Attachment B: 2025 Research and Analysis Plan for the Response Option and Error Message Design Test

American Community Survey Research and Evaluation Program

December 21, 2024

ACS Research & Evaluation Analysis Plan (REAP)

2025 Response Option and Error Message Design Test

United States™ Census

REAP REVISION LOG

Version	Date	Description	Author
0.1	October 2024	Initial Draft for Feedback	Rachel Horwitz Elizabeth May Nichols Lauren Contard
0.2	November 2024	Draft for Critical Review	Rachel Horwitz Elizabeth May Nichols Lauren Contard
1.0	December 2024	Final REAP	Rachel Horwitz Elizabeth May Nichols Lauren Contard

TABLE OF CONTENTS

1. INTRODUCTION.....5
2. BACKGROUND.....6
2.1 Response Buttons.....6
2.2 Edit Message Display.....7
2.3 ACS Internet Data Collection.....8
3. LITERATURE REVIEW.....9
3.1 Response Buttons.....9
3.2 Edit Messages.....10
4. RESEARCH QUESTIONS AND METHODOLOGY.....10
4.1 Sample Design.....10
4.2 Experimental Design.....10
4.2.1 Research Interest 1 – Response Buttons.....11
4.2.2 Research Interest 2 – Edit Message Format.....12
4.3 Research Questions.....13
4.3.1 Response Buttons.....13
4.3.2 Edit Messages.....13
4.3.3 Combined Effect.....13
4.4 Analysis Metrics.....14
4.4.1 Response Buttons.....14
4.4.2 Edit Message Format.....15
4.4.3 Standard Error of the Estimates.....15
4.4.4 Additional Analysis Metrics.....16
5. ASSUMPTIONS AND LIMITATIONS.....16
5.1 Assumptions.....16
5.2 Limitations.....16
6. TABLE SHELLS.....16
6.1 Response Button Table Shells.....17
6.2 Edit Message Format Table Shells.....18
6.3 Combined Effect Format Table Shells.....18
7. POTENTIAL CHANGES TO ACS.....20
8. REFERENCES.....20
Appendix A. Materials for the Experiment.....22

INTRODUCTION

In 2018, the U.S. Census Bureau formed a web survey design standards team as part of the time-limited Innovation and Operational Efficiency (IOE) program.[1] The IOE team's objective was to develop best practices for web design in an effort to reduce measurement error and respondent burden. Once the IOE program ended in 2021, the team continued its work under the Data Ingest and Collection for the Enterprise (DICE) program, where it currently lives.

The web survey design standards team in DICE was tasked with creating web survey design guidelines. The guidelines included the overall look and feel of the instrument (e.g., banners, font, color schemes), screen elements (e.g., navigation, branching, modals), question components (e.g., question stem, instructions, types of response choices), and specific screen types (e.g., dashboard, roster, summary). The guidelines were developed from existing literature and internally conducted experimental tests where the literature was insufficient or did not exist.[2] The DICE team's experimental tests were conducted using the Qualtrics survey platform. Qualtrics is an off-the-shelf survey solution, so customization of screens can be difficult. Because some of the potential design standards could not be replicated with Qualtrics' customization tools, the web standards team was unable to adequately test several research questions in Qualtrics.[3] Therefore, the DICE team was not able to develop standards for all aspects of survey design that they wished to address based on the Qualtrics results. Specifically, the designs of response buttons and edit messages were not able to be fully tested.[4]

Another issue with the Qualtrics experimental tests was the sample frame. The available sample came from a nonprobability panel. Using a nonprobability panel was not an issue for many of the standards, but it was for those that would impact some questions like the Race question. The nonprobability panel had a lower percentage of non-white respondents, which limited the analyses of some items pertaining to race and ethnicity.

The 2025 Response Option and Edit Message Design (RED) Internet Test will use the American Community Survey (ACS) to investigate the standards that were not adequately tested in the previous Qualtrics experimental tests. The purpose of the RED Internet Test is to determine the impact of our proposed standards on ACS response and respondent burden in the ACS internet response instrument. The Centurion platform, the Census Bureau's internet data collection survey system used for the ACS internet instrument, allows us to program features that we could not program in Qualtrics. Additionally, testing the standards with a probability sample and a nationally representative distribution of person-level demographic characteristics allows us to scientifically measure the impact of our proposed standards.

BACKGROUND

There are two specific standards that the web survey design standards team was unable to fully test in Qualtrics: response buttons and edit message display.

---

[1] The IOE program was created in 2010 to promote innovation at the enterprise level. Staff submitted projects that would provide benefit across the Bureau and those selected were funded at a corporate level. The IOE program ended in 2021 to focus on the transformation effort.

[2] The Census Bureau's web standards document (U.S. Census Bureau, 2024) provides citations for each standard.

[3] Section provides more details on the results and limitations of the tests conducted in Qualtrics.

[4] See Sections and for details on these items.

