Eye Tracking, Usability and Accessibility Results from the
Online Version of the 2018 National Survey of Children’s Health

Elizabeth M. Nichols
Yazmín A. García Trejo
Adriana Hernández-Viver
Lawrence A. Malakhoff
Alda G. Rivas
Rodney Terry

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

Report issued: June 27, 2019

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. The Census Bureau’s Disclosure Review Board and Disclosure Avoidance Officers have reviewed this product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release, CBDRB-FY19-CED001-B0010.
Abstract: This report documents the usability testing evaluation of the 2018 National Survey of Children’s Health. The PC-version and mobile-optimized version of the survey were pretested in both English and Spanish. The 12 English-speaking and 5 Spanish-speaking participants found similar problems. All participants found the questionnaire long with many questions seemingly unnecessary for healthy, typically developing children. While all English-speaking participants finished the survey, only one Spanish-speaking participant was able to finalize and submit the entire survey within the 90-minute usability session. The design of the page for listing the children in the household was complex. Participants did not know which button to select to add a child’s name. For a few questions, participants could answer either in metric or U.S. customary units (for example, height was measured in feet or meters), or in months or years. Participants were supposed to select one of the units to answer. However, participants answered using all the options available to enter information. For example, for children’s height, participants provided information in both measurement systems (meters and feet) and for age, they entered information in both years and months. Participants did not realize they needed to choose one option over the other. The spacing and size of response fields and response choices when accessed on iPhones was poor, with small and closely spaced radio buttons, which when selected did not appear selected. Participants did not use or understand the progress bar at the bottom of the screen on PCs. Like other surveys, the review screen was not helpful and not used by any participants to review their answers. In addition, many Spanish speakers did not find the language toggle to Spanish on the first page.

Keywords: online survey design, English and Spanish usability testing, effectiveness, efficiency, satisfaction

Executive Summary

In 2018, staff in the Center for Behavioral Science Methods (which at that time was called the Center for Survey Measurement) of the U.S. Census Bureau conducted usability testing of the 2018 National Survey of Children’s Health (NSCH) online survey in English and in Spanish. Twelve English-speaking and five Spanish-speaking participants took part in the usability testing. During the usability sessions, participants accessed the survey either on a Census Bureau provided laptop computer or on their own mobile device. During that testing, often cognitive issues arose with the questions and response choices. This report provides a summary of the usability, cognitive and accessibility methods and results from evaluating the NSCH online survey.

Similar problems were found by both the English-speaking and Spanish-speaking participants.

- All participants found the questionnaire long with many questions seemingly unnecessary for healthy, typically developing children. While all English-speaking participants finished the survey, only one Spanish-speaking participant was able to finalize and submit the entire survey within the 90 minute usability session.
- The design of the page for listing the children in the household was complex. Participants did not know which button to select to add a child’s name.
- For a few questions, participants could answer either in metric or U.S. customary units (for example, height was measured in feet or meters), or in months or years. Participants were supposed to select one of the units to answer. However, participants answered using all the options available to enter information. For example, for children’s height, participants provided information in both measurement systems (meters and feet) and for age they entered information both in years and months. Participants did not realize they needed to choose one option over the other.
- The spacing and size of response fields and response choices when accessed on iPhones was poor, with small and closely spaced radio buttons, which when selected did not appear selected.
- Participants did not use or understand the progress bar at the bottom of the screen on personal computers (PCs).
- Like other surveys, the review screen was not helpful and not used by any participants to review their answers.

There were a number of usability and cognitive challenges with the Spanish version of the instrument. For example, Spanish speakers had difficulty changing the instrument language to Spanish because they had trouble noticing the language toggle in the landing page of the survey (see page 11 for a more detailed explanation). While testing, researchers noted that there were misplaced response options and missing series of questions. Spanish-speaking participants found a number of grammatical errors in the online survey. A copy-editing process in Spanish is a must before finalizing the online instrument.
Other medium and low priority usability issues are reported for both languages, as well as comprehension issues regarding the questions themselves. Satisfaction ratings, which were generally high among the participants, are also reported.

Accessibility testing was conducted on the survey by a staff member familiar with the procedures of the accessibility software. The main purpose of accessibility testing was to confirm if what is vocalized by the screen-reader matches the visible text. One of the English-speaking participants for usability testing was visually impaired. She used the JAWS screen-reader for PC and confirmed many of the findings from the staff-conducted testing. Accessibility findings are documented in this report.
5. Accessibility Findings .................................................................................................................. 50

6. Cognitive Findings and Recommendations .............................................................................. 57

   6.1. Debriefing responses ........................................................................................................... 57

   6.2. Spontaneous comments ...................................................................................................... 73

       English-specific comments: ................................................................................................. 74

       Spanish specific comments: ............................................................................................... 75

7. Limitations ................................................................................................................................. 78

8. Future usability research on designing for branching logic ...................................................... 78

9. Acknowledgements .................................................................................................................. 79

10. References ............................................................................................................................... 79
1. Introduction

The National Survey of Children’s Health (NSCH) is sponsored by the Maternal and Child Health Bureau of the Health Resources and Services Administration, an agency of the U.S. Department of Health and Human Services (http://childhealthdata.org/learn/NSCH/resources/methods). The NSCH examines the physical and emotional health of children 0-17 years of age. This survey was conducted via an interviewer-administered telephone survey in 2003, 2007, and 2011/12 by the National Center for Health Statistics at the Centers for Disease Control (http://childhealthdata.org/learn/NSCH). In 2016, the U.S. Census Bureau carried out the data collection using a self-administered paper form and a Web questionnaire option for the first time. That Web questionnaire was optimized for PC only, meaning that it displayed as intended only on PCs. When displayed on a tablet or smartphone, the respondent would have to zoom in to see the questions.

For the 2018 NSCH, the survey was designed to render optimally on PC and mobile platforms, called ‘PC and mobile optimization.’ Optimization means that the screens automatically readjusted in size and in design depending on the display size of the device. When displayed on a tablet or smartphone, the respondent would not have to zoom in to see the questions.

In preparation for fielding the 2018 NSCH, usability testing of the English and Spanish versions of the proposed 2018 online Web instrument was conducted between April 11 and 30, 2018. The goal of the testing was to uncover usability issues with the current design (on both PC and mobile devices), and make recommendations for enhancements to ensure that the online survey performed optimally during data collection once the instrument was fielded in July of 2018. We also gathered feedback on new questions added in the survey. Twelve individuals participated in the usability testing of the English version of the NSCH and five individuals participated in testing the Spanish version.

As with many of our usability projects, we also evaluated the instrument for compliance to Section 508 regulations. Software applications comply if users can navigate through them and hear all visible directions and cues from a screen reader. In this study, a mobile user would hear feedback from TalkBack and a PC user would hear feedback from JAWS. The accessibility testing of the PC and mobile versions of the questionnaire was conducted only for the English version due to a lack of bilingual accessibility testing staff. A staff person trained in testing for accessibility conducted the testing and we had one English-speaking participant use assistive technology to access the questionnaire.

The NSCH online questionnaire used the same questions as in the paper questionnaire (see Attachment A for the English version and Attachment B for the Spanish version). Though the usability testing protocol targeted primarily usability issues and cognitive issues related to the new questions, during the course of testing we observed cognitive processing issues with some of the pre-existing questions from previous NSCH data collections. Testing results and recommendations for usability, cognitive and accessibility issues are described in this report.
2. National Survey of Children’s Health Online Questionnaire

The 2018 NSCH online questionnaire is one of many surveys the Census Bureau developed using an in-house codebase framework called Centurion. The NSCH contains a screening survey (S1) and then depending upon the ages of the children listed in the screener, respondents are directed to one of three different modules in the NSCH—module T1 covers a child 0 to 5 years old, module T2 covers a child 6 to 11 years old, and T3 covers a child 12 to 17 years old. If more than one child is listed in the screener, only one of them is selected based upon sampling criteria. Each module covers topics appropriate for the age of the child. Module T1 has sections on infant and child developmental milestones. Modules T2 and T3 are similar to each other, covering topics such as the child’s health, communication with health care providers, insurance, schooling and activities, and caregiver backgrounds. The number of questions differs by modules; T1 includes 182 questions, T2 includes 161 questions and T3 includes 171 questions. Within each module, some questions are skipped depending upon the respondent’s answers.

The average expected time to answer the NSCH (with the screener and one module) is 33 minutes. The survey is available in English and Spanish with a language toggle link on the web instrument login screen where the respondent can change from the English to the Spanish version or vice versa by clicking the link. The default setting is the English version. The instrument includes functionality for respondents to save their answers, logout and return later to finish the survey. For each section of the survey, there were multiple web pages with a ‘Next’ button to navigate forward to the next web page and a ‘Previous’ button if the respondent wanted to return to a prior page. Most web pages in the survey contained only one question; however, several web pages contained multiple related questions with skip sequences built into the page where the questions would become enabled or disabled depending upon the answer to the filter question on that page. There was also some branching logic between web pages within the survey, and if that logic was triggered, whole sections of questions could be skipped. There were very few, if any, edit messages or alerts within the survey, meaning that participants could leave questions blank and navigate to the next page.

3. Methods

We conducted one round of usability testing on the English and Spanish versions of the NSCH from April 11-30, 2018. This section describes the participants, the devices they used to access the survey, and the usability evaluation methods implemented during the testing sessions. The last section of this report describes the accessibility testing methods.

3.1. Participants
As shown in Table 1, twelve participants took part in usability testing of the 2018 NSCH English version. One of the 12 participants was vision impaired and she completed the T2 instrument using the JAWS screen-reader. In terms of child health issues, participants included someone with a child with ADHD and someone with an obese child. Another participant had a child with
health issues, but that was not the selected child for the survey. Other than these participants, the remaining English speakers had healthy, typically developing children.

