

RESEARCH REPORT SERIES
(*Survey Methodology* #2018-07)

**Redesigning Grids in the Mobile-Optimized
National Survey of College Graduates:
Results of a Usability Evaluation**

Elizabeth M. Nichols
Brian Falcone
Ivonne Figueroa

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

Report issued: April 2, 2018

Disclaimer: 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.

Abstract

In 2016, staff in the Center for Survey Measurement of the U.S. Census Bureau usability tested the 2017 National Survey of College Graduates (NSCG) online survey. During the usability sessions, respondents accessed the survey either on Census provided laptop computers or on their own mobile devices, including tablets and smartphones. This report documents the observations concerning the use of a grid response format with small mobile devices. A grid response format is a matrix where there is one question stem; each row heading contains a unique question topic or item, and the column headings contain a response option or value. Several questions on the NSCG paper form used grids and for mode consistency purposes, grids were used for those same questions for the NSCG online reporting option. During usability testing, we did not observe any difficulty with respondents using grids while answering on a PC. However, when respondents accessed the survey via a mobile phone, the grids were more difficult to use because respondents had to scroll back and forth horizontally to see the question items and all the response options. Because of this difficulty, halfway through the usability testing, the matrix design was changed to an item-by-item design for smaller mobile devices. In this design, each question that was originally presented in the grid became a stand-alone item with the response options placed vertically underneath the item as a series of radio buttons. While this design increased the amount of vertical scrolling needed on the page to see and answer each question item, the usability of the question improved with the elimination of the horizontal scrolling. Other researchers have also documented the advantages of an item-by-item solution for grids on mobile (de Bruijne 2015; Borger and Funke 2015; McClain & Crawford, 2013; Lattery & Park Bartolone, 2013). This report provides another source for documenting the item-by-item solution for grids on mobile as it relates to usability and user experience. This report does not make any claims about the impact of the design on the resulting survey response distributions as the research was qualitative in nature.

Keywords: grid response options, item-by-item design, National Survey of College Graduates, mobile optimized, survey design

Suggested Citation: Elizabeth M. Nichols, Brian Falcone, Ivonne Figueroa (2018). **Redesigning Grids in the Mobile-Optimized National Survey of College Graduates: Results of a Usability Evaluation** *Research and Methodology Directorate, Center for Survey Measurement Study Series (Survey Methodology #2018-07)*. U.S. Census Bureau. Available online at <http://www.census.gov/content/dam/Census/library/working-papers/2018/adrm/rsm2018-07.pdf>

1. Introduction

The National Survey of College Graduates (NSCG) is a longitudinal survey sponsored by the National Center for Science and Engineering Statistics (NCSES) within the National Science Foundation (NSF). The Census Bureau has historically served as the NSCG data collection contractor on behalf of NCSES. The survey takes an average of 30 minutes to complete. The survey examines “various characteristics of college-educated individuals in the workforce, including occupation, work activities, salary, the relationship of degree field and occupation, and demographic information.” (<https://www.nsf.gov/statistics/srvygrads/overview.htm>). Most responses are received online. The survey offered an online reporting option in the 2010, 2013 and 2015 NSCG data collections. For those years, the online survey design was optimized for a PC. Beginning in the 2017 survey year, the NSCG will also be optimized for mobile devices.

In preparation for fielding the 2017 NSCG, staff in the Center for Survey Measurement (CSM) of the Census Bureau conducted two rounds of usability testing of the online NSCG instrument on PC and mobile devices between June and September 2016. Thirty individuals participated in the testing, 15 in each round with changes made between the rounds. Testing results and recommendations were provided to the sponsor shortly after testing was finished. The intention of this report is to document the use and redesign of a grid response format for surveys accessed via smaller mobile devices.

2. Background on the use of grid response formats

Grid response options use a row and column matrix design where there is typically one question stem above the matrix, each question item or topic is a label on a row, and the column headings have unique response options that do not vary across those question items. Grids are often used for rating items as shown in Figure 1. In survey design, grids help save space on paper forms because the same question stem and response options are used across items and the response option labels only have to be printed once. Grids also have the unique feature of easily allowing respondents to compare their answers across question items. For mode consistency, often, when a survey uses a grid format on paper, that same format is used in any online response mode as shown in Figure 2 for a PC and Figure 3 for a mobile phone.

While grids have positive space saving features, they are not always recommended by survey designers. The research in fact is mixed. While some researchers have found more missing data, break offs, or straightlining (when the same answer is selected for each of the question topics) with grids, others have not found all those problems (Kolbas 2015; Stern et al. 2015; Wang et al. 2015; Guidry 2013; Toepoel et al. 2009; Callegaro et al. 2009; Couper et al. 2001). Respondents can complete grids faster than other question designs, even when completing a grid on a smaller mobile device (de Bruijne 2015; Couper et al. 2013; Saunders et al. 2013), but Dillman et al. (2014) report that completing a grid is cognitively challenging, no matter the device, especially if the axis labels, either the top labels or the left item labels, scroll off the page. Other researchers such as Smyth and Olson (forthcoming) tested a fully labeled grid on PCs to prevent that from happening. When grids are presented on small mobile devices, having the labels scroll off the visible screen

either horizontally or vertically is a real concern. Figure 3 shows how the some of the labels are not visible and a horizontal scroll is needed to see them.

A28. Thinking about your principal job held during the week of February 1, please rate your satisfaction with that job's...

Mark one answer for each item.

	Very satisfied	Somewhat satisfied	Somewhat dissatisfied	Very dissatisfied
1 Salary	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>
2 Benefits.....	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>
3 Job security	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>
4 Job location	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>
5 Opportunities for advancement.....	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>
6 Intellectual challenge.....	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>
7 Level of responsibility	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>
8 Degree of independence..	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>
9 Contribution to society.....	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>

Figure 1: Question A28 from the 2017 NSCG paper form demonstrating a grid response design.



NATIONAL SURVEY OF COLLEGE GRADUATES

[Save & Exit](#)

Part A - Principal Job

A28. Thinking about your principal job held during the week of February 1, please rate your satisfaction with that job's...

Select one answer for each item.