Response Buttons

To date, the Census Bureau has not used response buttons in any of its web surveys. Response buttons are similar to radio buttons (Figure 1) and check boxes. They can be programmed to allow for either the selection of a single response option like radio buttons, or multiple selection like check boxes. However, response buttons have an outline around the clickable area and highlight once a selection is made (Figure 2 and Figure 3). Respondents can click anywhere within the outlined box to make a selection. Response buttons are frequently used in non-governmental surveys because the large clickable area is easier to select and it is more apparent to the respondent that they have made a selection because the entire response is highlighted (Antoun et al., 2020).

The DICE team attempted to test the use of response buttons in the Qualtrics experiment. Qualtrics has a response button default for its surveys (Figure 2), but it could not be customized to accommodate write-in follow-up questions, like the Race and Place of Birth questions in the ACS have. Additionally, Qualtrics uses an opt-in panel for recruitment, so the demographic makeup of the Qualtrics sample was more homogenous than the general population, making it difficult to assess follow-up questions for minority populations. These limitations in Qualtrics resulted in limited research to develop a standard on whether response buttons should be used in Census Bureau web surveys moving forward.

Figure 1. Standard Radio Button Format

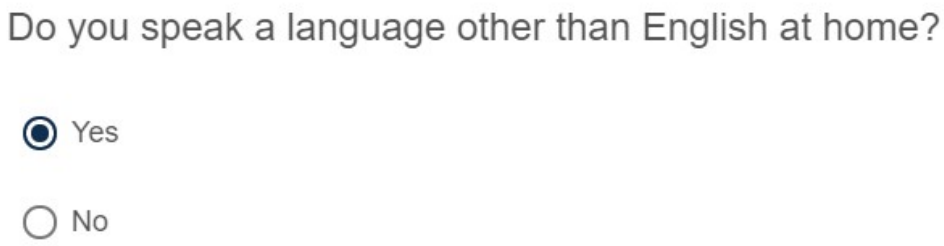


Figure 2. Response Button Format – Select One

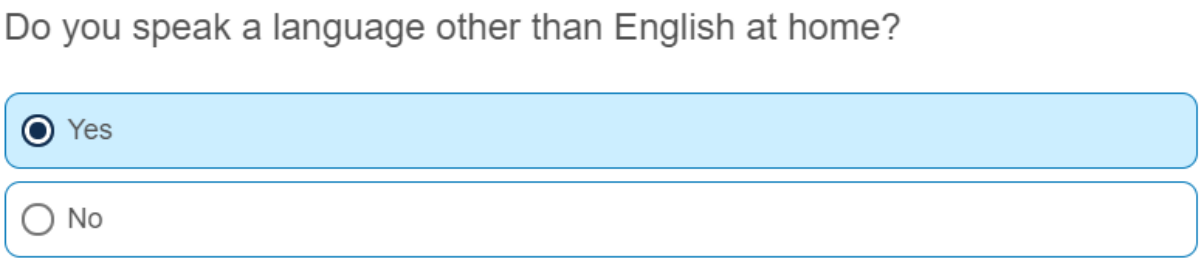
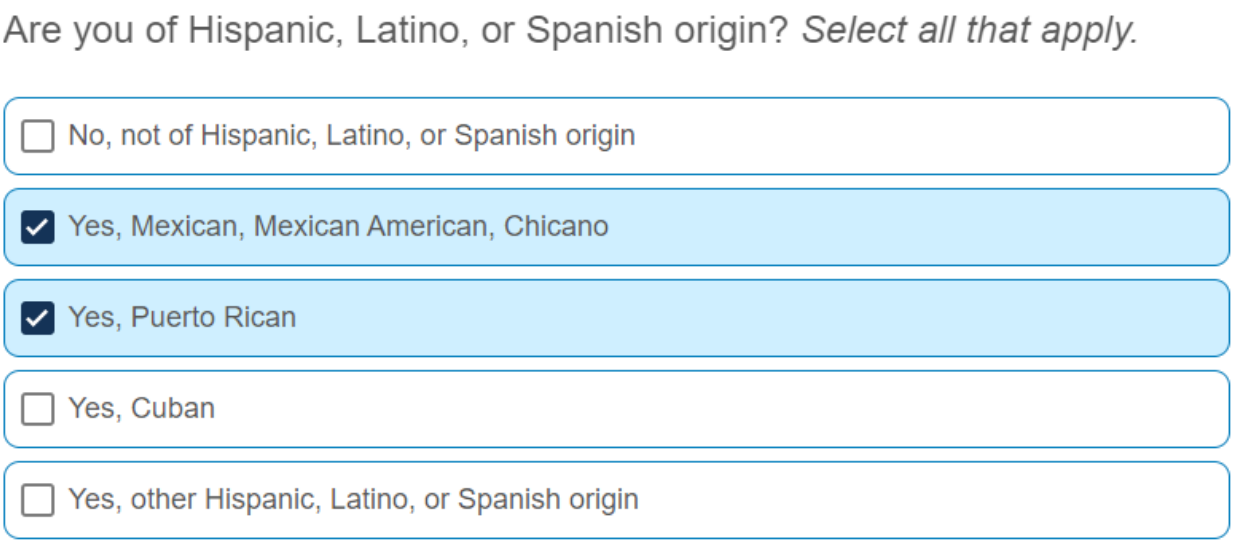


