc. When conducting the analysis, we did not have access to these data and therefore could not control for these factors in the comparison across tests.

More work is needed to uncover the reasons for the break-offs and lack of response time improvement, whether they differ by the demographics of the respondents, the specific device type, the connection speed, or as yet unknown reasons. Future work should also look at what other metrics may work for measuring the usability of mobile devices. As the current efficiency metric indicates, it is not sufficient to identify a working design. Other efficiency considerations could include changing answers on the screen, or backwards navigation.

References

Anderson, M. "Technology Device Ownership: 2015." Pew Research Center, October, 2015, Available at: http://www.pewinternet.org/2015/10/29/technology-device-ownership-2015

Baker, R., (2014). Some thoughts on the 2015 Census for mobile. Internal correspondence.

Community Facts, American Fact Finder, 2013, URL: http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml, last accessed December 18, 2015.

Couper, M. P., & Kreuter, F. (2013). Using paradata to explore item level response times in surveys. Journal of the Royal Statistical Society, 176, 271–286.

Couper, M. P., and Peterson, G. (2015). Exploring why mobile Web surveys take longer. Presentation at the GOR 2015 Conference. http://www.gor.de/category/gor-15

Couper, M. P., Antoun, C. and Mavetova, A (2015). Mobile Web Surveys: A Total Survey Error Perspective. Presentation given at the 2015 International Total Survey Error Conference. Baltimore, Maryland.

de Bruijne, M. and A. Wijnant. 2013. Comparing survey results obtained via mobile devices and computers: an experiment with a mobile web survey on a heterogeneous group of mobile devices versus a computer-assisted web survey. Social Science Computer Review 31(4): 482–504.

File, Thom. 2013. "Computer and Internet Use in the United States." Current Population Survery Reports, P20-568. U.S. Census Bureau,Washington, DC.

Geisen, E. and Olmsted, M. (2011). 2020 Census Coverage Study: Survey on Mobile Devices Literature Review. RTI International. RTI Project number 0212349.

Gummer, T., & Rossmann, J. (2015) Explaining Interview Duration in Web Surveys: A Multilevel Approach. *Social Science Computer Review 33*(2) 217-234

Horwitz, R., (2015). Usability of the American Community Survey Internet Instrument on Mobile Devices. U.S. Census Bureau. #ACS15-RER-04 Last accessed: https://www.census.gov/content/dam/Census/library/working-papers/2015/acs/2015_Horwitz_01.pdf

Jue, A., and Luck, K. (2014). Update: Participation of mobile users in online surveys. Decipher White Paper. Retrieved from https://www.decipherinc.com/n/uploads/images/pages/Decipher_Mobile_Research_White_Paper_Update.pdf.

ISO 9241-11. International Organization for Standardization. *https://www.iso.org/obp/ui/#iso:std:21922:en* Retrieved January 20, 2016.

Link, M., Murphy, J., Schoer, M. F., Buskirk, T. D., Childs, J. H., and Tesfaye, C. L. (2014) Social Media and Public Opinion Research: Report of the AAPOR Task Force on Emerging Technologies in Public Opinion Research. https://www.aapor.org/AAPORKentico/Education-Resources/Reports.aspx

Martin, E., Childs, J. H., DeMaio, T., Hill, J., Reiser, C., Gerber, E., Styles, K., and Dillman, D. (2007). *Guidelines for Designing Questionnaires for Administration in Different Modes* U.S. Census Bureau, Washington, DC 20233.

Nichols, E., Olmsted Hawala, E., and Keegan, R. (2015). Usability Testing of the 2014 Census Test Online English Instrument. U.S. Census Bureau. *Survey Methodology #2015-04* http://www.census.gov/srd/papers/pdf/ssm2015-04.pdf

Nichols, E. and Olmsted Hawala, E. (forthcoming). Usability Testing of the 2015 Census Test Online English Instrument. U.S. Census Bureau.

Olmsted-Hawala, E., Nichols, E., and Holland, T. (2015). "A Usability Evaluation of the American Community Survey Online Instrument for the 2014 American Community Survey (ACS) Internet Test" U.S. Census Bureau. Survey Methodology

Smith, A. (2013). Smartphone Ownership 2013 - Update. Pew Research Center's Internet & American Life Project. http://pewinternet.org/Reports/2013/Smartphone-Ownership-2013.aspx

Pew Internet project (2014). Mobile Technology Fact Sheet as of October 2014. http://www.pewinternet.org/fact-sheets/mobile-technology-fact-sheet/

Peytchev, A. and Hill, C. A. (2010). Experiments in Mobile Web Survey Design: Similarities to Other Modes and Unique Considerations. *Social Science Computer Review. 28*(3) 319-335. DOI: 10.1177/0894439309353037.

Rempel, H. G. and Bridges, L., (2013). That was then, this is now: Replacing the mobile-optimized site with responsive design. *Information Technology and Libraries 32*(4), 8-24.

Seckler, M. Heinz, S., Bargas-Avila, J. A., Opwis, K., and Tuch, A. N. (2014). Designing Usable Web Forms – Empirical Evaluation of Web Form Improvement Guidelines DOI: 10.1145/2556288.2557265 Conference: In Proceedings of the 2014 annual conference on Human factors in computing systems (CHI'14). ACM, New York, NY, USA.

U.S. Census Bureau. (2013). Community Facts, American Fact Finder [Data file]. Retrieved from http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml

U.S. Census Bureau, (2014). "2020 Research & Testing Program Research Results Report: 2012 National Census Test Contact Strategy Results; Optimizing Self Response (4.103)" DSSD 2020 Decennial Census R&T Memorandum Series #E-04

Yan, T., and Tourangeau, R. (2008). Fast times and easy questions: The effects of age, experience and question complexity on Web survey response times. *Applied Cognitive Psychology*, 22, 51–68.