	Very Satisfied	Somewhat Satisfied	Somewhat Dissatisfied	Very Dissatisfied
1. Salary	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Benefits	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Job security	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Job location	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. Opportunities for advancement	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Intellectual challenge	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. Level of responsibility	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. Degree of independence	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. Contribution to society	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[Previous](#)

[Next](#)

[Burden Statement](#) | [Accessibility](#) | [Privacy](#) | [Security](#)

Figure 2: Question A28 from the 2017 NSCG PC optimized version for usability testing demonstrating a grid response option

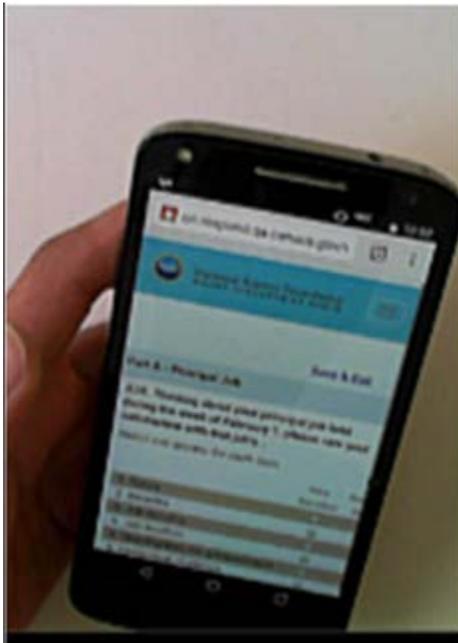


Figure 3: Question A28 from the first round of usability testing for the 2017 NSCG mobile optimized version - demonstrating a grid response option where the response option labels fall off the right side of the page.

Not all grids have more than two response options, and sometimes the entire grid can fit on a page of a mobile phone as shown in Figure 4, where the grid is a series of yes/no questions.

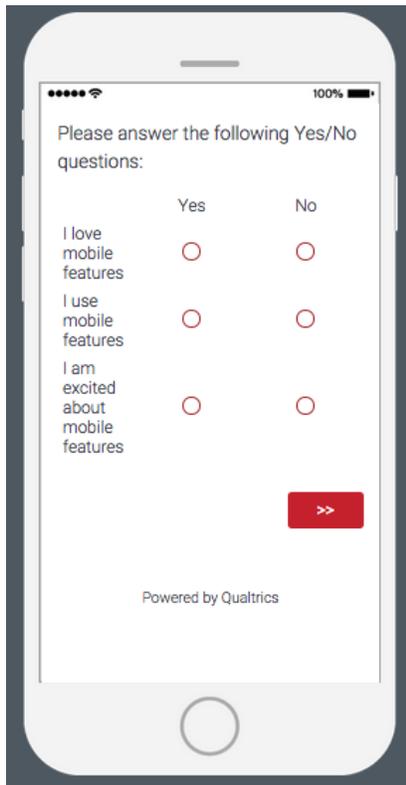


Figure 4: A yes-no grid on a smartphone. Source: Qualtrics. <https://goo.gl/images/hM65L8>

However, because that is not always the case and because of the inherent difficulties with grids, researchers have investigated alternative designs for the grid for mobile, even for yes/no grids. One such design called the stacked grid or banked items displays each item with the responses presented horizontally under each element as shown in Figure 5. Richards et al. (2016) found that the stacked design produced less straightlining than the traditional grid on mobile.

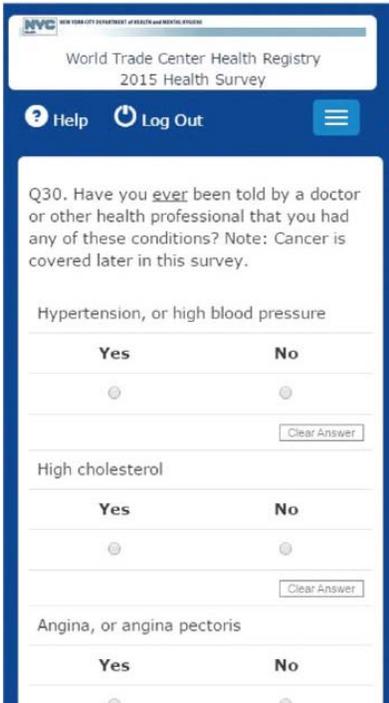


Figure 5: Example of a stack design. Source: Richards et al. 2016.

An alternate design is the item-by-item format where the response options are presented below the question stem, but in a vertical layout on one long scrolling screen as shown in Figure 6. Compared to the traditional grid, the item-by-item design has been shown to produce less straightlining (Borger and Funke 2015; McClain & Crawford, 2013), stronger correlation among items (Lattery et al., 2013), and better survey evaluations (de Bruijne 2015) as summarized by Antoun et al. (2016).

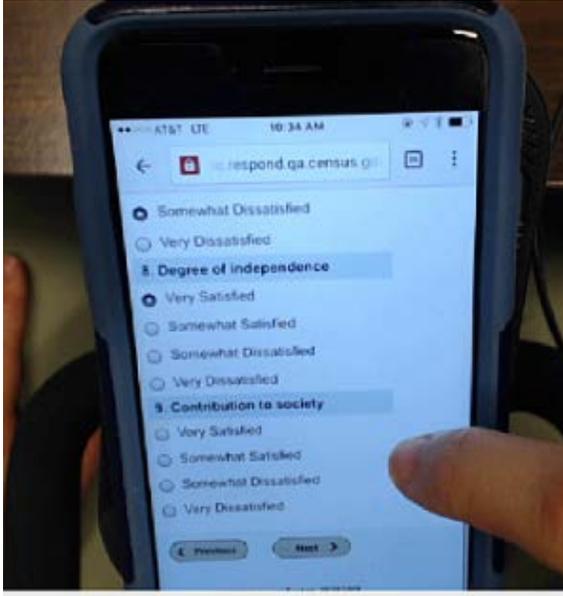


Figure 6: Question A28 from the 2017 NSCG mobile version in Round 2 of usability testing demonstrating an item-by-item design.

This report documents the usability issues with the traditional and the item-by-item response designs as observed in the 2017 NSCG usability testing.

3. National Survey of College Graduates Online Questionnaire

The NSCG online questionnaire is one of many surveys the Census Bureau developed using an in-house codebase framework called Centurion. The survey was designed so that there is generally one question per screen with a 'next' button to navigate forward to the next question (i.e., paging design). A 'previous' button was also available if the respondent wanted to return to a prior question. Skip logic was used within the online survey to navigate the respondent to the next appropriate question in the series based on the answer provided. In 2017, the survey was designed to be both mobile and PC optimized. Optimization means that the screens automatically readjusted in size and in design depending on the display size of the device. Prior to 2017, the survey was optimized for PC only. A respondent could have accessed the survey on their mobile device, but they would have had to zoom in to be able to read the questions on the screens. The goal of mobile optimization was so the respondent would not have to zoom in or out, or scroll left or right, to see, read and answer the survey questions on a mobile device. Figure 7 is an example of a typical screen format for the PC version of the NSCG and Figure 8 is that same question optimized for a mobile phone.