Five participants took part in usability testing of the 2018 NSCH Spanish version. Six participants were originally scheduled, but one Spanish-speaking participant refused to participate after signing the consent form. This participant felt uncomfortable with the equipment (e.g. laptop and recording devices) and thought the test administrators were in reality immigration enforcement officials. None of the participants for the testing in Spanish had children with health issues.

### Table 1: Distribution of 2018 NSCH Usability Testing Participants across Survey Modules

<table>
<thead>
<tr>
<th>Screener and T1</th>
<th>Screener and T2</th>
<th>Screener and T3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 English</td>
<td>4 English</td>
<td>4 English</td>
<td>12 English</td>
</tr>
<tr>
<td>2 Spanish</td>
<td>1 Spanish</td>
<td>2 Spanish</td>
<td>5 Spanish</td>
</tr>
<tr>
<td>6 Participants</td>
<td>5 Participants</td>
<td>6 Participants</td>
<td>17 Participants</td>
</tr>
</tbody>
</table>

Source: 2018 NSCH usability testing

Participant characteristics are described in Table 2.

### Table 2: Demographic characteristics of 2018 NSCH Usability Testing Participants

<table>
<thead>
<tr>
<th>Demographics</th>
<th>English n=12</th>
<th>Spanish n=5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>Male</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Age</strong> – Mean (Std. Deviation) (Range)</td>
<td>38 (8) (32-58)</td>
<td>38 (12) (19-52)*</td>
</tr>
<tr>
<td><strong>Race/Hispanic origin</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White/nonHispanic</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>White/Hispanic</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Black/nonHispanic</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Black/Hispanic</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Asian</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Hispanic – Mexico</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Hispanic – Central America</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Hispanic – South America</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Bachelor’s</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Post Bachelor’s</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Missing</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Source: 2018 NSCH usability testing

*Age calculations include four of the five participants because one person gave an unrealistic birth year and therefore it was not taken into account in the calculations.

English-speaking participants were recruited through advertisements on Craigslist, personal connections, an email blast that was distributed to all Census Bureau employees in March, and a community mom’s listserv posting. Spanish-speaking participants were recruited using two sources: contacts from our existing recruitment database of potential participants screened for
previous studies; and participants recruited onsite at one local community center in Maryland that serves the local Latino community living in the Washington, DC metropolitan area.

To be selected, all participants had to have a child that fell into one of the age ranges for the modules, and they needed at least one year of experience using the Internet. All Spanish speakers selected reported being of Latino, Hispanic, or Spanish origin; that Spanish was their native language and that they spoke Spanish better than or equally well to English. All participants lived in the Washington, D.C. area.

3.2. Devices used
Six English-speaking participants answered the survey using a census-provided laptop; and six used their own smartphone. Three Spanish-speaking participants answered the survey using a census-provided laptop; and two used their own smartphone (see Table 3 for the summary of devices by module).

Table 3: Devices used in the usability testing of the NSCH English and Spanish versions

<table>
<thead>
<tr>
<th>Device</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Census provided laptop</td>
<td>2 English</td>
<td>2** English</td>
<td>2 English</td>
</tr>
<tr>
<td>Dell Latitude E6430, Windows 7</td>
<td>2 Spanish*</td>
<td></td>
<td>1 Spanish</td>
</tr>
<tr>
<td>Smartphone (BYOD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>iPhone 8</td>
<td>0</td>
<td>0</td>
<td>1 English</td>
</tr>
<tr>
<td>iPhone 6</td>
<td>1 English</td>
<td>0</td>
<td>1 English</td>
</tr>
<tr>
<td>iPhone 5</td>
<td>1 English</td>
<td>0</td>
<td>1 English</td>
</tr>
<tr>
<td>iPhone – no version specified</td>
<td>0</td>
<td>1 Spanish</td>
<td>1 Spanish</td>
</tr>
<tr>
<td>Android one+five</td>
<td>0</td>
<td>1 English</td>
<td>0</td>
</tr>
<tr>
<td>Android Samsung 7</td>
<td>0</td>
<td>1 English</td>
<td>0</td>
</tr>
</tbody>
</table>

Source: 2018 NSCH usability testing

*Both participants started with smartphones (iPhone and Android) and switched to PC during the survey. One participant had a very slow internet connection and the other participant never could proceed past the security warning message and the researcher was unable to help the individual.

**One participant used the screen-reader JAWS to access and answer the survey.

3.3. Usability Testing Procedure
The usability test involved the participant completing the draft online version of the 2018 NSCH while being observed by a Census Bureau test administrator. All sessions included video and audio recording with participant consent. Each usability session lasted approximately one-and-a-half hours. To offset the costs of parking and travel, participants received an incentive of $40 for their participation. Eight English sessions occurred in the usability lab at the Census Bureau’s headquarters building, and the other four English sessions occurred at libraries in the Washington DC metro area. All Spanish sessions took place at two community centers in the area.

Several sessions were observed by other staff, including staff who work on the survey. Observers watched and listened in a separate area for sessions conducted at headquarters but were in the room at the sessions conducted away from headquarters.
During the usability sessions, participants were instructed to complete the survey as they would if they were home with two exceptions.

- For participants with multiple children of different ages, we had to instruct them which child to list in the screener so that the correct topical module would be triggered because the NSCH only selects one child in the household about whom to ask detailed questions. If the participant had multiple children in the same age range—we let them list all of those children. Other than this instruction, participants were instructed to answer the questions as they applied in their real life circumstances.
- Approximately midway through the topical module, participants were asked to exit the survey pretending that they would resume later to complete it. This was to test the resume survey feature with the PIN and/or security question.

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 participants’ expectations and reasoning. A participant engaging in thinking aloud verbalizes his or her available, conscious thoughts and decisions while completing the tasks so that the researcher can understand the participants’ cognitive processes as they interact with the web survey 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. After the first section in the module, participants were asked to exit the survey pretending that they would resume later to complete it.

We collected eye-tracking data for three English-speaking participants who completed the survey on the census-provided laptop using the Tobii X2-60 system. The Tobii system was also used to record the audio and video of the participant answering the survey. We did not collect eye tracking for the mobile devices because the effort needed to analyze that data outweighed any insights learned about the design on such a small screen. Instead we used Camtasia on another Census laptop to record the screen capture of the participant answering the NSCH on their mobile device. To make the video recording, a Webcam attached to a tripod recorded participants using their hand-held device.

Eye tracking data were collected in the English-speaking sessions because these data can be useful in confirming difficulties with the design or with the question wording that are uncovered during testing. For example, unusual patterns from the eye tracking heat maps can identify challenging words, phrases, and design layouts. A more typical pattern on a screen can also confirm that there were no issues with either the words or the design used.

We decided not to gather eye-tracking data from Spanish-speakers because sessions with monolingual or Spanish-dominant Spanish-speakers take longer to conduct than English-speaking sessions and adding the eye tracking calibration task, which can sometimes take 5 minutes to

---

1 We did not collect eye-tracking data from the JAWS screen reader participant. We also did not collect eye tracking in the offsite locations that used the laptop.
complete, would lengthen the session even more. Additionally, because eye-tracking is not well known, sometimes the addition of that method makes Spanish-speaking participants even more wary of participating in a government-run session.

After the participant finished the survey, he or she completed a satisfaction questionnaire including opinion questions about general usability aspects of the survey (e.g., 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, focusing on the new questions that were added in the 2018 NSCH. The test administrator followed a protocol during the session and the questions and protocol were approved by OMB on March 23, 2018 using the generic clearance for pretesting 0607-0725.

3.4. Accessibility Testing Procedure
Accessibility testing uses methods to determine if a disabled user can use the system. The accessibility tester is looking to see if there are any instances where the content presented on the screen visually is not presented orally; whether the tab sequence allows users to navigate to every object on the screen and whether the user can navigate both forward and back and enter data correctly. Sometimes what is spoken by the JAWS software is not seen on the screen.

Concurrent to the NSCH usability testing, accessibility testing was performed on the NSCH using the Job Access With Speech (JAWS) 18 screen-reader and the Internet Explorer 11 browser for the Windows 7 desktop platform and the Chrome 66 browser and Google TalkBack 6.1 screen-reader running on the Android 7 platform for smartphones by a staff member trained in testing for accessibility. This staff member navigated through the survey for three times on the desktop, once for each age group (0-5 years, 6-11 years, and 12-17 years), and once on the smartphone for the oldest age group. Navigation for JAWS users was accomplished through tab and arrow keys and for TalkBack users by left and right finger swipes.

Although usually there is not a need for someone with low vision to do the testing, for this testing, one participant who called to take part in the usability testing had low vision. We had her use JAWS as she completed the NSCH screener and T2 survey on a desktop.

3.5. Usability Testing Analysis
We focused our analysis on the three measures of usability: effectiveness, efficiency and satisfaction.

To measure effectiveness, we used observations of the sessions (both from real time and from the audio and video recording), heat maps of the eye-tracking data, think-aloud comments and debriefing probes. Eye tracking data were collapsed across the three English-speaking participants to form heat maps. These maps are color-coded. The colors are red, green and yellow and show the general pattern of where participants looked on the screen. Red areas are
where the participants looked the most followed by yellow and then green. The more intense the color, the more they looked at that place.

To measure efficiency, we recorded spontaneous participant comments about the length of the survey. We recorded general time to complete the survey, but we did not use an exact timing measure because our sample size was too small to make generalizations about the time needed to complete the survey.

Immediately after each usability session, the researchers summarized the usability and cognitive findings and provided the summary to the sponsor. The summary was based upon the observed behavior of the participant and the verbal feedback. These summaries were used to create a “Quick Report” which included user-centered design issues observed, cognitive problems with the questions and response options, and recommendations. Accessibility testing and analysis took place over the course of a few days by the researcher.

4. Usability Findings and Recommendations

4.1. Effectiveness

The following findings and recommendations are in order of severity from the highest priority usability issues to the lowest. Unless noted otherwise, these findings were uncovered both with English and Spanish speakers.