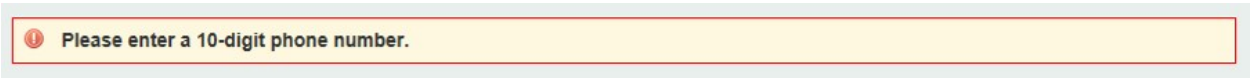
Figure 3. Response Button Format – Select All That Apply



Edit Message Display

Edit messages are used in web surveys to alert respondents that a survey response is incorrect or missing on a screen. This can include incorrect formatting (typing a character in a numeric field), leaving an item blank, or providing a response that is out of a predefined range. These edit messages may be hard or soft edits. Soft edits allow respondents to continue forward in the survey without making a correction while hard edits require a response or a valid response before the respondent can move on. The ACS internet instrument contains only soft edits (see example in Figure 4).

Figure 4. ACS Edit Message Display



The display of edit messages is inconsistent across the Census Bureau's different survey internet instruments. The various displays for edit messages do not appear to be problematic for Census Bureau surveys as respondents make corrections after receiving the messages (Horwitz et al., 2013). However, they are not consistent with the recommendation from the U.S. Web Design System (USWDS), "an active open source community of government engineers, content specialists, and designers," whose "contributors both in and out of government support dozens of agencies and nearly 200 sites" (USWDS).

When conducting the Qualtrics experimental tests of potential design standards, the DICE team found that Qualtrics does not allow the display of edit messages to be modified, nor does it allow for customized highlighting of specific fields that need to be attended to. For example, if a respondent is on the Place of Birth screen in the ACS internet instrument and they select the first radio button but select 'Next' before selecting a state, they receive an edit message at the top of the screen and the state field is highlighted to identify where the response is missing (Figure 5). The additional highlighting is an important feature to adequately convey to respondents what additional response is needed before moving to the next screen. The inability to manipulate that feature in Qualtrics meant a suitable experiment was not feasible.

Figure 5. Example of Edit Messages on the ACS Place of Birth Screen

ACS Internet Data Collection

The RED Internet Test will be conducted using the ACS self-response operation's internet mode. The monthly ACS production sample consists of approximately 295,000 housing unit addresses, which we refer to as a panel. Data collection for each panel occurs over three months. The first two months comprise the self-response period, and in the third month, the Computer-Assisted Personal Interviewing (CAPI) nonresponse follow-up operation begins.

A total of up to six mailings are sent to sampled households during the self-response period. The sooner a household responds to the ACS, the fewer mailings it receives. At a minimum, all households in sample receive the first two mailings. The first two mailings encourage households to respond online and provide a URL and an internet user ID that respondents enter to access the internet instrument. The third mailing contains a paper questionnaire that can be filled out and mailed back, but also informs households that they can still respond online. The fourth, fifth, and sixth mailings encourage online response,[5] and inform households that if they do not self-respond, a Census Bureau interviewer may visit them to complete an interview.

Of the remaining nonresponding addresses, a subsample is selected to be included in the CAPI operation. In CAPI, Census Bureau field representatives (FR) attempt to conduct interviews by phone or in-person visit. However, FRs also encourage households to self-respond, and internet responses are accepted until the end of the CAPI operation.

Additional information about the ACS data collection methodology is found in the ACS and Puerto Rico Community Survey (PRCS) Design and Methodology Report (U.S. Census Bureau, 2022).

LITERATURE REVIEW

This section describes the research that has previously been conducted by the DICE Web Standards Team regarding response buttons and edit messages. In some cases, the research is

[5] The fourth and fifth mailings also state that households can still respond using the paper questionnaire.

limited or nonexistent, which provides the motivation for the 2025 ACS RED Internet Test.

Response Buttons

While little research has been done directly comparing standard radio buttons and check boxes to response buttons, many surveys and platforms have moved to response buttons for a more modern look, including Qualtrics, Amazon Returns, and the United States Postal Service website. However, Antoun et al. (2020) did a direct comparison using smartphone respondents and testing four response option versions:

- The standard radio button format currently used in the ACS internet instrument and in the internet instruments for other Census Bureau surveys.
- The standard radio button format, but with larger radio buttons.
- Response buttons with no radio button inside the response box.
- Standard response buttons with a radio button inside the response box.[6]

In both the third and fourth versions, the response button area highlighted once it was selected.