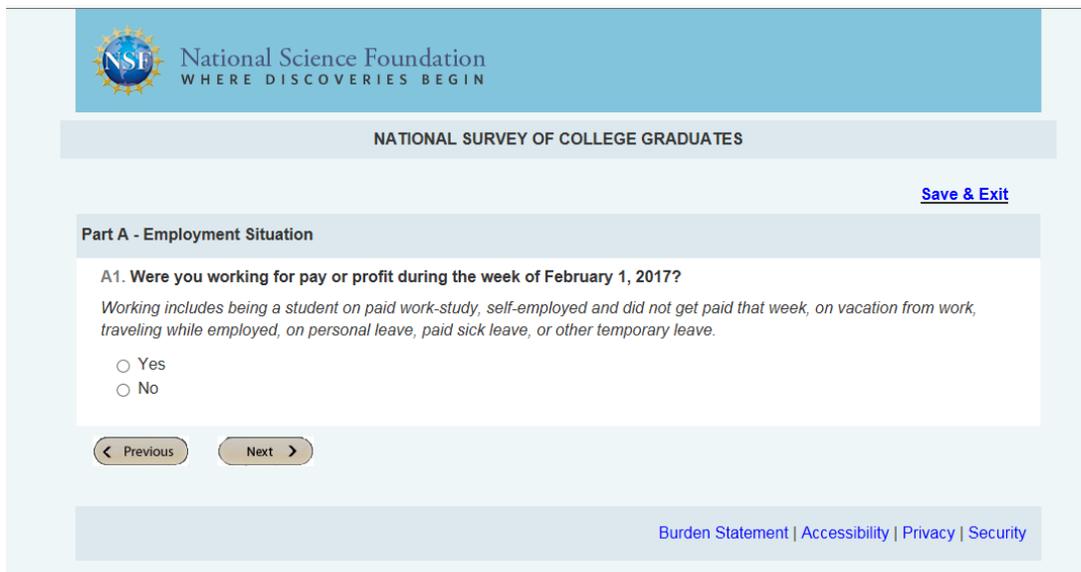


Figure 7: Question A1 in the 2017 NSCG on a PC

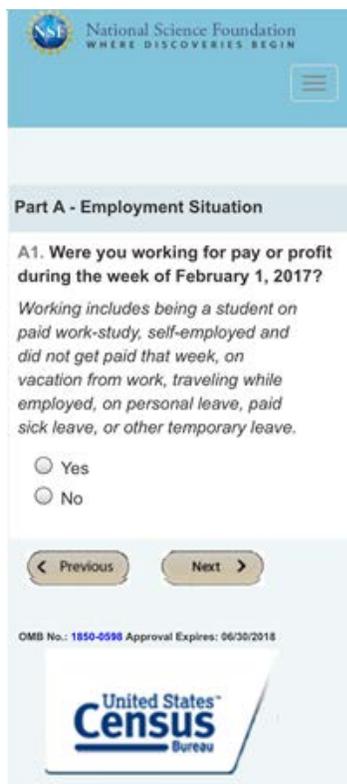


Figure 8: Question A1 in the 2017 NSCG on a mobile phone

The NSCG is a relatively long survey, taking respondents approximately 30 minutes to complete. In total, there are a maximum of 106 questions; however, the number of questions any one respondent receives is often fewer because of skip patterns. Of those 106 questions, 14 of them were “Yes – No” questions using a grid response format and three questions were formatted as

grids but with more response choices as shown earlier in Figures 1-3. ¹ The majority of the questions were either select all that apply, choose one, or other open-ended response questions.

4. Methods

We conducted two rounds of usability testing on the NSCG between June and September 2016, with changes to the online survey made between the two rounds. This section describes the participants, the devices they used to access the survey, and the procedure implemented during the testing sessions.

4.1. Participants

Thirty participants took part in usability testing of the 2017 NSCG. Participants were required to have a minimum education level of a bachelor's degree because this is a requirement for selection into the NSCG sample. Some of the participants had post-bachelor's degrees. Many of the participants had a science, technology, engineering, or math degree and all our participants were employed. We recruited a diverse distribution of respondents, who varied by age, sex and race characteristics. Participant characteristics for each testing round are described in Table 1.

Participants were recruited through advertisements on Craigslist and in an email blast that was distributed to all Census employees. In Round 2, an email to a George Mason University psychology list serve allowed us to recruit several post-bachelor students who were working while attending school. The mean age was lower in Round 2 compared to Round 1 because of these students. All respondents lived in the Washington, D.C. metropolitan area, but many had gone to colleges outside the local geographic area.

¹ We evaluated the new cohort questionnaire for the 2017 NSCG. It contained five sections: Part A—Employment Situation; Part B - Past Employment; Part C—Other Work-Related Experiences; Part D—Education Background; and Part E—Demographic Information.

Table 1: Demographic characteristics of 2017 NSCG Usability Testing Participants

Demographics	Round 1 (n=15)	Round 2 (n=15)
Sex		
Female	10	12
Male	5	3
Age – Mean (Std. Deviation) (Range)	42 (12) (26-67)	32 (7) (24-49)
Race/Hispanic origin		
White	5	6
Black	5	3
Asian	5	4
Hispanic	0	1
Multiracial	0	1
Education		
Bachelor’s degree	7	8
Post-bachelor’s degree	8	7
Employed	15	15

4.2. Devices used

In Round 1, five individuals answered the survey using their smartphone, one used his/her tablet, and the other nine used a Census provided laptop. In Round 2, 10 individuals used their own smartphone; two used their own tablet, and three individuals answered the survey on the census provided laptop (see Table 2 for the summary of devices).

Table 2: Devices used across both rounds of NSCG usability testing

Device	Round 1 (n=15)	Round 2 (n=15)
Census provided laptop	9	3
Dell Latitude E6430, Windows 7		
Smartphone (BYOD)		
iPhone 4S	1	0
iPhone 5	1	2
iPhone 5S	0	3
iPhone 6	1	1
iPhone 6S	0	1
iPhone 6SPlus	0	1
Android Turbo 2	1	0
Android Samsung Galaxy	1	2
Tablet (BYOD)		
iPad 10 inch	1	1
Galaxy S Tablet	0	1

4.3. Procedure

Fifteen participants took part in the first round of usability testing. Results from Round 1 were presented to stakeholders consisting of internal Census Bureau project staff and NCSES staff. Based on the findings from Round 1, changes to the instrument were made prior to the start of Round 2 testing. Fifteen additional participants completed the survey in the second round.

Each usability test conducted in Round 1 and 2 involved the respondent completing the *new cohort version* of the draft online version of the 2017 NSCG while being observed by a Census Bureau test administrator. Participants were instructed to answer the questions as they applied to them in their real life, and we employed a concurrent think-aloud method. All sessions included video recording and audio recording with respondent consent. Each usability session lasted approximately one-and-a-half hours. To offset the costs of parking and travel, participants received an incentive of \$60 for their participation.