4.1.1. Radio buttons and field length design issues

Mobile phone users had problems selecting radio buttons because they were too close together (see Figure 1); that is, even participants with good dexterity could not always activate the small radio button they wanted because sometimes their finger press would inadvertently activate a different response. On iPhones there was not enough visual distinction between whether a radio button was selected or not selected (see the very faint distinction between the selected ‘Yes’ and the not selected ‘No’ in Figure 2). Errors were made in selection, that is, the wrong choice was selected and sometimes no choice was selected at all. Three Spanish-speaking participants spontaneously mentioned having difficulty selecting radio buttons and/or seeing their selected responses.

The lack of spacing between radio buttons was also an issue that arose for Spanish speakers using the PC, but not for English speakers on the PC.

Additionally, on iPhones some fields were not big enough to show the entire entered answer. The top arrow in Figure 2 points to an open-text box where the number ‘10’ has been entered but the ‘0’ is not fully-displayed; it is essentially cut in half.
Figure 1. Example of proximity of the radio buttons on an Android device
Source: 2018 NSCH usability testing
Recommendation: We recommend increasing the space between radio buttons to match the space used in the National Sample Survey of Registered Nurses for both PC and mobile (see Figure 3). That spacing worked well during testing of that instrument (Nichols, Kephart, and Malakhoff, 2018). We recommend increasing the response field size to match the expected length of the entry.
4.1.2. Programming errors were found in Spanish.

In the Spanish instrument, participants were unable to enter a response in some of the fields (such as in the example in Figure 4). For this specific example, the participant was asked to report the times he or she has moved to a new address. Every time the participant tried to enter information into the response field, this person was unable to do so even after clicking several times. This barrier to enter information on the text field appeared to be an instrument programming malfunction.
Recommendation: More robust user acceptance testing would have uncovered any inaccessible text field. During that type of testing, we recommend checking all enabled text fields to make sure that they can be accessed by the user for entering text.

4.1.3. Spanish speakers did not use the language toggle link

After starting the survey on the default English language page, three Spanish-speaking participants did not see the link to change the language (see Figure 5). This happened both with PC and mobile users. The test administrator had to lead participants to show them how to switch the language. Testing has consistently showed that this type of design is not optimal (Lykke and García Trejo, 2018).
Recommendation: Create a separate page for selecting the preferred language as shown in Figure 6 and add a language toggle in the top bar as shown in Figure 7. We also recommend to continue testing how respondents react to having the separate page for selecting the preferred language. For more information about the recommendation of creating a separate page for selecting preferred language see Lykke and García Trejo (2018).
Figure 6. The first page of the 2020 Census Planning Survey contains the language toggle. This design worked well for users. 
Source: 2020 Census Planning Survey (Lykke and García Trejo, 2018)

Figure 7. The login page of the 2016 Census Test with the language toggle in the top right corner of the blue menu bar with the black arrow pointing to it. 
Source: 2016 Census Test
4.1.4. Dashboard design leads to usability issues

After answering some basic questions about the household, the respondent is presented with a ‘dashboard’ which collects the names, ages and basic demographics of the children in the household. After the respondent enters the information about all the children in the household, the survey then selects only one child. Then, the respondent will be routed to one of three modules (depending on the age of the child) containing health-specific questions. When the respondent arrives at the dashboard, they are supposed to select ‘Start’ (see bottom left corner of Figure 8). Then, they will be able to enter the name, sex and age for the child(ren). If there are still more children to add when they come back to the dashboard, they are supposed to click the ‘Add Child’ button at the bottom of the screen. However, during the test sessions, several English-speaking participants incorrectly selected the button ‘Add a Child’ to begin, and one participant selected ‘Continue’ instead of selecting the ‘Start’ button. The eye tracking heat map in Figure 9 lends support to the idea that participants did not know where to look.

Three Spanish speakers also expressed confusion about what button to press and what to do next the first time they viewed the dashboard. One participant commented about whether ‘Comenzar’ [start] is a button that takes the user somewhere or performs an action. This participant tried to type directly below the ‘First Name, Initials, or Nickname’ column heading in Figure 8. Another Spanish speaker added a child by mistake as she was exploring the different buttons and clicking everywhere. On the final interaction with the dashboard, another participant was confused on what to do next or what button to press to continue. Overall, we found that there are too many buttons on this page.

While everyone ultimately selected ‘Start’ and then added their children from youngest to oldest, completing the task was difficult. Having such a difficult task in the beginning of the survey could be seen as a deterrent to completing the rest of the survey.

Several participants using a mobile device also commented that there was a lot of text on the screen as demonstrated in Figure 8, but PC participants seemed to read this text as shown in Figure 9.
Figure 8. Screenshot of the dashboard on a mobile phone. There are three buttons at the bottom of the screen. The button labeled 'Start', in reverse print on the bottom left of the screen, is the one a respondent is supposed to select to enter a child’s information. A bigger button labeled ‘Add a Child’ on the bottom of the screen, centered, is supposed to be used to enter information for additional children. The button below that one labeled ‘Continue’ is supposed to be used to move to the next screen, after information on ALL children in the household is entered. It was unclear to participants what they should do first.
Source: 2018 NSCH usability testing
Below we offer three recommendations in order of our preferences, with Recommendation 1 being the most ideal.

**Recommendation 1:** The current dashboard used for NSCH is not intuitive for users as demonstrated during the usability testing. To avoid usability challenges and confusion, we recommend following the 2020 Census design where the total number of people is requested on one screen and the next screen includes correct number of boxes to enter their names (see Figure 10). The NSCH could be designed similarly. First, collect the number of children in the household on one page. Then, on the next page, display that exact number of fields with instructions to list the names of the children in order from youngest to oldest. After that page, we recommend using a dashboard design with all the names listed and a ‘Start’ button for each name (see Figure 11).
Figure 10. The 2017 Census Test screen where the respondent can enter names. The same design is used in the 2020 Census. Source: 2017 Census Test

Figure 11. Dashboard for 2017 Census Test. The respondent selects start for a person, completes the data entry for that person, and then is taken back to the dashboard where the respondents selects another person and proceeds until he or she is finished. The same design is used in the 2020 Census. Source: 2017 Census Test
**Recommendation 2:** Keep the dashboard, but make the action ‘Start’ button more noticeable and bold or make this button more prominent in the instructions.

**Recommendation 3:** Consider showing a list of steps to make the instructions shorter.

**Recommendation about the Spanish translation issue on the buttons:** Both “Continuar” [continue] and ‘Comenzar’ [start] imply an action to start something in Spanish. Change ‘Comenzar’ [start] to ‘Inicio’ [start]. Change ‘Continuar’ [continue] to ‘Siguiente’ or ‘Próximo’. On the final dashboard shown, change the name of the button ‘Continuar’ to ‘Siguiente sección’ [next section].

4.1.5. Errors in entering information about the adults in the household

Toward the end of the survey there were questions about the child’s adult caregivers. There were separate questions for up to two adults. The questions for each caregiver were asked separately and labeled Adult 1 and Adult 2, with Adult 1 questions first, followed by the same questions for Adult 2. Participants were expected to understand who Adult 1 and Adult 2 were. However, this assumption led to user confusion. For example, during sessions, several English-speaking participants entered their information on the Adult 1 screen and then entered the same information on the Adult 2 screen (see Figure 12).

There were several reasons why these questions were confusing to participants. First, the label for the first screen in the ‘Adult 2’ section said, ‘About You- Adult 2’. Almost always, the participant entered his or her information in the Adult 1 section, so that label implied the participant had to enter the information again. Second, the instruction on this screen said, ‘Complete the following question for each of the two adults...’ This instruction was incorrect at this point in the survey because the respondent has already entered information for the first adult. The instruction also assumed the household is a two-parent family household, and this household type will not always be the case. In one session, there were three adults in the household, including the parent, and the grandparents and the participant who was the parent decided not to list either one of her parents at this point in the questionnaire.
Recommendation: Remove ‘About You’ in the label and use only the words ‘Adult 2’ in the label. Additionally, take advantage of the automated skips allowed in the online form, to ask ‘Is there another primary caregiver in this household?’ and then if ‘Yes’ is selected, enable the remaining questions. Consider allowing for more than two primary caregivers. Consider asking for the adult’s name and filling that name in the appropriate questions.

4.1.6. Language response choices are not sufficient

One question in the screener (see Figure 13) asked what languages are spoken in the household, with the choices being English, Spanish, and other. The other had an open text field where the participant could enter the language. In the online survey, the responses were designed as radio buttons, where only one selection could be made. Two Spanish-speaking participants were unable to enter a complete response for the additional languages that they spoke. This happened in the screener for T1 and T3. Both participants spoke two languages: English and Spanish. They both initially selected the radio button for Spanish, but also wanted to click on the ‘English’ option signifying that they use both languages at home. They tried also to type in English in the open text field; however, because the question used a radio button design, only one response was allowed. They should have selected the third choice and then used the open-text field to type that they both speak English and Spanish at home but none of them did that. These participants
did not understand how the radio button design worked for this specific question and therefore did not realize they were supposed to select the third option and then type in both the languages. Because the design was not usable for these participants, incomplete data was collected and users were frustrated. We did not observe this problem for the English-speakers participants because none of them spoke more than one language.

![Image of the screen with the languages spoken at home question. Notice the red ban icon in the grey field, indicating that the field is not accessible](Source: 2018 NSCH usability testing)

**Recommendation:** There are two recommendations for this issue. If the sponsor would like to gather information on multiple languages spoken at home, replace the radio buttons with checkboxes, which allow users to choose one or more options, and add an instruction of ‘Select all that apply.’ If the sponsor would like to gather information on only one language, consider adding additional instructions about the type of answer that is desired such as, ‘Seleccione sólo una respuesta’ [Select only one answer], or if you speak more than one language please select the language you speak most often at home.