The researchers found more mishits (that is, respondents selecting an option other than the one they intended) in the third version, the response button without a radio button. Additionally, participants did not like that condition as much as the radio button and response button with radio button formats. There were no significant differences between the radio button and response button with a radio button format. However, this research was conducted only on mobile devices, and its respondents were older than the general population (aged 59-80). Since older respondents may be more likely to have trouble using mobile devices, this research is not necessarily representative of how response buttons perform with the general population on mobile devices. It also does not provide information on how response buttons perform on computers and larger devices such as tablets.

In an effort to address these limitations, Horwitz et al. (2022) conducted a limited web study in Qualtrics using questions from the ACS comparing standard radio buttons to response buttons with radio buttons (Figure 1 and Figure 2, respectively). They found the response buttons yielded significantly faster response time, both overall and across different question types (select one, select all, write-ins), compared to radio buttons. They also found that more respondents in the response button condition found the survey to be “very easy,” and more respondents preferred the response button design. In terms of data quality, the response format did not have a significant impact on answer changes, mishits, or response distributions. However, there was significantly higher item nonresponse with the response button format for follow-up write-ins on questions like Race and Place of Birth. Additionally, while the Qualtrics sample had a more representative age distribution compared to Antoun et al. (2020), the overall racial and ethnic makeup of the Qualtrics sample was more homogenous than the general population (having more white participants and fewer Hispanic). This could have impacted the findings for the Race and Hispanic Origin questions.

Edit Messages

As described in Section , edit messages are programmed into web surveys to indicate possible errors, missing responses, or inconsistencies in the data reported. The messages may be triggered once the respondent selects a certain response option, or when they attempt to leave the survey page and navigate to another page, depending on the question. They are typically used to remind respondents of missing answers, or to alert the respondent to inconsistencies between different questions.

---

[6] This is similar to the treatment response button format being used in this test.

In general, edit messages are very beneficial to web surveys. They decrease item nonresponse for closed-ended questions, numerical answers, and frequency questions, and they increase overall data quality by correcting inconsistent or invalid responses (Holland, 2009; Couper, 2012). In the 2011 ACS Internet Test, over 90 percent of errors that triggered an edit message were corrected. Additionally, the edit messages did not seem to frustrate respondents (based on the patterns of changes they made in response) and did not lead to an increase in breakoffs (Horwitz et al., 2013).

While there is consensus that edit messages are beneficial (Horwitz et al., 2013), there does not seem to be much empirical research on the optimal format and display of these messages. The current edit messages used by the ACS seem to be working as intended; however, the USWDS recommends a different format for edit messages in general, and differentiates between hard and soft edits.[7]

USWDS recommends using red formatting for hard edits and yellow formatting for soft edits. This is consistent with common color uses suggesting red means stop and yellow means warning. The banner appears at the top of the screen and the outlined box surrounds a specific item that needs to be attended to. For example, if the state write-in was missed (a soft edit), that field would be outlined in yellow to help the respondent find the field with the issue. Currently, the ACS uses a black outline, yellow fill, and an arrow.

RESEARCH QUESTIONS AND METHODOLOGY

This section discusses the sample design, experimental design, and lists the research questions for the 2025 RED Internet Test. The goal of this test is to assess whether response buttons and new edit message formatting improve data quality and the user experience.

Sample Design

The 2025 RED Internet Test will be conducted using the August 2025 ACS production panel. The monthly ACS production panel consists of approximately 295,000 housing unit addresses and is divided into 24 nationally representative groups (referred to as methods panel groups) of approximately 12,000 addresses each. This test will use all 24 methods panel groups. Each group will be randomly assigned to one of the four treatments (control, response buttons, edit message formatting, combined response buttons and edit message formatting), so that each treatment uses six randomly assigned methods panel groups.

Experimental Design

This test will include a control group and three treatment groups. The control group will receive the 2025 ACS production internet instrument. The three treatment groups will have the following changes to the production internet instrument, as outlined in Section :

- Treatment 1: Replace radio buttons and ‘select all that apply’ checkboxes with response buttons. Response buttons outline the touch or click area and are colored once a selection is made.
- Treatment 2: Updated edit message formatting (yellow formatting and outline of missing response, where applicable)
- Treatment 3: Use response buttons and updated edit message formatting

Example screenshots of all treatments are shown in Error: Reference source not found. This

[7] See Section for an overview of edit messages in the current ACS.

design will be fully factorial, allowing us to measure both the impact of each treatment as well as the overall impact of the combined changes to response buttons and edit messages, which is how they would be implemented on the production ACS to follow the DICE web standards.

Research Interest 1 – Response Buttons

As mentioned in Section , response buttons have become widely used in survey and web design. Unlike standard radio buttons or check boxes (Figure 6), response buttons have an outline around the entire clickable area and highlight the response option when it is hovered over or selected, making it very clear to respondents which option they are selecting. Figure 7 depicts the response button format that will be used in this test. In Treatments 1 and 3, all standard radio buttons and check boxes in the internet instrument will be replaced with response buttons, including those with write-ins associated with the radio button or check box (e.g., Place of Birth, Race, Hispanic Origin).[8] Figure 8 is an example of a question with response buttons and a write-in field as they will appear in this test.