Participants were instructed to “think aloud” while completing the survey. The think-aloud technique is modeled on Ericsson and Simon’s (1993) approach to collecting verbal feedback. Our think-aloud protocol was used to maintain a running verbal commentary of the respondents’ expectations and reasoning. A respondent engaging in thinking aloud verbalizes his or her available, conscious thoughts and decisions while completing the tasks so that the researcher can understand the respondents’ cognitive processes as they interact with the interface. The test administrator encouraged the participants to continue to think aloud, using prompts such as, “Keep talking” if they became silent for more than ten seconds.

We collected eye-tracking data as the participant completed the survey. In the first round of testing, we used the Tobii X260 system for eye tracking for laptops and the SMI eye tracker for mobile devices. The SMI eye-tracker requires the participant to wear glasses, but the participant can hold their device as they would if they were at home. In the second round of testing, we used the Tobii X260 system for both the laptop and the mobile data collections. The Tobii X260 mobile eye-tracking system does not require glasses, but the participant must place their phone on a mobile stand and use it from that stand.

After the participant finished the survey, he or she completed a satisfaction questionnaire including opinion questions about general usability aspects of the survey like using forward and backward navigation and comprehension of general survey terminology. Then, the participant was shown a PowerPoint presentation with some screenshots taken from the survey and asked a few questions about those screens. The test administrator followed a protocol during the session and the questions and protocol were approved by OMB using the generic clearance for pretesting 0607-0725.

5. Findings and Recommendations

In Round 1, the grid design used in PCs was also used on smartphones. As the pictures in Figure 9 show, the participant had to scroll to the right (on the phone, to move horizontally the

participant swipes with his or her fingers and not a scroll bar) to see the answer choices when his phone was held in the portrait position for grids with four response choices. One participant held his phone in the landscape or horizontal position when he arrived at grid formatted questions to minimize the horizontal scrolling needed (see Figure 10). The online form auto-rotated allowing the grid to fit better on the screen. However, for respondents who did not use the auto-rotate feature, the usability testing demonstrated what other researchers already documented: when grids are used on small mobile devices, having the labels scroll off the visible screen is a concern.

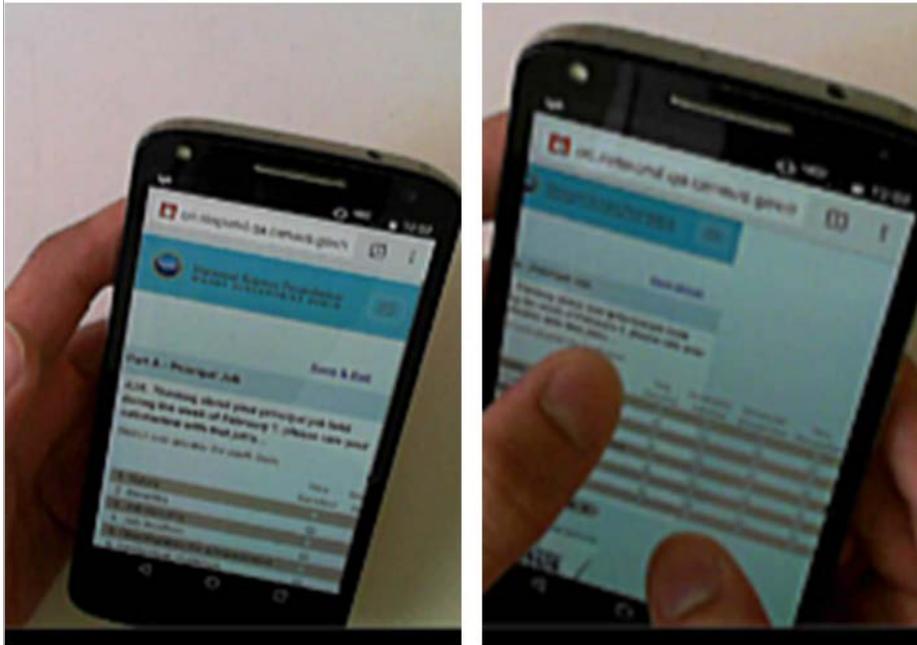


Figure 9: Question A28 from the 2017 NSCG mobile version in Round 1 of usability testing demonstrating a traditional grid response option. Notice that the participant has to scroll to the right to see the response choices and then the question topics scroll off the screen.



Figure 10: Example of a participant who rotated his phone to the landscape orientation and the form auto-rotated so that he could see the entire grid

The need for horizontal scrolling occurred even for grids with yes or no response options on tablets as shown in Figure 11. The horizontal scroll occurred on the tablet because the screen was not optimized for the portrait orientation of this sized tablet. The item labels column was

very wide and had it been programmed to be more narrow for this screen size, the response choices would have been visible on the screen. The optimization for this sized device was corrected prior to fielding the 2017 NSCG.

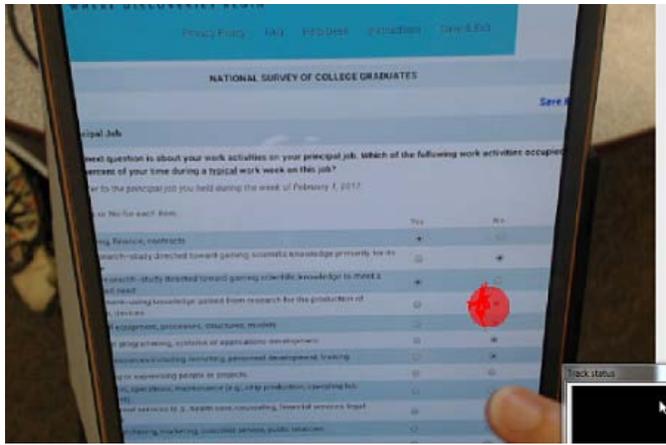


Figure 11: Example of a yes/no grid on a tablet in Round 2 of the NSCG testing. This screen also has suboptimal spacing. Notice that the participant has to scroll to the right to see the response choices. (The red dot was eye tracking data and the small black “Track status” box in the lower right was also present only to the task administrator. The participant could not see the dot nor the box.)

Participants also made spontaneous comments during the session in Round 1 when they came to a grid with the horizontal scrolling needed.

“Maybe it is my smartphone, but having to scroll all the way over and back to see which one it was, also maybe because of my gorilla thumbs, but I’m having to click pretty carefully. .. Maybe holding it portrait or landscape would help.” “... (when asked whether that was something he would normally do) ... No that is not what I would normally do.” Later at another scroll grid, “Same thought process about that menu, having to scroll left and right.” He later changed it to a landscape position for the next grid. To do this, he went into settings and set it to auto-rotate. He said that he usually sets his phone to not rotate because he doesn’t want it rotating when he is using his GPS in his car.