We preferred the first recommendation because it seems the intent of the question is to collect multiple languages. If the design remains the same, we predict underreporting of languages spoken as not all respondents will understand how the radio button design operates, as we observed.
4.1.7. Security/verification questions are too long to display on phones and there are too many of them.

Soon after logging into the survey, the respondent is shown a PIN number, with the vague instruction of ‘make note of the PIN’ immediately preceding the PIN. Below the PIN, the respondent is supposed to choose three security questions to answer (see the first security question in Figure 14). Information collected on this screen would allow the user to re-enter and resume a partially completed survey in the event the respondent needs to exit the survey prematurely. Several English-speaking participants commented that some of the security question options were truncated on their phone display so they could not read the entire questions (see Figure 14). Several participants also commented that three security questions was overkill and not typical. Several participants commented that the security questions did not apply to them, that is, they did not have a pet (question is not shown in Figure 14) or a hero. Spanish speakers did not comment on the security questions and were able to select all three questions. However, three of the five Spanish speaking participants did not ‘make note of their PIN’ by writing it down. The eye tracking heat map in Figure 15 shows that English-speaking participants focused on the PIN and it also shows that three different security questions are asked.

Figure 14. Screenshot shows how the security question choices are cut off on the mobile phone, so much so, that the participant cannot guess what the question is for some of them. For example, ‘What was the last name of your third...’
Source: 2018 NSCH usability testing
Figure 15. Three participants’ eye tracking data on the PIN screen. Notice that participants focused on the PIN and then at each of the verification questions.

Source: 2018 NSCH usability testing

Recommendations: 1. Reduce the number of time the participant needs to select a security question. Currently, respondents need to select three different security questions to answer and we recommend that they should only need to select one security question because only one security question is needed to re-enter the survey. 2. Choose security question choices that fit on mobile phones. 3. Change the instructions about making note of the PIN to something direct like, ‘Write down this PIN’ and highlight those instructions. 4. Additionally, even though it was not a usability problem, we recommend using only one term, ‘Verification’ or ‘Security’ to describe the questions. Currently both terms are used on the screen.

4.1.8. Font color was inconsistent on the Spanish version of the questionnaire

In Question A39 (see Figure 16), much of the text was in blue, not grey as on the English version.

Recommendation: Change the blue color to grey on questions where there is disabled font to ensure consistency with the rest of the instrument and with the English version of the questionnaire.
4.1.9. Banner was inconsistent in Spanish version

A banner appeared on every web page of the PC and mobile versions of the survey. In English the same banner was used on every page on the PC design; this did not happen on the Spanish version. Figures 17 and 18 show that the banner changed across web pages in the Spanish instrument. Both types of banners appeared on various screens through the Spanish version of the survey. There was no consistency on when one banner was used over the other.
**Recommendation**: Keep the banner consistent throughout both the Spanish and English versions of the survey.

4.1.10. **Review screen was not used and Spanish translation on the screen is incorrect**

At the end of the survey, a review screen was offered. Using this screen, participants could have navigated back into the survey to change their answers. The screen (see Figure 19) had each major questionnaire section listed with a ‘+’ button to the left of the section description. To see the questions in that section, a respondent had to click or touch the ‘+’ button to open the section. After touching/clicking the ‘+’, a series of blue links for each question would appear (see the links under the section ‘Experiences with this child’s health care provider’ as an example in Figure 19). To the right of the blue link was not the survey answer but a green check mark with the word ‘Ok’ if the respondent answered the question. To actually see or change an answer, the participant would have to click on the blue question link which would navigate the respondent back to the question.

During testing, none of the English-speaking participants used the review screen to change their answers and many said they would not use it. One participant clicked on some of the ‘+’ buttons to open sections, but did not change anything. The participant commented that he/she wanted to see their answer next to the item and not the word ‘Ok’ as shown in Figure 19. Only one Spanish-speaker completed the entire survey (T2) and saw this page. When she received this page, she too did not make any changes, but wanted to submit her results. The submission button in the Spanish version had the wrong label. It said ‘Vuelva a la revision’ [Return to the revision] instead of ‘Submit’. This participant was confused and did not know how to finalize and submit her answers.

Figure 20 shows the heat map from three participants’ eye tracking data on the PC screen. It shows that participants did not read all the sections of the questionnaire.
Figure 19. The review screen on a PC with the arrow pointing to the ‘Ok’ button indicating that an answer was provided to the survey.

Source: 2018 NSCH usability testing
Recommendation: Because the review screen was not used by participants during testing, we recommend reviewing the paradata for this screen once the survey has been fielded to determine how many respondents make changes to their answers. If most respondents do not use this screen, consider deleting it. If it is used by respondents, to make it more usable, include the answers on this page instead of the word ‘Ok.’ On the Spanish version, change the name of the button to ‘Terminar’ [Finish] or ‘Enviar’ [Submit] or consider a combination of both ‘Terminar y Enviar’ [Finish and submit]. If two terms are used, we recommend that the change be implemented in both English and Spanish versions. Similarly, any changes to the English version should apply to the Spanish version.

4.1.11. Remind participants that the survey only asks questions about one child

The survey only collects detailed information about one child per household. Two English-speaking participants were surprised that the survey did not continue to ask about the other children in their households. One of the participants did not read/recall the information provided at the beginning of the topical module stating that only one child is selected for the survey. Figure
Figure 21 shows the arrow pointing to the statement. Although the eye tracking of that screen in Figure 22 shows that at least some of the participants read that information, it is easy to miss, especially on mobile devices.

Figure 21. The arrow points to the informational statement that only one child is selected in the survey. This statement comes at the beginning of the topical module.
Source: 2018 NSCH usability testing
Figure 22. The arrow points to the informational statement that only one child is selected in the survey. This figure includes three participants’ eye tracking shows that that the instruction was read by at least some of the participants.
Source: 2018 NSCH usability testing

**Recommendation:** At the thank you screen (see Figure 23 with the heat map), remind respondents that only one child was selected (if there is more than one child living in the household). If there is only one child in the household, then do not include that text. Keep the text at the beginning of the survey (Figure 21) that lets the respondent know questions will be about only one child; however, remove that instruction if there is only one child in the household (currently that text is there regardless of the number of children.)

Consider allowing respondents the ability to answer for another child if they would like.
4.1.12. Improve the visibility and labeling of the feature to exit and resume later and correct the Spanish translations of the headings.

When given the task midway through the session to stop answering the questionnaire and pretend that they would resume the next day, three of the 11 English-speaking participants did not log out because they did not know whether they could come back later. The word ‘Logout’ (Figure 24 shows the arrow pointing to the button) did not imply to them ‘Save and logout’.

Placement of the logout feature did not seem to be where participants expected it. When presented with the task, the first action most participants did was look at the bottom of the screen for something that looked like it could help them with the task. Next, many participants looked in the FAQs to determine if they could exit and resume later. However, the instructions for that task were not in FAQs, they were in the ‘Instructions’ link that appears at the top of the screen.

All five Spanish speaking participants needed help to logout from the instrument. Two participants closed the window instead of selecting ‘Logout’. One participant was expecting a ‘Save’ option before logging out. On the Spanish version of the survey, the headings are not in Spanish which makes the links useless for Spanish-only speakers.
On the re-entry page, there is an incorrect word in the Spanish instructions. It says ‘Name’ [nombre] instead of ‘Number’ [número] when referring to the ID number. This is one example of the numerous translations issues with the survey. Although this behavior was not observed during this testing, a Spanish-speaking respondent who is attempting to log back to the instrument and reads ‘Name’ instead of ‘number’ on the log in instructions will not know what information to enter. This could block Spanish speakers from logging back into the instrument.

**Recommendation**: Use the phrase ‘Save and Logout’ or ‘Guardar y salir’ as the button label, put all the content in the instructions into the FAQs button and, eliminate the button called ‘Instructions’. On the Spanish re-entry page, replace ‘Nombre’ with ‘Número’. Make sure the ‘Logout’ button label (and all labels) are translated into Spanish when the survey switches to Spanish. Perform a copy edit process to improve the quality of the instrument.

4.1.13. Progress indicator was rarely used by participants and was not a standard design

There was a progress indicator at the bottom of the survey when accessed on a PC as shown in Figure 25. There was no such design on the mobile; however, one participant commented that she would like to see a progress indicator at the bottom of the mobile screen. The only time the progress indicator was used was when one participant, who got back into the survey to make sure she had completed it, tried to click on the progress indicator to go someplace, but it did not take her anywhere. Eye tracking across the pages in the previous and subsequent figures does not indicate that the participants looked at the progress indicator much, if at all.
Figure 25. Progress indicator appeared at the bottom of the screen on PCs below the next and previous buttons. The colors attempted to indicate where the respondent was in the survey: bright blue was topics answered; light blue was the current topic section being answered; grey was the topics to answer. Source: 2018 NSCH usability testing

**Recommendation:** There are mixed results with progress indicators (Villar, Callegaro, & Yang, 2013). It might be good to include a split panel test of using a more typical progress indicator (a horizontal line as shown in Figure 26) compared with a design without a progress indicator.

![Figure 26](https://example.com/figure26.png)

**Figure 26.** Example of a progress indicator near the next and previous buttons.

4.1.14. Participants answered questions incorrectly with screens that scrolled in two Spanish cases

While scrolling down a page, two Spanish-speaking participants using a laptop had difficulty remembering what the question was asking about and provided incorrect answers. These screens had a stem question and multiple follow-up questions. Participants read the stem question, as this was the first thing they saw on the screen, but then needed to scroll down to keep answering all the questions on the screen. Since the content of the screen was long, as participants scrolled down they lost sight of the question. Participants kept answering but forgot what the question was about. This was noticeable in the think-aloud process because they were providing incorrect responses while at the same time trying to do multiple tasks (e.g. remembering the original question at the top of the screen, scrolling down, reading and selecting the response options). In one case, a participant was unable to scroll back to the question located at the top of the screen and started providing wrong responses (Question A4 in topical module T1). This finding was not observed in the English-speaking sessions.