Figure 6. Current ACS Radio Button Format

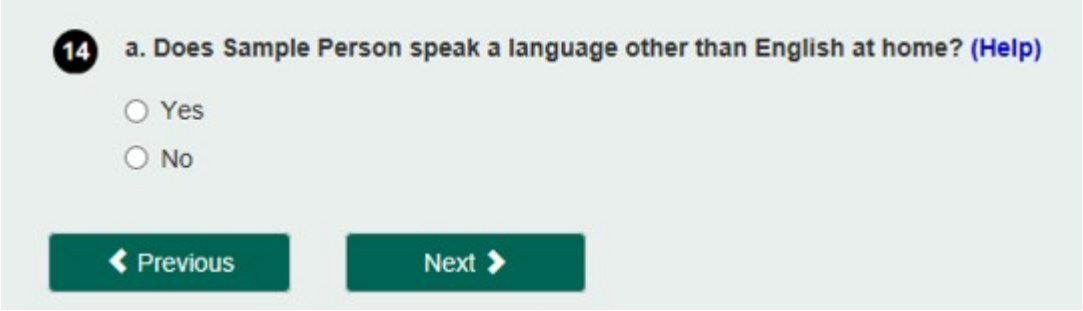


Figure 7. RED Test Response Button Format

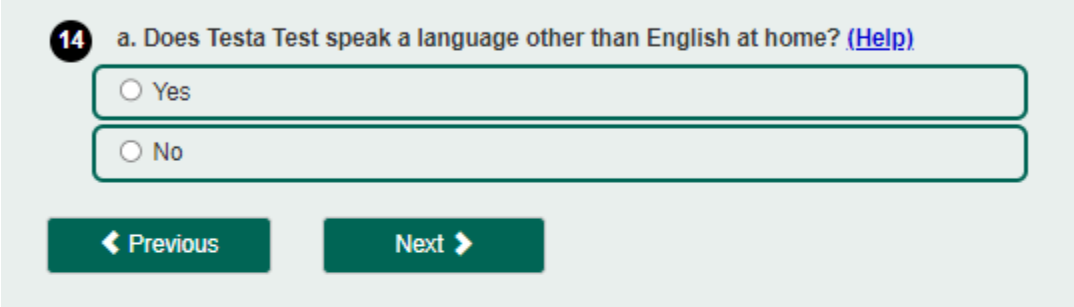
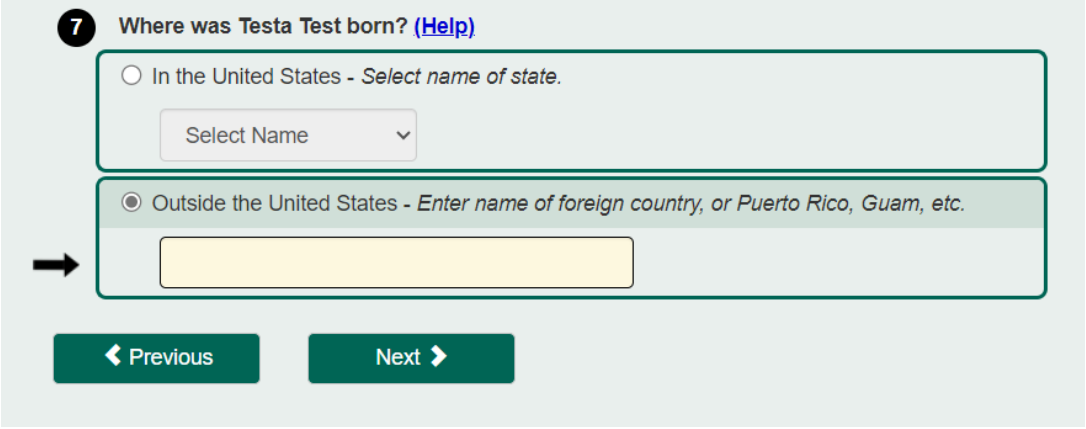


Figure 8. RED Test Response Button Format with Write-In



Research Interest 2 – Edit Message Format

While the current ACS edit messages (Figure 9) seem to function as intended as respondents

[8] The only questions that will not be affected are grids and text-entry items.

often provide a response where one was missing or correct incorrect/out of range information (Horwitz et al., 2013), the USWDS recommendation for edit messages will replace the current edit message format in Treatments 2 and 3 (Figure 10).[9] In addition to updating the color of the banner at the top of the screen, any missing field will be highlighted with a yellow border. Currently in the ACS, the field area is highlighted yellow. If the edit message is associated with a follow-up question (e.g., 'Other, specify' or 'State/Country of birth'), there is also a black arrow pointing at the highlighted field and the field has a black outline. The wording of the messages will remain as they currently are. The only difference between Treatments 2 and 3 and the control group will be the formatting.