Participants answered every question in the grids in both rounds of testing so we did not observe any missing data on these questions. It could be that participants would otherwise fail to answer a question on their own but felt pressure to do so in the testing circumstance. We suspect that although there was a high incidence of participants scrolling right and left, the shading associated with every other line (i.e., a swim lane design) helped keep respondents on the correct row also reducing any missing data. However, even the *swim lane* design could be confusing to respondents. Between Rounds 1 and 2, the brownish swim lane color (see Figures 9 and 10) was replaced with a blue color as shown in Figures 11 and 12. Although we observed no errors in answering these questions using either swim lane color, one participant asked whether questions

with the blue swim lanes were more important than other questions. In Round 2, blue coloring was used for a variety of purposes within the survey design, such as in the banner at the top of each page and for item labels described later in the report. The various uses of this color could have confused this participant.

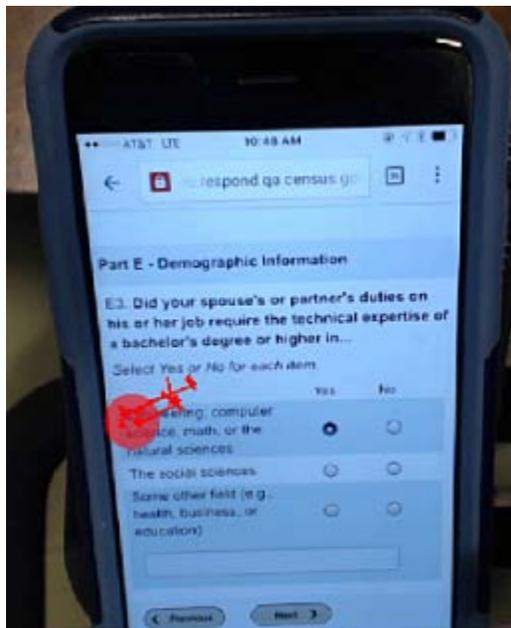


Figure 12: Example of a Yes-No grid on a smartphone in Round 2 of the NSCG testing. The colored swim-lane design appeared to help participants line up their response with the correct category, but one participant wondered aloud whether some topics were more important than other topics. (Red dot is eyetracking data.)

In Round 2, to resolve the horizontal scrolling problem, changes were made to grid questions with four or more response choices. As noted earlier, three grid questions on the NSCG questionnaire had four or more response choices. Those three grid-designed questions became item-by-item questions for both smartphone and tablets. The response choices were placed vertically below the topic of interest and the topic was highlighted in a blue color as shown in Figure 13. There were no usability issues identified with this design during testing and as previously mentioned, all questions were answered, so there was no missing data. Respondents were able to scroll vertically to see and answer all the choices and respondents made no negative comments about the design during the debriefing. The blue highlighting for each question topic appeared to help participants navigate the page. Participants were able to navigate to each item to answer. This item-by-item with blue highlighting change should be considered a success.

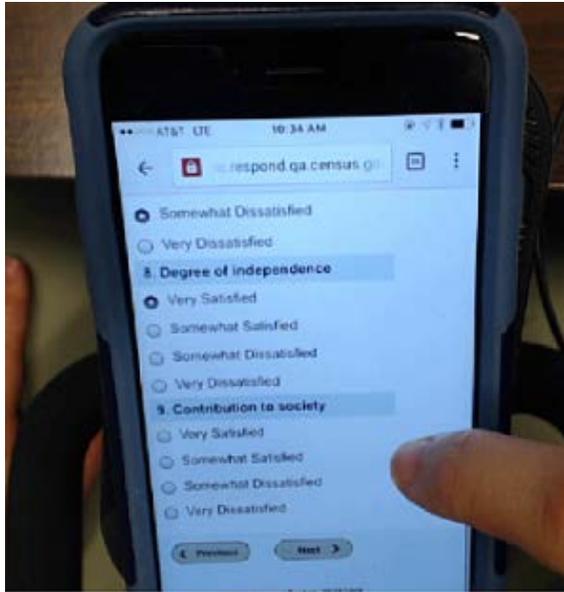


Figure 13: Example of a grid question on a mobile phone in Round 2. There were no usability issues identified with this design. The use of color to denote the items worked well and helped guide the respondent to the next set of response options.

The yes-no questions remained as grids as shown in Figure 12 and aside from one participant’s curiosity regarding swim lane color, no usability issues were observed with this design².

We tested one other design solution with regard to grids for online questionnaires. The functional disabilities question in the NSCG is near the end of the survey and on the paper form, the question is formatted as a grid question as shown in Figure 14.

E25. The next several questions are designed to help us better understand the career paths of individuals with specific functional limitations.

What is the usual degree of difficulty you have with...

Mark one answer for each item.

	None	Slight	Moderate	Severe	Unable to do
	↓	↓	↓	↓	↓
1 SEEING words or letters in ordinary newsprint (with glasses/contact lenses, if you usually wear them)	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>
2 HEARING what is normally said in conversation with another person (with hearing aid, if you usually wear one).....	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>
3 WALKING without human or mechanical assistance or using stairs.....	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>
4 LIFTING or carrying something as heavy as 10 pounds, such as a bag of groceries	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>
5 CONCENTRATING, REMEMBERING, or MAKING DECISIONS because of a physical, mental, or emotional condition	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>

Figure 14: Use of a grid for the functional disabilities question in the 2017 NSCG paper questionnaire.

² There are other possible solutions to the forced-choice (e.g., Yes/No grid) response options, such as using check-all-that-apply options. See Callegaro, Murakami, Tepman, and Henderson (2015) for a meta-analysis comparing these two response option designs.

It was not a grid question on the PC design during the 2015 NSCG, as each question was on a separate screen with the question stem on its own screen as shown in the example in Figure 15. While there were no usability problems identified during the usability testing for that survey iteration, paradata (i.e., production data detailing user actions on each screen) from the 2015 NSCG showed that the question stem screen had a high percentage of break-offs. Based on that finding, survey sponsors agreed to remove that screen and place all the questions in that series on one screen (Holland et al., draft 2016).