**Recommendation:** Limit screen scrolling by reducing the number of questions on a page. Repeat the question in the middle part of the sub-sections to remind respondents about the questions they have already been asked. Other solutions such as freezing a pane on the screen would be ideal; however, we are unaware of this ability in a web form.
4.1.15. Field or text placement impedes answering questions correctly

We observed participants missing questions and entering incorrect data because of the field placement in English and the text placement in Spanish. The four instances are below.

Other specify
For several Yes/No questions, more detailed information is sought with a ‘Yes’ response. For example, the last question in Figure 27 asks about mental health support, including hard-coded responses such as peer/support groups. When respondents get to the last question, the ‘Other’ question, they are asked to specify the type of (support) in an open text box if they respond ‘Yes.’ On mobile phones, the specify box is between the ‘Yes’ and ‘No’ response choices (see Figure 27). One participant commented that it was an awkward placement on mobile and we observed several participants not answering the ‘other’ question (with either a ‘Yes’ or a ‘No’) on mobile phones perhaps because of the ‘specify’ field placement came directly in-between those answer fields.

Figure 27. Screenshot shows the specify box between the ‘Yes’ and the ‘No’ response options
Source: 2018 NSCH usability testing
**Recommendation**: Because there is item nonresponse (neither ‘Yes’ or ‘No’ are selected) to these type of questions, reorder the response options on questions E2g and H11 so that ‘Yes’ and ‘No’ follow each other, like the other response choices on the page. Place the specify field below both choices as shown in Figure 28. You might also want to add ‘If yes, please specify’ as the label instruction for that field. Initially, the specify field and label should be disabled and then only enabled if ‘Yes’ is selected. If ‘No’ is selected, the label and box remains disabled.

![Mockup of recommended reordering of response fields](image)

*Figure 28. Mockup of recommended reordering of response fields. The label should be disabled as well until ‘Yes’ is selected and then the label and the field should become enabled.*

*Source: 2018 NSCH usability testing*

**Years and months**

On the age question (see Figure 29), participants are given the option of recording age in terms of years OR months. This is to accommodate respondents with both older children (3 and up) whose age is generally described in terms of years, and younger children whose age is often given in terms of months. Four participants entered years and also entered months. The participants should have only entered data in one field, not both. When both are selected, an error message is triggered, and then both fields are cleared. In one case, the participant moved forward without re-entering a response. We notice in the heat map in Figure 30 that participants focused on both of the fields. Spanish participants did not have difficulty with this screen.
Figure 29. Screenshot of age question with years and months on the same horizontal line.
Source: 2018 NSCH usability testing
Figure 30. Three participants' eye tracking on the age screen. Notice that participants looked at both fields, when they really only needed to answer the years field.
Source: 2018 NSCH usability testing

**Recommendation:** We recommend testing an alternative design where the label ‘Age in Years’ is above the field. Next, the italicized instruction for age in months would appear below the field mentioned above followed by the label ‘Age in Months’ (see Figure 31). Another alternative is to disable the italicized instruction and the ‘age in months’ field and only enable it if the respondent selects “0” in the first field.
**Question C5 – Height - feet and meters**

For the height question shown in Figure 32, the respondent is supposed to enter feet and inches, or meters and centimeters, but not both. One participant wanted to enter 5 feet 6 inches. She entered 5 into the feet field, but then entered the 6 in the meter’s field. This action then removed the 5 in the feet field. She did not realize her data had erased and she continued with the survey, leaving the item blank. Eye tracking, shown in Figure 33, also suggests that participants looked at all of the fields, not just those that were appropriate.
Figure 32. Screenshot of height question showing that respondents can answer either in feet or in meters. Source: 2018 NSCH usability testing

Figure 33. Three participants’ eye tracking on the height and weight screen. Notice that participants looked at feet/inches and meters/centimeters. Source: 2018 NSCH usability testing
Below we offer three recommendations.

**Recommendation 1:** Consider adding an edit instead of automatically erasing the data when people fill out both fields.

**Recommendation 2:** Another alternative is to have only two fields, but have the user be able to select the dimensions (consider either a radio button with feet and inches as the default) or a dropdown.

**Recommendation 3:** The third alternative is if the user enters data into the feet or inches field then meters and centimeters disables and vice versa.

*Placement of ‘OR’ or ‘o’ in Spanish on Topical Module T1/Questions B2, B5, B6, etc.*

Two Spanish-speaking participants did not realize that they only need to enter information in one field because of the placement of the word ‘o’. In Spanish, the word ‘o’ (the translation for ‘or’) is left justified on both PC and mobile and looks more like a bullet in the online Spanish version (see the PC version in Figure 34).

![Figure 34. Placement of 'o' is left justified making it look like a bullet instead of a word. Source: 2018 NSCH usability testing](image)

**Recommendation:** Align the word ‘OR’ or ‘O’ to the open field boxes like in the paper-based questionnaire. This recommendation applies to all modules (T1, T2, T3) of the questionnaire and all questions that include the two options with the word ‘OR’ in between as shown in Figure 35.
4.1.16. Low priority usability issues

*Instructions are in the wrong place on Question A22.*

While test participants did not have any usability issues or input incorrect answers, the instruction ‘Examples of educators are teachers and school nurses’ (see the arrow in Figure 36) is in the wrong place and should be below the question stem. Eye tracking shows no unusual patterns in Figure 37.
Figure 36. Question with example in the wrong place (denoted by the arrow). It should be below the question stem.  
Source: 2018 NSCH usability testing
Figure 37. Heat map of question with example in the wrong place (denoted by the arrow). It should be below the question stem.
Source: 2018 NSCH usability testing
The blood disorders question has too much greying out on one screen on mobile phones

Several participants commented that there was too much grey on the blood disorders question when they answered on their mobile phone. On mobile phones, the screen can end up looking like everything is grey if they answer the earlier filter question with ‘no’ that they do not have a blood disorder (see Figure 38). Consider breaking that question up into separate screens.

Figure 38. Screenshot of the blood disorders questions when the first question (not shown) is answered a ‘No’ and then the subsequent questions do not enable. When scrolling down to the next button there are so many disabled questions that the entire screen can show all grey questions at one point.
Source: 2018 NSCH usability testing
Use a keypad for numeric entries

The QWERTY keyboard comes up on mobile phones even when the answer must be a number. While we observed no participants answering incorrectly, the type of keyboard presented to the respondent communicates the type of answer requested. For questions requiring a number answer such as Login ID, telephone number, income, ZIP code, etc., using a keypad is better than using a QWERTY keyboard (see Figure 39). Other surveys, such as the National Survey of College Graduates, the American Community Survey and the 2016 Census Test all have adopted the keypad design.

Figure 39. Example question with QWERTY keyboard coming up even when answer must be a number.
Source: 2018 NSCH usability testing
Match the field length to the estimated response length

For Questions K1 and K2, the field length is larger on mobile phones than the expected answer length. The field length should match the two digit length shown on the paper form as shown in Figure 40.

Figure 40. Paper form design showing that the field length matches the expected value.
Source: 2018 NSCH usability testing draft paper questionnaire

Consider a year dropdown field instead of an open-text field

Question J5 ‘When did you come to live in the United States?’ is asking for a calendar year response, but one participant initially entered 20 years because she misinterpreted the question. Dropdowns have their own limitations and can take longer to answer, but using a dropdown in this question instead of an open-text field would eliminate any misunderstanding of the question intent.
Take advantage of the automation

Many of our participants answered questions about healthy children. Some questions used double-barrel response choices to capture whether the child had a health issue or was healthy as shown by the example in Figure 41 where the first response choice should be chosen if the child does not have a condition. The additional effort participants with healthy children spent on these types of questions added to the length of the survey. We recommend using a filter question for these sections to first ask is the child has any health condition, and then if not, automatically skip questions within the survey, instead of using the double-barrel response choices. If the sponsor is concerned about not identifying all health issues, then we recommend monitoring how often questions are answered consistently—that is a child is always identified as healthy at all questions; and how often questions are answered inconsistently—in some questions, the child is identified as having a health issue and at other questions they are not.

**Figure 41. Example of a double-barrel question with the answer choice, “This child does not have any health conditions.” (This screen shot is from the paper form, but the same design was used online.)**

Source: 2018 NSCH usability testing draft paper questionnaire
On the income question (Figure 42), if the participant had income in any of the categories, he or she selected ‘Yes’ and then the income field for that category became enabled and the participant entered the correct amount. The problem was the total amount did not automatically sum all the individual amounts. One participant commented that the income sources should sum automatically on the income question screen’s web page. This is a user expectation. The eye tracking heat map of the page shows that participants spent time looking at the total income (Figure 43). Had this been an automated sum, participants most likely would not have spent as much time looking at that field. The fact that the field did not autosum increased respondent burden.

Figure 42. Income screen where the arrow points to the total income field
Source: 2018 NSCH usability testing
Figure 43. Three participants’ eye tracking data on the income screen. The incomes did not sum automatically and there are fixations near the final response box (the green heat map) indicating that the participants looked there. Source: 2018 NSCH usability testing
4.2. Efficiency
We experienced a much longer survey completion time than expected. In terms of the English
speakers, a few participants commented on the length of the survey. Only one English speaker
took less than the estimated 30 minutes to complete the survey on the laptop and the quickest
mobile phone user took 35 minutes to complete the survey. The other five mobile phone users
took one hour to an hour and a half to complete the survey. English speaking participants who
completed the survey on the laptop took between 35 and 50 minutes.

All five Spanish speakers spontaneously mentioned that the survey was long. Only one of the
five Spanish speakers finished the survey within the 90 minute time frame of the test session.

**Recommendation**: During production, keep track of the amount of time needed to complete the
survey on different devices and if the average time by device exceeds 30 minutes, readjust the
OMB hours or eliminate questions.