Figure 9. Current ACS Edit Message Format

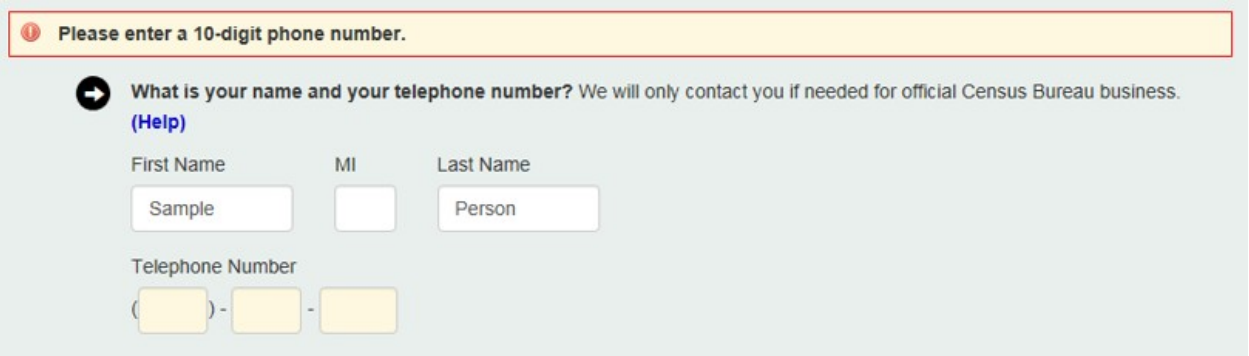
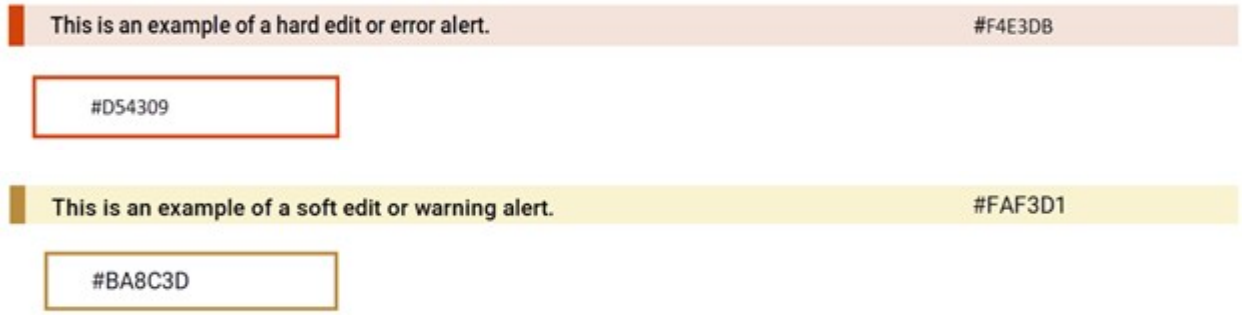


Figure 10. USWDS Recommended Hard and Soft Edit Message Format



Research Questions

The 2025 RED Internet Test will answer the following questions, grouped by the change being tested: response buttons, edit messages, and combined effect.

Response Buttons

1) What is the impact of response buttons on data quality compared to standard radio buttons?

   a) Is the write-in item nonresponse rate different between Control and Treatment 1?

   b) Is the rate of multiple selections for 'select all that apply' questions different between Control and Treatment 1?

   c) Is the rate of edit message triggers different between Control and Treatment 1?

[9] The USWDS formats for both hard and soft edits are displayed here for information, but only soft edits are used in the ACS.

d) Is the rate of breakoffs different between Control and Treatment 1?

2) What is the impact of response buttons on efficiency (response time) compared to standard radio buttons?

   a) Is the average time on screen per question different between Control and Treatment 1?

   b) Is the rate of answer changes different between Control and Treatment 1?

3) Is there a difference in response distributions for individual questions between Control and Treatment 1?

Edit Messages

4) Does the Treatment 2 edit message format result in fewer corrected responses compared to Control?

5) Do respondents spend more time on a screen with edit messages in Treatment 2 compared to Control?

Combined Effect

This section of the analysis will focus on questions that have both response buttons and an edit message.

6) Is the write-in item nonresponse rate different between Control and Treatment 3?

7) Is the rate of multiple selections for ‘select all that apply’ questions different between Control and Treatment 3?

8) Is the rate of edit message triggers different between Control and Treatment 3?

9) Is the rate of corrections following an edit message different between Control and Treatment 3?

10) Is the rate of breakoffs different between Control and Treatment 3?

11) Is screen completion time different between Control and Treatment 3?

12) What is the impact of response buttons on efficiency (response time) compared to standard radio buttons?

   a) Is the average time on screen per question different between Control and Treatment 3?

   b) Is the rate of answer changes different between Control and Treatment 3?