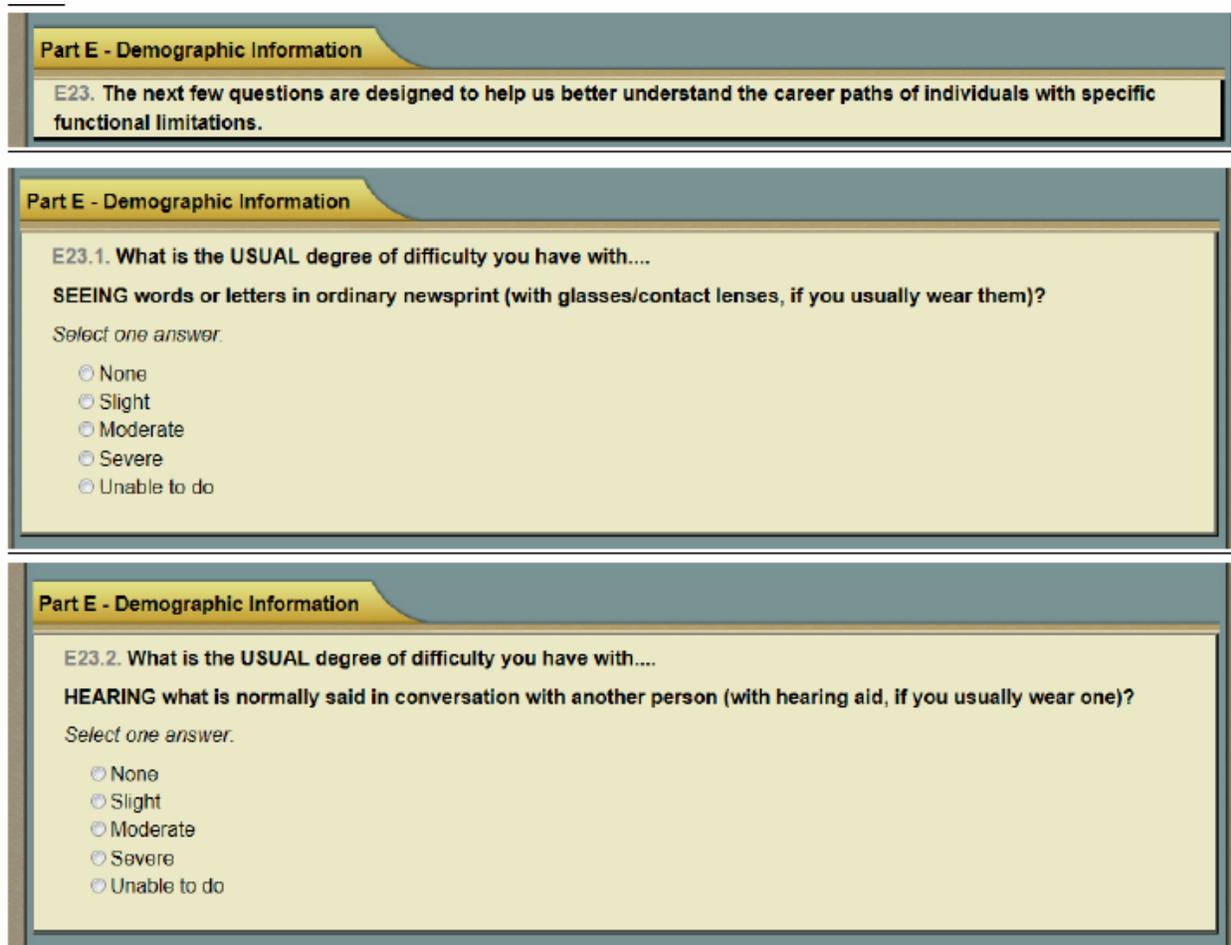


Figure 15: Example of how the functional disabilities questions appeared in the 2015 NSCG on a PC. The example contains the question stem and the first two questions in the series. Each of these items was on a different screen.

For the usability testing in preparation for 2017, the Round 1 design included a radio button design (Figure 16), but then mid-way through that round, the design was changed to a dropdown design (Figures 17 and 18) for both PC and mobile. This decision was made in an effort to save space on smartphones and reduce the amount of vertical scrolling needed. However, there are a few problems with dropdowns, some of which came out during testing.

- With dropdowns, the respondent has to make two touches or clicks for each question to select an answer, which a few participants (but not all) commented on as burdensome or “annoying.”
- A second issue is that the response choices are not immediately visible with dropdowns. One participant commented that he was not sure what the response choices would be when he was presented with the dropdowns. He said that he was expecting some sort of example, like hearing aid or glasses, as an answer choice. This comment suggests that the question is ambiguous as this participant was not able to anticipate the unseen answer choices. Although the respondent appeared to answer the question accurately, the dropdown response design caused additional cognitive burden for him because he originally thought that the question meant one thing, was anticipating an answer choice prior to selecting the dropdown, but then once he saw the actual response choices, he had to return to the question to determine what it was really asking. Had he immediately seen the response choices below the question, as he would in a radio button design, he might not have originally misinterpreted the question.
- A third potential issue, which would only affect iOS device users is that the dropdown choices appear in the bottom half of the screen. In other usability testing, we have found respondents unfamiliar with that design (Nichols, et al. 2015).
- The U.S. Design Web standards also do not recommend dropdowns for questions with fewer than seven response choices regardless of device size (<https://standards.usa.gov/components/form-controls/#dropdown>).

For these reasons and for consistency within the survey, we recommended using the item-by-item design with the radio buttons used for the other grid questions in mobile and PC. This suggestion (as shown in Figure 16) was adopted for the 2017 NSCG.

The screenshot shows the National Science Foundation logo and the title 'NATIONAL SURVEY OF COLLEGE GRADUATES'. Below the title is a 'Save & Exit' link. The main content area is titled 'Part E - Demographic Information' and contains three questions, each with radio button options:

E14.1. What is the usual degree of difficulty you have with...
SEEING words or letters in ordinary newsprint (with glasses/contact lenses, if you usually wear them)?
Select one answer:
 None
 Slight
 Moderate
 Severe
 Unable to do

E14.2. What is the usual degree of difficulty you have with...
HEARING what is normally said in conversation with another person (with hearing aid, if you usually wear one)?
Select one answer:
 None
 Slight
 Moderate
 Severe
 Unable to do

E14.3. What is the usual degree of difficulty you have with...
WALKING without human or mechanical assistance or using stairs?
Select one answer:
 None
 Slight
 Moderate

Figure 16: Question E14 as a radio button design for PC (This occurred for the first few cases in Round 1)