4.3. Satisfaction
Participants rated their satisfaction with the online survey immediately after completing the
NSCH. Spanish speakers filled out a paper satisfaction questionnaire and English speakers
completed the questionnaire online. This paper questionnaire contained 12 different measures
of satisfaction for which participants rated their subjective experience on a 9-point Likert scale
with only the endpoints labeled. The online questionnaire contained the same 12 different
measures of satisfaction, but with a 7-point Likert scale. For the purposes of this report, this
section includes only the comprehensive measure of satisfaction: overall reaction to the web
survey. The results of the other satisfaction measures can be found in Attachment C, which report
ratings on more specific criteria (e.g. organization of text).

Figures 44 and 45 display satisfaction reports for participants’ overall reaction to the web survey
for Spanish and for English. All responses were on the positive end of the scale, with most
participants reporting the highest possible level of satisfaction for both the Spanish and English
versions.
Figure 44. Satisfaction scores for the overall reaction of Spanish speakers to the Spanish web survey by device type
Source: 2018 NSCH usability testing

Figure 45. Satisfaction scores for the overall reaction of English speakers to the English web survey by device type (the JAWS user did not complete the satisfaction questionnaire)
Source: 2018 NSCH usability testing
5. Accessibility Findings

In our accessibility testing, we determine if a person with vision impairment, who needs a screen reader, can use the system. In this survey, form validation, that is when the survey provides feedback on the answers in the form of edit messages, functioned correctly because the error messages were directly above the questions with missing data. Screens with gray text functioned correctly without affecting previous/next navigation as seen in other Web survey applications tested in the past. The main issues encountered are summarized here with recommendations.

The JAWS’ user had difficulty with some of the stem and leaf type questions, specifically, she had trouble remembering what the original question was asking when rating frequency or level of agreement. This problem was more prominent in questions with many leaves (sub-questions) and in the later sub-questions.

**Screens that allow multiple inputs for units**

During testing, our sole JAWS participant entered height or weight values in U.S. customary units (see Figure 32 for the screen) then navigated to the next question by keyboard commands. The tab sequence took the participant through the metric units fields, which she left blank because she had answered the English units. However, when tabbing through the metric unit fields, the English units data was erased without any notification. The participant never corrected her response. Users do not expect the same question to be asked again in different units after just responding, that is, asking for height and/or weight in metric units after responding with English units, or answering in years and then having another question appear asking for months.

Figures 46 and 47 show other examples of this design. For example, in Figure 46, the tab sequence takes the respondent from the age in years to the age in months. After entering age in years, screen-reader users do not expect to enter the age in months. In Figure 47, after entering the weight in pounds and ounces, the respondent is asked to enter the weight in kilos and grams because of the tab sequence. When the keyboard focus is placed inside the kilograms field, pressing the down arrow once will delete the entries in the pounds and ounces fields and the user will never know this has occurred.
Figure 46. Age in years and months
Source: 2018 NSCH usability testing

Figure 47. Weight in U.S. customary units and metric units
Source: 2018 NSCH usability testing
**Recommendation:** The alternative unit field should be disabled once an entry is made into the unit of choice. For example, in Figure 47, make the kilograms and grams fields unavailable if pounds and ounces are already entered. This recommendation was made earlier in the report.

**Awkward labels for units**

The JAWS participant remarked that labels for height, weight, months, and years were each announced using both the full word and an abbreviation which the participant said was confusing. The JAWS screen reader reads the title text from the source code for the web page. Title text never appears on the screen, but it is detected by JAWS. It is not the same as the label. Title text further defines the input needed. Title text comes before the field needing input. After the title text comes the word, “edit.” Edit is a cue to JAWS users to enter their response. In other words, when JAWS users hear “edit” they know they should type something. After the word “edit” comes the label. There is always a label on fields, but fields do not always need title text. The question in the red box shown in Figure 48 has several labeling issues:

- When using the tab key, for the pound field the JAWS user hears ‘pounds lbs edit pounds’ where the title text is ‘pounds lbs’ and the label is the ‘pounds’ after the word ‘edit.’
- JAWS users also hear ‘ounces oz edit ounces’, ‘kilograms kg edit kilograms’, and ‘grams g edit grams’.

Questions C4 and C5 shown in Figure 32 have similar issues. When using the tab key for the meters field, the JAWS user hears ‘meters m edit meters’ instead of just ‘meters edit meters.’ Similarly, the JAWS user hears ‘centimeters cm edit centimeters’ instead of just ‘centimeters edit centimeters.’

Questions C30 and C33 shown in Figure 48 also have a similar issue. JAWS users hear ‘Month Mos edit Months’ instead of just ‘Months edit Months’. During testing on mobile devices, TalkBack users users hear ‘Yrs’ and ‘Mos’ for labels instead of ‘Years’ and ‘Months.’
Recommendation: Use ‘pounds’ for the title text and ‘pounds’ for the label, so a JAWS user would hear ‘pounds edit pounds’ and similarly for the other examples.

Out of sequence tab orders
Problem 1: On the first screen of the survey there is a tab order problem. Once respondents navigate to the login fields, they will not proceed further. Screen-reader users (JAWS and TalkBack) will not read any of the text below the sign-in button (in the red box in Figure 49) and as an example of this issue, the participant who used a screen reader did not read that text.
**Recommendation:** Change the tab order of the page so that the text is read first, then the login ID and then the sign in button.

Problem 2: On the income question shown in Figure 50, a series of income sources are listed and under each source (shown left to right) is the Yes/No response choice of to indicate if the respondent had that income, then the amount box, and then a checkbox if the amount was a loss. If respondents have a negative income, they are supposed to select the radio button for ‘Yes’, enter the negative income (without the minus sign) and check the ‘Loss’ box. However, the ‘Loss’ checkbox can be selected even if the ‘No’ option is selected.

**Recommendation:** The ‘Loss’ checkbox should be unavailable if the ‘No’ response choice is selected.
Incorrect instructions, labels and misspellings

Problem 1: The instructions on the review screen for both PC (shown in Figure 51) and mobile are incorrect. The instruction references links, but none of the controls on this page are defined as links. The instruction to select a link to make a correction is not possible unless the ‘+’ sign is pressed first to reveal the list of links for a topic on a completed page.

![Figure 51. Review screen with incorrect instructions. Source: 2018 NSCH usability testing](image)

Recommendation: Delete the sentence: ‘Listed below are links to completed pages.’ Replace it with ‘To review your responses for each topic, click on the “+” sign.’
Problem 2: On the screener dashboard (see the red box on Figure 52), when the screen reader reads the label for deleting a child, it reads aloud redundant text. It says ‘delete a child button button.’

Figure 52. Screener dashboard where the red box indicates the element with the incorrect label
Source: 2018 NSCH usability testing

Recommendation: The screen reader detects the delete button as a button and inserts the word itself, so the screen reader label should just be ‘delete a child’.

Problem 3: Question I13 contains a misspelling of the word ‘severely’. Question A7 contains a misspelling of the word ‘decisions’.

Problem 4: On the address verification, the address is difficult to understand because of the abbreviations of Court (CT), Suite (STE), and all other address abbreviations.

Recommendation: Correct labels, misspellings and expand all abbreviations to words.
In the desktop version of Question B2 requesting weight in pounds, see Figure 53, source code appears in the yellow box over the field. That yellow box is called a tool tip.

**Recommendation:** Reprogram the tool tip so only the label appears.

### 6. Cognitive Findings and Recommendations

This section includes comments about the new questions added to the 2018 NSCH. Participants made these comments in response to probes during the structured debriefing and they made spontaneous comments while completing the survey. The questions in English and in Spanish are available in the Appendices.

#### 6.1. Debriefing responses

During the debriefing, we focused on the new questions. The comments made in this section were in response to probing questions. We highlight each new question with a summary of the participants’ understanding of the question and the eye-tracking heat map, if eye-tracking was captured on that screen.

**Blood disorders**

This question was a Yes/No question (see Figure 54) asking whether the child had any blood disorder disease diagnosed by a health care provider. If ‘Yes,’ there were subsequent questions about how it was diagnosed and the specific disease. Initially, on the screen only the filter question was enabled. If the participant answered ‘Yes’, then all the grey text (the disabled text) became bold and enabled so the participant could record his or her answer to each of the subsequent questions. There were no cognitive difficulties found with this question other than
many English-speaking participants did not know what Thalassemia was. No one answered ‘Yes’ in either English or Spanish. Participants made spontaneous comments about the amount of disabled text, as described in Section 4.1.16. Eye tracking (see Figure 55) indicates that participants read the stem questions and glanced through the disabled text.

Figure 54. Blood disorder question
Source: 2018 NSCH usability testing
**Cystic Fibrosis**

The cystic fibrosis question was also a Yes/No question asking whether the child has been diagnosed with cystic fibrosis by a health care provider as shown in Figure 56. No one answered ‘Yes’ in either English or Spanish. There were no cognitive difficulties found with this question across languages. Most English-speaking participants had heard of cystic fibrosis. Most Spanish speakers had not heard this term, however they correctly answered ‘No’ if they did not know the meaning, as a caregiver would certainly know the meaning if his or her child had the condition.

---

**Figure 55. Heat map using eye tracking data of the blood disorders screen**

*Source: 2018 NSCH usability testing*
The eye tracking data (shown in Figure 57) suggests that participants read the stem and glanced through the disabled text.

Figure 56. Cystic fibrosis question
Source: 2018 NSCH usability testing
**Other genetic condition**

The final question in this series was a catchall question about any other genetic or inherited condition. It too was a Yes/No question with subsequent questions if the answer to the filter question was ‘Yes.’ While there were no cognitive difficulties with this question (shown in Figure 58), there was one case of measurement error. One participant said that her child could be a carrier of a serious condition. She did not know for sure, but answered the question affirmatively, which implied that the child has this condition when she does not. The addition of a ‘Don’t know’ response option with a write-in field could possibly be an addition to this question to get at these nuanced situations. Eye tracking data in Figure 59 shows that most of the focus is on the question stem.
Figure 58. Other genetic or inherited condition screen
Source: 2018 NSCH usability testing

Figure 59. Heat map using eye tracking data of the other genetic disorders screen
Source: 2018 NSCH usability testing
**Bullying**

The bullying questions appear in Figure 60. Those questions asked about whether the child had been bullied and if so, how frequently, and if the child was a bully and if so, how frequently.