13) Is there a difference in response distributions for individual questions between Control and Treatment 3?

Analysis Metrics

All internet response analyses will be weighted using the ACS base sampling weight (the inverse of the probability of selection). Cases in the CAPI subsample that respond by internet during the CAPI period will have a CAPI subsampling factor that will be multiplied by the base weight.

The research questions on response buttons will be tested using two-tailed t-tests. The sample size will be able to detect differences of approximately 0.31 percentage points between the write-in item nonresponse rates of the experimental treatments for the Race question, 1.22 percentage points between the write-in item nonresponse rates for the Year Built question, and 0.69 percentage points between the rates of multiple selection for the Health Insurance question (with 80 percent power and α=0.1).

The primary purpose of the edit messages portion of this test is to confirm that the changes to the formatting of the edit messages do not hurt data quality. Our intention is to implement the changes as long as no problems are found. Therefore, the research questions on edit messages will be tested using one-tailed t-tests to check if data quality with the treatment edit messages is *worse* than in Control. We will use a significance level of α=0.1 when determining significant differences between treatments.

Response Buttons

Response buttons will be evaluated both in terms of data quality and efficiency. To determine the effect on data quality, we will examine item nonresponse (particularly for questions with specify write-in responses like Race), the rate of multiple selections for 'select all that apply' questions, the rate of edit message triggers, and the overall breakoff rate.

$$Write\text{-}In\ Nonresponse\ Rate = \frac{\begin{array}{c}Number\ of\ respondents\ that\ selected\ a\ response\\ option\ with\ a\ write\text{-}in\ write\text{-}in\ blank\end{array}}{\begin{array}{c}Total\ number\ of\ respondents\ that\ selected\\ a\ response\ option\ with\ a\ write\text{-}in\end{array}} * 100$$

$$Multiple\ Selections\ Rate = \frac{\begin{array}{c}Number\ of\ respondents\ that\ selected\\ at\ least\ two\ response\ options\end{array}}{\begin{array}{c}Total\ number\ of\ respondents\ that\ answered\\ a\ question\ with\ multiple\ response\ options\end{array}} * 100$$

$$Edit\ Message\ Rate = \frac{Number\ of\ respondents\ that\ received\ at\ least\ one\ edit\ message}{Total\ number\ of\ internet\ respondents} * 100$$

$$Total\ Edit\ Message\ Rate\ for\ screen\ X = \frac{Number\ of\ edit\ messages\ received\ on\ screen\ X}{Total\ number\ of\ edit\ messages\ on\ screen\ X} * 100$$

$$Breakoff\ Rate = \frac{Number\ of\ respondents\ that\ started\ but\ did\ not\ finish\ the\ survey}{Total\ number\ of\ respondents\ that\ successfully\ logged\ into\ the\ survey} * 100$$

To assess efficiency, we will focus on paradata measures including time on screen and answer changes.[10] This analysis will not include screens where responses are provided in grids or write-in fields. To calculate the response time on each screen, we take the difference between the time the respondent selected the 'Next' button to leave the screen and the time they entered the screen. The average time per screen will be the sum of the screen-level response times divided by the total number of applicable screens in the instrument.

$$Average\ response\ time = \frac{\sum_{i}^{\square} (response\ time\ in\ seconds\ for\ screen\ i)}{Total\ number\ of\ applicable\ screens}$$

where *i* is each applicable screen

[10] Some screens contain more than one question, but we are not able to determine how respondents divided their attention between questions on the same screen, so our analysis will measure total time on each screen.

$$Overall\ Changed\ Answer\ Rate = \frac{Number\ of\ respondents\ that\ changed\ one \vee more\ answers}{Number\ of\ respondents} * 100$$

We will also evaluate response distributions to see if the response button format affects how respondents answer.

$$Response\ Distribution = \frac{\sum_{i}^{\square} (selections\ for\ response\ option\ i)}{Total\ number\ of\ responses} * 100$$

Edit Message Format

To evaluate the edit message format, we will compare the percentage of edit message triggers that are corrected in each display using the following formula:

$$\%\ Edit\ Messages\ Corrected = \frac{\sum(Corrected\ edit\ triggers)}{Total\ number\ of\ edit\ triggers} * 100$$

We will calculate the average time from when an error is triggered to when the respondent selects "Next" to measure attention paid to the edit message. If respondents spend more time on a screen after triggering an edit message, we assume they are acknowledging and focusing on the message. The time from trigger to the "Next" button selection also takes into account the time taken to respond or change a response to the question. This rate will be compared between the treatment and control groups.

$$Time\ on\ screen = Timestamp(next) - Timestamp(error\ trigger)$$

Standard Error of the Estimates

We will estimate the variances of the point estimates and differences using the Successive Differences Replication (SDR) method with replicate weights – the standard method used in the ACS (see U.S. Census Bureau, 2022, Chapter 12). In calculating the different rates, we will use replicate subsampling adjusted weights, which account for the initial sampling probabilities and the subsampling during the CAPI operation. We will calculate the variance for each rate and for the difference between rates using the formula below:

$$Var(X_0) = \frac{4}{80} \square_{r=1}^{80} (X_r - X_0)^2$$

where:

X_r = ¿ the estimate calculated using the r^{th} replicate

X_0 = ¿ the estimate calculated using the full sample

The standard error of the estimate (X_0) is the square root of the variance.