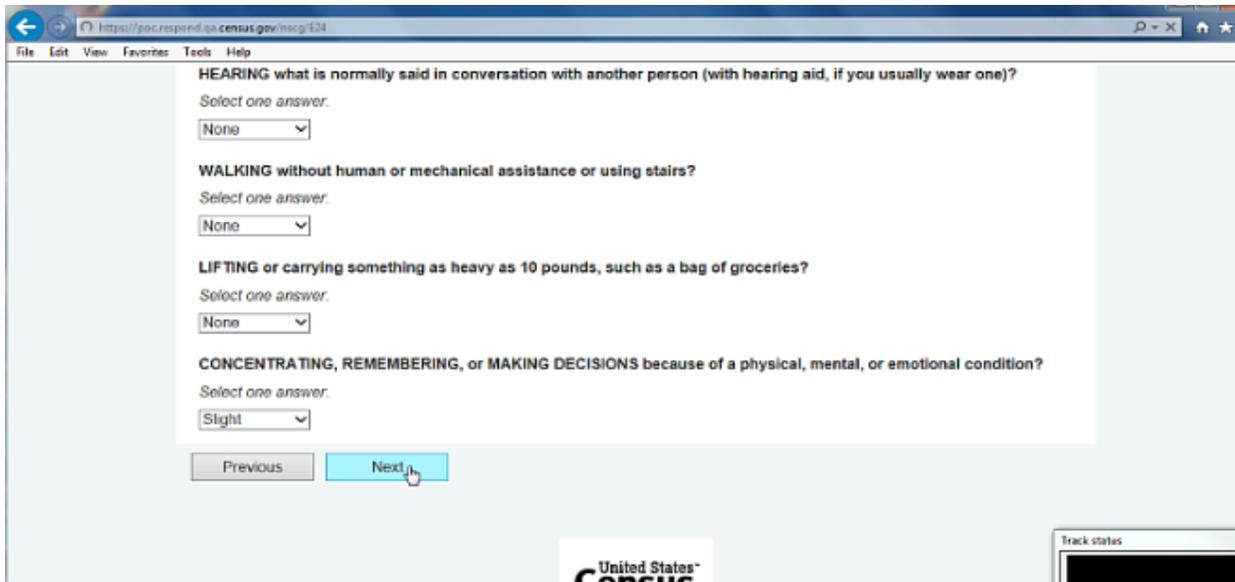


Figure 17: Question E14 as a drop down for PC with answer choices selected in Round 2

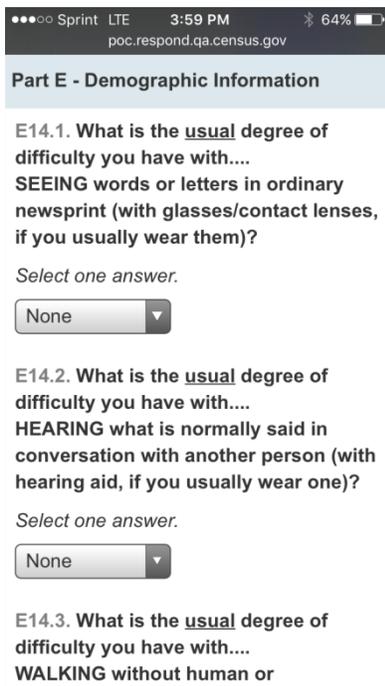


Figure 18: Example of how the design looked on a smartphone with the “None” dropdown response choice selected.

6. Summary and Discussion

The PC-optimized NSCG survey utilized a number of grids for response options. Respondents had some difficulty using these grids when completing the survey on their smartphones. Grids containing four or more response choices were converted to a long vertical list of questions and response options using radio buttons for smartphones, called an item-by-item design. Blue

highlighting identified the question items within the long scrollable page. This design appeared to relieve the usability issues with the former matrix-designed grids on smaller devices. The NSCG continues to use grids for PCs and for yes/no questions for all device sizes.

While the item-by-item design caused no user problems during qualitative usability testing, only the quantitative data on break-offs will be able to prove whether this design works well in the field. We learned from the 2015 NSCG survey results that solving the grid problem by placing each question topic on a different page increased break-offs even though we observed no break-off problems during usability testing conducted on that version of the online survey. Usability testing was not well-suited to uncover the break-off issue. We mention that even though the current item-by-item design with the question topic highlighted worked during the usability testing documented in this report, only analysis of the break-off rates in the 2017 NSCG survey data will allow sponsors to measure whether the item-by-item design maintains engagement within the survey and does not lead to more item nonresponse.

7. Future usability research on grids

Future research should explore comparisons of eye-fixations and fixation time between different grid designs. Eye tracking can collect eye fixations and fixation duration in qualitative study designs. Penzo (2006) has compared fixation data across different designs. He asserts that fixations are an important measure because the number of fixations is associated with cognitive load. For example, he recommends choosing a design with fewer fixations in an experiment because fewer fixations are indicative of less cognitive load. However, we do not have information on whether fewer fixations leads to more accurate reporting.

Figure 19 contains the number of fixations on the screen with grid question for a smartphone (on the left) and the same question on a tablet (on the right) for two participants. The smartphone displayed the question in an item-by-item vertical format. This particular participant had 72 gaze fixations. The participant who used the tablet with a horizontal grid for the same question had 97 fixations. In the present study, the eye tracking data was not consistent enough to do a comprehensive quantitative comparison between vertical and horizontal grids, but future research could investigate this aspect more as it is an independent source of data to measure burden. For example, we could compare fixation data for an item-by-item design for a yes-no response format against the current grid format or we could further the research on the dropdown response options by comparing fixation data from that design to an item-by-item design to provide more data to determine whether the dropdown design increases cognitive load.

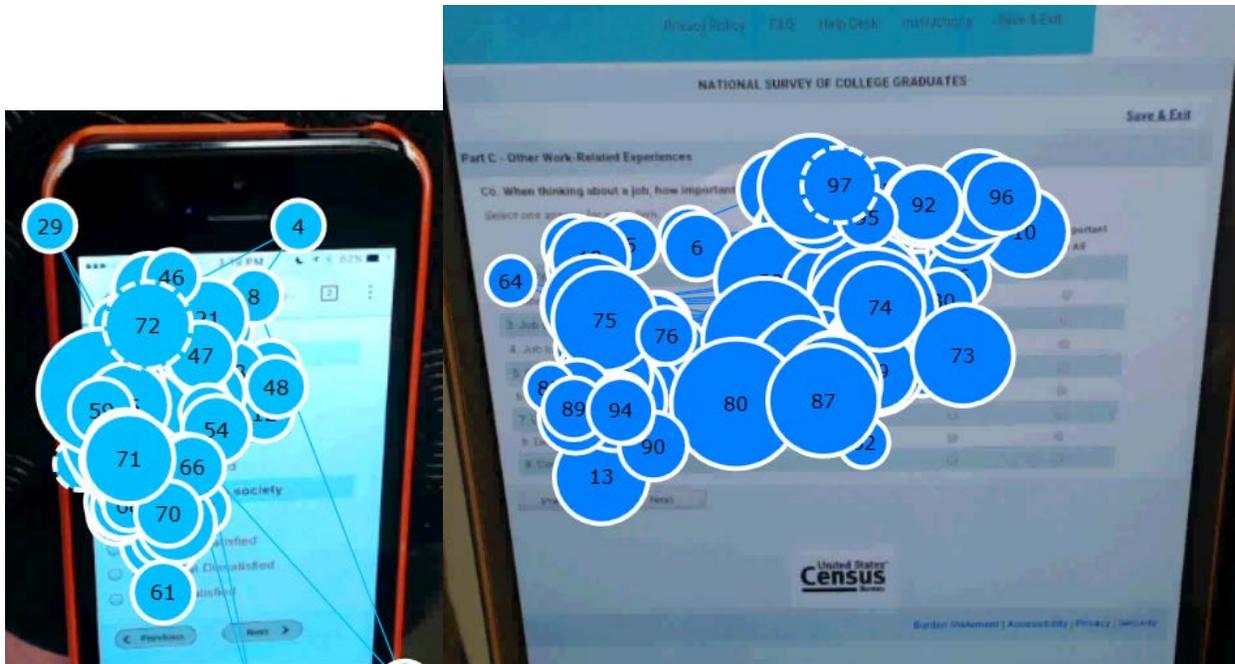


Figure 19: On left, example of a vertical grid question on a mobile phone; on the right, a horizontal grid question on a tablet in Round 2.