During the usability sessions, these questions appeared to produce a lot of measurement error. Only one participant described a bullying situation involving her child and people outside of the family where the episode was so traumatic that the participant declined to describe it. Other participants also indicated that their children were bullied, but they seemed to focus on episodes of their child being excluded during typical play situations. When asked whether their child bullies anyone, participants focused on their child bullying siblings. Depending upon the interests of the sponsor, these could be false positive answers. We recommend reviewing the bulling questions in the 2017 School Crime Supplement of the National Crime Victimization Survey to see if that set of questions would better measure school bullying.

Results of the eye-tracking data for that screen is found in Figure 61. The bullying questions are located at the bottom of a screen about other childhood behaviors and the eye tracking data does not identify anything unusual in the reading pattern associated with the questions.
Figure 61. Heat map using eye tracking data of the screen with the bullying questions (n=2)
Source: 2018 NSCH usability testing
Obesity

The obesity questions (see Figure 62) first asked if the adult caregiver was concerned over the child’s weight and then asked if a health care provider had told the caregiver that the child is overweight. There were no cognitive difficulties with this series of questions, but it seemed like participants decided how to answer after reading both questions. In one instance, the second question might have influence the answer to the first question: the father said that his daughter falls within the typical limits, but he thinks that she is slightly chubby for his culture. He responded “No, I am not concerned.” to the first question even though he expressed some concern verbally. Eye-tracking data in Figure 63 shows no unusual patterns. Other participants, both English and Spanish, did not have any comments or concerns with these questions so we have no recommendations based on this testing.

Figure 62. Weight questions
Source: 2018 NSCH usability testing
**Overnight stays at the hospital**

This question asked whether the child was admitted to the hospital and stayed overnight in the past 12 months as shown in Figure 64. We did not observe any cognitive difficulties with this question and parents appeared to remember how often their child had been to the hospital. Eye tracking shown in Figure 65 reveals participants focused on the text, “At least one night.”
Medical History

This question asked whether the respondent and the child received a copy of the medical history. We did not observe any cognitive difficulties with the medical history question in English (see Figure 66). When asked to think aloud while answering this question, participants mentioned brochures on weight control, immunization records for school, etc. We do not have eye tracking results on this question because we did not eye track the English-speaking participant who received this question on the laptop. Participants did not know the meaning of “medical history” when translated in Spanish as “resumen medico.” Further testing in Spanish is necessary for this term.
**Transition to adulthood**

This question asked about conversations had with health care providers about older children transitioning to other doctors. While we did not observe participants having any cognitive difficulties while answering the transition to adulthood question (shown in Figure 67), the researchers observed that it was not really appropriate for younger children (for example, a 12-year-old) in the T3 module because seeing doctors who treat adults is still six or more years away. None of our participants answered ‘Yes’ to this question in the T3 module. We did not collect eye tracking on this screen.

![Figure 67. Transition to adulthood question](source: 2018 NSCH usability testing)

**Speech and language milestones**

In the T1 module, an entire screen contained 11 questions about whether the child had made particular milestones with speech and language (see Figure 68). The participant had to answer ‘Yes’ or ‘No’ to each question. There were no cognitive difficulties with the speech and language question in either English or Spanish. One participant said the list sounded similar to what her doctor asked her at checkups. One participant spent more time answering the question of whether the child can tell a story with a beginning, middle and end. The eye tracking data shown in Figure 69 confirms this. We do not believe this is a concern.
Figure 68. Speech and language milestone questions
Source: 2018 NSCH usability testing
Figure 69. Heat map using eye tracking data of the speech and language milestone questions screen (n=1)
Source: 2018 NSCH usability testing
**Screen time**

Participants spent time answering the screen-time question (see Figure 70). This question asked how much time the child spends on electronics, outside of school work, on any given weekday. English-speaking participants appeared to be able to separate screen time their child spent on homework from screen time for enjoyment. Participants reported older children as having much more screen time, with participants explaining that their kids watch TV and use their phones at the same time. These participants did not double count the number of hours when it was simultaneous. Two Spanish-speaking participants had issues calculating screen time for their children. One participant explained that it was difficult for her to calculate because the screen time was not continuous. For example, her child picked up her phone and used it and then got bored. Another participant had trouble calculating the screen time because she was unsure whether FaceTime counts as screen time and this is the way this respondent communicates with her daughter.

The eye tracking data in Figure 71 show a typical pattern of behavior, but that time was spent on this question in comparison to other questions. We have no recommendations for improving this question or response choices.

---

*Figure 70. Screen-time question*  
*Source: 2018 NSCH usability testing*
Figure 71. Heat map using eye tracking data of the screen-time question screen. The screen-time question is at the top of the page (n=3)  
Source: 2018 NSCH usability testing

Plan of care
Participants showed some confusion with the plan of care series of questions, shown in Figure 72. The question attempted to gain information on whether there was a plan of care as the child matures. Several participants answered ‘Yes’ to the first two questions, indicating that there was a plan of care and they had access to it, but in each instance, the plan was either an oral plan or a referral and not something with tasks and goals and a schedule. All participants answered ‘No’ to the third question about whether the plan addresses transitions to doctors who treat adults. We did not collect eye tracking on this screen. Spanish speakers did not know the meaning of ‘plan of care’ as translated in Spanish as ‘plan de cuidado.’ We recommend that the term ‘plan of care’ be changed to ‘instrucciones a seguir’ and be further tested.
6.2. Spontaneous comments

While participants completed the survey, they were instructed to think aloud. A few English-speaking participants commented on the length of the survey. All five Spanish speakers spontaneously mentioned that the survey was long. There could be several factors contributing to these comments.

Participants with typically developing, healthy children received a lot of questions that perhaps did not need to be asked. For example, questions about coordination among doctors and the amount of care given did not seem appropriate for healthy children. For example, if a child did not see more than one health care provider, it is not apparent why the survey asks whether it was difficult to coordinate the care. There is another question about how many hours a week the provider spends administering care, but for children without any ongoing conditions, it seems
like the lack of skip patterns in the survey risks generating false positive data. Several of our participants reported spending an hour a week giving vitamins or bandages to their children. Including an earlier filter question that addresses things like how many doctors the child saw last year (excluding dentists), could be used to skip respondents appropriately.

Two Spanish speakers wondered aloud about the presence of the greyed out questions. However, they realized there was a response pattern and that questions would be activated only if they answered ‘Yes’ on specific questions. It could be that being able to see the extra, greyed out questions led to the perception of a long survey.

**English-specific comments:**

*Question D5 is cognitively difficult.*

This question asked ‘During the past 12 months, did this child need any decisions to be made regarding his or her health care, such as whether to get prescriptions, referrals, or procedures?’ At least two participants did not know what the question meant. One participant reported ‘Yes,’ saying that a prescription was filled, but it is not clear whether that is the intent of the question. We recommend more cognitive testing of this question.

*A response option in Question E4 needs a different reference person.*

This question asked about what type of health insurance covers the child. The question was, ‘Is this child CURRENTLY covered by any of the following types of health insurance or health coverage plans?’ It was a ‘Yes/No’ forced-choice question design with these categories: (a) Insurance through a current or former employer or union; (b) Insurance purchased directly from an insurance company; (c) Medicaid etc.; (d) TRICARE or other military health care; (e) Indian Health Service; Other. The respondent is supposed to select ‘Yes’ or ‘No’ for each of those insurance types. One participant commented that the option that says, ‘Insurance through a current or former employer or union’ implies that it is the child’s current or former employer, and that is not correct. It should be the parent’s employer. We recommend this option changing to, ‘Insurance through a parent or guardian’s current or former…’ because the child is not the person working.

**Module T3: Question C9, C11 and C12 have a skip error.**

Question C9 asked if there was a place the respondent took the child when he or she was sick. If the answer was no, then the next question (C11) was if there is a regular place where the respondent took the child for well-child or preventative care. If the answer was ‘Yes,’ then the following question (C12) asked whether that place was the same place as the place visited when the child was sick. That question does not make sense if there was no place where the respondent takes the sick child. We recommend skipping Question C12 if C9 is answered ‘No.’

*Question I5 needs a reference home.*

This question asked about mold in the home in the last 12 months, but for people who moved within the last 12 months, it was not clear which home they should consider. One participant had this issue. The home she moved out of had mold; but her new home did not. She answered
for the most recent home. We recommend adding an instruction specifying to report for your current home in the reference year.

*Question A24 has unfamiliar terminology.*
In all modules, there was a series of questions about whether the respondent had ever been told by a doctor or health care provider that the child had particular health issues. In the T2 and T3 module, Question A24 asked whether the respondent had been told that the child has ‘Substance use disorder?’ Two participants said they had heard of substance abuse, but not substance use disorder. Everyone appeared to answer correctly, so we have no recommendations based on this testing.

*Question G5 in T2 and Question I6 in T1 have questions not relevant to small children.*
Question G5 in module T2 asked whether the child has any regular paid work, such as babysitting or cutting grass. The age range for this module is 6 to 11 year olds. For children at the younger end of this age range, there typically is not any paid or volunteer work. Similarly, in Module T1, Question I6 asks, ‘When your family faces problems, how often are you likely to do each of the following? (a) Talk together about what to do; (b) Work together to solve our problems; (c) Know we have strengths to draw on; and (d) Stay hopeful even in difficult times.’ The reference to family was confusing for young families because these parents do not hold discussions with their very young children. At this time, we recommend more monitoring of the questions.