Finally, future research should explore whether response distributions are any different under the alternative designs. Ideally, this research would include some kind of validation data source to shed some light on which format produces more accurate data.

9. References

- Antoun, C., Katz, J., Argueta, J., Keegan, R., Gareau, M., & Wang, L., (2016). Questionnaire Design Considerations for Mobile Web Surveys – A Literature Review. U.S. Census Bureau. Internal Presentation April 26, 2016.
- Borger, C., & Funke, F. (2015). Responsive Questionnaire Design for Higher Data Quality in Mobile Surveys, General Research Conference (GOR) 2015.
- Callegaro, M., Murakami, M. H., Tepman, Z., & Henderson, V. (2015). Yes-no answers versus check-all in self-administered modes. A systematic review and analyses. *International Journal of Market Research* 57(2): 203-223. DOI: 10.2501/IJMR-2015-014.
- Callegaro, M., Shand-Lubbers, J. & Dennis, J.M. (2009). Presentation of a single item versus a grid: effects on the vitality and mental health subscales of the SF-36v2 health surveys. Paper presented at the annual meeting of the American Association for Public Opinion Research, Hollywood, FL. Available at http://www.websm.org/uploadi/editor/1368003524Callegro_2009_Presentation_of_single_item.pdf
- Couper, M.P., M.W. Traugott & M.J. Lamais. (2001). Web survey design and administration. *Public Opinion Quarterly* 65(2): 230–253.
- Couper, M.P., R. Tourangeau, F.G. Conrad & C. Zhang. (2013). The design of grids in web surveys. *Social Science Computer Review* 31(3): 322–345.

- De Bruijne, M.A. (2015). Designing web surveys for the multi-device internet Tilburg: CentERdata, Center for Economic Research available online at: [https://pure.uvt.nl/portal/en/publications/designing-web-surveys-for-the-multidevice-internet\(19e4d446-a62b-4a95-8691-8e75eb69b611\).html](https://pure.uvt.nl/portal/en/publications/designing-web-surveys-for-the-multidevice-internet(19e4d446-a62b-4a95-8691-8e75eb69b611).html)
- Dillman, D.A., J.D. Smyth & L.M. Christian. 2014. *Internet, phone, mail, and m-mode surveys: the tailored design Method*. John Wiley & Sons, Hoboken, NJ.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis: Using Verbal Reports as Data*. Cambridge, MA: The MIT Press.
- Guidry, K. R. (2012). Response Quality and Demographic Characteristics of Respondents Using a Mobile Device on a Web-based Survey. Paper Presented at the AAPOR annual conference, Orlando, USA. Available at <http://cpr.indiana.edu/uploads/Response%20Quality%20and%20Demographic%20Characteristics%20of%20Respondents.pdf>.*
- Holland, T., Reiser, C., Horwitz, R., & Tancreto, J., (draft). Web Enhancement for the 2017 National Survey of College Graduates. U.S. Census Bureau. Internal Memorandum.
- Kolbas, V. (2015, July). *The measurement effect in PC smartphone and tablet surveys*. Paper presented at the ESRA biannual conference, Reykjavik, Iceland.
- Lattery, K., & Park Bartolone, G. (2013). Optimizing Surveys for Smartphones: Maximizing Response Rates While Minimizing Bias. Paper presented at the 2013 CASRO Online Research Conference.
- McClain, C. & Crawford, S., (2013). Grid Formats, Data quality, and Mobile Device Use: Toward A Questionnaire Design Approach. Paper presented at 2013 AAPOR, Boston MA. http://www.aapor.org/AAPOR_Main/media/AnnualMeetingProceedings/2013/Session-F-4-4-McClain.pdf
- Nichols, E., Clark Fobia, A., Holzberg, J., Holland, T., Malakhoff, L. & DeMaio, T. (2014). 2015 National Survey of College Graduates. Cognitive and Usability Pretesting Results: Rounds One and Two. U.S. Census Bureau: Internal Memorandum (September 30, 2014).
- Nichols, E., Olmsted-Hawala, E., Horwitz, R. & Bentley, M. (2015). Optimizing the Decennial Census for Mobile, - A Case Study. Federal Committee on Statistical Methodology. Washington D.C. Available at: http://fcsml.sites.usa.gov/files/2016/03/12_Nichols_2015FCSM.pdf
- Penzo, M. (2006). Label Placement in Forms from, <http://www.uxmatters.com/mt/archives/2006/07/label-placement-in-forms.php> (accessed 8/18/2016)
- Richards, A., Powell, R., Murphy, J., Nguyen, M., & Yu, S. (2016). Gridlocked: The impact of adapting survey grids for smartphones. *Survey Practice*, 9(3). Retrieved from <http://www.surveypractice.org/index.php/SurveyPractice/article/view/329>
- Saunders, T., Chavez, L., Chrzan, K., & Brazil, J. (2013). Scale Orientation, Grids and Modality Effects in Mobile Web Surveys http://www.aapor.org/AAPOR_Main/media/AnnualMeetingProceedings/2012/03_Saunders_F1_Maritz-Decipher-Mobile-Web-Survey-Presentation-5-18.pdf

- Smyth, J., & Olson, K., (forthcoming.) A Comparison of Fully Labeled and Top Labeled Grid Question Formats. University of Nebraska-Lincoln. Presented at QDET2 – Miami FL 2016
- Stern, M. J., Sterrett, D., Bilgen, I., Raker, E., Rugg, G., & Baek, J. (2015, May). *The effects of grids on Web surveys completed with mobile devices*. Paper presented at the AAPOR annual conference, Hollywood, FL.
- Toepoel, V., M. Das & A. Van Soest. (2009). Design of web questionnaires: the effects of the number of items per screen. *Field Methods* 21(2): 200–213.
- Wang, M., McCutcheon, A. L., & Allen, L. (2015, May). *Grids and online panels: A comparison of device type from a survey quality perspective*. Paper presented at the AAPOR annual conference, Hollywood, FL.