*Spanish specific comments:*
*Race and ethnicity question*
One participant did not know exactly how to answer the race and ethnicity question in the screener. The wording of the question is ‘Es este(a) niño(a) de origen Hispano, Latino o Espanol?’ The response options include ‘No, no es de origen hispano, latino o espanol’, ‘Si, mexicano(a), mexicano(a) Americano(a), chicano(a)’, ‘Si, puertorriqueno(a)’, ‘Si, cubano(a)’, ‘Si, de otro origen hispano, latino, o espanol’. The participant was confused about how to answer the Hispanic origin question because she was expecting the response options to be ‘Yes’ or ‘No’ and none of the ‘Yes’ options specified ‘Bolivian.’ After reading the remaining options, the participant selected the option for another Hispanic origin. We recommend a consultation with analysts in the Census Bureau’s Population Division to make sure the survey is using the standard question version to measure race/ethnicity. Figures 73 and 74 show the standard ethnicity and race questions in Spanish on the American Community Survey.
NOTA: Por favor, conteste la Pregunta 5 sobre origen hispano Y la Pregunta 6 sobre raza. Para esta encuesta, origen hispano no es una raza.

¿Es la Persona 2 de origen hispano, latino o español?

☐ No, no es de origen hispano, latino o español
☐ Sí, mexicano, mexicano americano, chicano
☐ Sí, puertorriqueño
☐ Sí, cubano
☐ Sí, otro origen hispano, latino o español – Escriba el origen en letra de molde, por ejemplo, argentino, colombiano, dominicano, nicaragüense, salvadoreño, español, etc.

¿Cuál es la raza de la Persona 2? Marque (X) una o más casillas.

☐ Blanca
☐ Negra o africana americana
☐ India americana o nativa de Alaska – Escriba en letra de molde el nombre de la tribu en la cual está inscrita o la tribu principal.

☐ India asiática
☐ China
☐ Filipina
☐ Japonesa
☐ Coreana
☐ Vietnamita
☐ Nativa de Hawaií
☐ Guameña o Chamorro
☐ Samoana
☐ Otra asiática – Escriba la raza en letra de molde, por ejemplo, hmong, laosiana, tailandesa, paquistaní, camboyana, etc.

☐ Alguna otra raza – Escriba la raza en letra de molde.
**Question A25**
The introduction to Question A25 is confusing. The introduction in T1 and T2 reads ‘Alguna vez un medico o un educador le ha dicho a usted que este(a) nino(a) padece de ....Algunos ejemplos de educadores son maestros(as) y enfermeros(as) escolares.’ One participant was confused by the examples because in Spanish it is just a repetition of the question. We recommend removing the examples.

The words ‘conducta’ and ‘comportamiento’ are used interchangeably in Module T1/Question C8-Figure 75.
The question in the paper version of the instrument reads ‘Durante los ultimos 12 meses, le preguntaron los medicos o proveedores de atencion medica de este(a) nino(a) si usted estaba preocupado(a) por el aprendizaje, el desarrollo o la conducta de este(a) nino(a)?’. The word ‘comportamiento’ [behavior], as it appears in the online instrument, is closer to the meaning in the English instrument. We recommend choosing ‘comportamiento’ as it is the closest term to the English wording.

![Image](image_url)

**Figure 45. Question 8 in the T1 module in Spanish**
*Source: 2018 NSCH usability testing*

**Mode differences in Spanish paper and online forms**
The usability testing team for the Spanish cases found wording differences in at least six questions between paper and online forms. The mode differences were reported promptly and the NSCH survey team addressed them properly.
7. Limitations

Due to time limitations we were not able to recruit participants with many of the health issues that are included in the questions. Thus, not every question was tested. We did not prepare vignettes ahead of time for every question as these vignettes would be difficult to administer if the participant did not have direct knowledge of the condition.

8. Future usability research on designing for branching logic

This survey used disable and enable branching logic, where questions are grey until a filter question is answered in such a way to make them applicable, and then they turn black and can be answered. This design allows skip sequences to be on the same page and it allows the respondent to see all the questions to help them shape their understanding of the filter question.

We did not experience any usability issues with this design in the English testing; nor did we experience any usability issues when we tested this design with the National Sample Survey of Registered Nurses (Nichols, Kephart, and Malakhoff, 2018). However, the disabled text does make the questionnaire look longer than it would be for some respondents. In the case of the Spanish testing, participants commented often about the length of the survey and we do not know up to what extent the branching design contributed to this.

Further testing about the relationship between grey out areas and length of the instrument is necessary before reaching out to conclusions on this topic. It is unclear if this design adds to the time it takes to complete the survey compared to unfolding the questions or paginating the questions where each question is on a different screen. Eye tracking with English speaking test participants on the NSSRN instrument showed that participants did not spend much time reading the greyed out questions. While we did not observe participants reporting in such a way to avoid answering subsequent questions (and in fact observed the opposite), further study with real respondents in a situation where there is no payment is needed in order to confirm that respondents do not modify their answers so that they do not have to answer subsequent questions.

We recommend testing different types of conditional branching options for online, self-administered questionnaires (e.g. automatic graying out of skipped questions vs. selective revealing (expand-contract options)). Quantitative analysis can examine any human errors when using conditional branching, such as the tendency of respondents to answer in ways to avoid long sections, time-on-tasks differences, user satisfaction and likelihood to break off before finishing the questionnaire (Norman, 2001).
9. Acknowledgements

We thank the staff of Leah Meyer in ADDP for the opportunity to usability test the NSCH. We also thank our reviewers including Leah Meyer, Lin Wang, Patricia Goerman, Joanne Pascale, and Paul Beatty. We also thank Kevin Younes, our recruiter for this project.

10. References


Module T1 (for children 0-5 years old)
Attachment A: 2018 NSCH draft paper questionnaires – English version
Module T2 (for children 6-11 years old)
Attachment A: 2018 NSCH draft paper questionnaires – English version
Attachment B: 2018 NSCH draft paper questionnaires – Spanish version

Screener (Spanish)
Attachment B: 2018 NSCH draft paper questionnaires – Spanish version
Attachment B: 2018 NSCH draft paper questionnaires – Spanish version

Module T2 (for children 6-11 years old; Spanish)
<table>
<thead>
<tr>
<th>I. Acerca de su familia y su hogar</th>
</tr>
</thead>
<tbody>
<tr>
<td>DURANTE LA SEMANA PASADA, ¿dónde vive su familia?</td>
</tr>
<tr>
<td>- En su casa/donde vive desde niño/a</td>
</tr>
<tr>
<td>- En el hogar de los abuelos</td>
</tr>
<tr>
<td>- En un albergue/pabellón</td>
</tr>
<tr>
<td>- En un hogar infantil</td>
</tr>
<tr>
<td>- Otro (especifique)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>J. Cuidador(es) de este(a) niño(a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>¿Cómo está el grado de interacción con el niño(a)?</td>
</tr>
<tr>
<td>- Bien</td>
</tr>
<tr>
<td>¿Cuáles son las condiciones de vida del niño(a)?</td>
</tr>
<tr>
<td>- Pobre</td>
</tr>
<tr>
<td>¿Cuál es la edad del niño(a)?</td>
</tr>
<tr>
<td>- Menor de 1 año</td>
</tr>
<tr>
<td>¿Cuáles son las habilidades del niño(a)?</td>
</tr>
<tr>
<td>- Menos</td>
</tr>
<tr>
<td>¿Cuáles son las fortalezas del niño(a)?</td>
</tr>
<tr>
<td>- Pequeñas</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>¿Cuál es su estado civil?</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Soltero/a</td>
</tr>
</tbody>
</table>

| ¿Cuáles son los valores de la familia en su entorno social? |
| - Positivos | Negativos | Mixtos |

| ¿Cuáles son las expectativas de futuro del niño(a)? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |

| ¿Cuáles son las expectativas de futuro de la familia en su entorno social? |
| - Baja | Media | Alta |
Satisfaction Results

Screen Layouts for English speakers

Source: 2018 NSCH usability testing

Screen layouts for Spanish speakers

Source: 2018 NSCH usability testing
Use of terminology throughout the survey for
English speakers

Source: 2018 NSCH usability testing

Use of terminology throughout the survey
for Spanish speakers

Source: 2018 NSCH usability testing
Instructions displayed on screen for English speakers

Source: 2018 NSCH usability testing

Instructions displayed on the screen for Spanish speakers

Source: 2018 NSCH usability testing
Questions displayed on screen for English speakers

Source: 2018 NSCH usability testing

Questions displayed on screen for Spanish speakers

Source: 2018 NSCH usability testing
Questions can be answered in a straightforward manner for English speakers

Source: 2018 NSCH usability testing

Questions can be answered in a straightforward manner for Spanish speakers

Source: 2018 NSCH usability testing
Organization of questions, instructions and response categories for English speakers

Source: 2018 NSCH usability testing

Organization of questions, instructions and response categories for Spanish speakers

Source: 2018 NSCH usability testing
Forward navigation for English speakers

Source: 2018 NSCH usability testing. One participant did not answer the question.

Forward navigation for Spanish speakers

Source: 2018 NSCH usability testing
Overall experience of completing the survey for English speakers

Source: 2018 NSCH usability testing

Overall experience completing the survey for Spanish speakers

Source: 2018 NSCH usability testing
Correcting your mistakes for English speakers

Source: 2018 NSCH usability testing

Correcting your mistakes for Spanish speakers

Source: 2018 NSCH usability testing
Attachment B: 2018 NSCH draft paper questionnaires – Spanish version

Reading text for English speakers

Source: 2018 NSCH usability testing

Reading Text for Spanish speakers

Source: 2018 NSCH usability testing
Source: 2018 NSCH usability testing. One participant selected not applicable.

Source: 2018 NSCH usability testing
Attachment B: 2018 NSCH draft paper questionnaires – Spanish version

Source: 2018 NSCH usability testing. One participant selected not applicable.

Source: 2018 NSCH usability testing
Source: 2018 NSCH usability testing. Six participants selected not applicable.