Additional Analysis Metrics

Prior to answering the research questions, we will investigate the underlying data to check that there are no differences between treatments in metrics (as designed) that could affect the research question results. Specifically, we will look at demographic distributions of Person 1 (who is typically the respondent) from internet responses with at least a "sufficient partial" level of completeness.[11] We will also test for any device differences (i.e., PC, tablet, and smartphone)

[11] In general, a sufficient partial internet response is one that has at least minimal information, which indicates an attempt to respond. The specific definition of a sufficient partial internet response is sensitive and for Census Bureau internal use only.

between the control and each of the treatment groups.

ASSUMPTIONS AND LIMITATIONS

Assumptions

1. A single ACS monthly sample is representative of an entire year (twelve panels) and the entire frame sample, with respect to both response rates and cost, as designed.

2. A single methods panel group (1/24 of the full monthly sample) is representative of the full monthly sample, as designed.

Limitations

1. This test will only collect data from people who choose to respond by internet. We will not be able to assess how the changes to response buttons and edit messages would work for those who choose to respond by paper or CAPI. If respondents similar to those currently in the paper and CAPI response universes respond by internet in the future, they may react differently to the changes than those who respond by internet in this test.

2. We will only be able to assess the effect of the changes to edit messages for respondents who trigger an edit message. Respondents who do not trigger any edit messages will be left out of this analysis.

TABLE SHELLS

Below are samples of tables that will be used in the final report to show results from this test.

Response Button Table Shells

Table 1. Sample Table for Write-in Item Nonresponse Rate

Question	Control	Treatment 1	Difference	P-value
Ethnicity	%%.%	%%.%	%%.% (#.#)	#.##
Race	%%.%	%%.%	%%.% (#.#)	#.##
Place of Birth	%%.%	%%.%	%%.% (#.#)	#.##
Year Built	%%.%	%%.%	%%.% (#.#)	#.##
Device	%%.%	%%.%	%%.% (#.#)	#.##

Source: U.S. Census Bureau, American Community Survey, 2025 RED Internet Test, DRB #########
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. An asterisk (*) indicates a statistically significant result. Significance was tested based on a two tailed t-test at the α=0.1 level.

Table 2. Sample Table for Percentage Selecting Multiple Options

Question	Control	Treatment 1	Difference	P-value
Ethnicity	%%.%	%%.%	%%.% (#.#)	#.##
Race	%%.%	%%.%	%%.% (#.#)	#.##
Device	%%.%	%%.%	%%.% (#.#)	#.##

Source: U.S. Census Bureau, American Community Survey, 2025 RED Internet Test, DRB #########
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. An asterisk (*) indicates a statistically significant result. Significance was tested based on a two tailed t-test at the α=0.1 level.

Table 3. Sample Table for Average Time on Screen

Question	Control	Treatment 1	Difference	P-value
Question 1	xx.x	xx.x	xx.x	#.##
Question 2…	xx.x	xx.x	xx.x	#.##
…Question Y	xx.x	xx.x	xx.x	#.##
Overall	xx.x	xx.x	xx.x	#.##

Source: U.S. Census Bureau, American Community Survey, 2025 RED Internet Test, DRB #########
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. An asterisk (*) indicates a statistically significant result. Significance was tested based on a two tailed t-test at the α=0.1 level.

Table 4. Sample Table for Answer Change Rate

Question	Control	Treatment 1	Difference	P-value
Question 1	%%.%	%%.%	%%.% (#.#)	#.##
Question 2…	%%.%	%%.%	%%.% (#.#)	#.##
…Question Y	%%.%	%%.%	%%.% (#.#)	#.##
Overall	%%.%	%%.%	%%.% (#.#)	#.##

Source: U.S. Census Bureau, American Community Survey, 2025 RED Internet Test, DRB #########
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. An asterisk (*) indicates a statistically significant result. Significance was tested based on a two tailed t-test at the α=0.1 level.

Table 5. Sample Table for Breakoff Rate

Question	Control	Treatment 1	Difference	P-value
Question 1	%%.%	%%.%	%%.% (#.#)	#.##
Question 2…	%%.%	%%.%	%%.% (#.#)	#.##
…Question Y	%%.%	%%.%	%%.% (#.#)	#.##
Overall	%%.%	%%.%	%%.% (#.#)	#.##

Source: U.S. Census Bureau, American Community Survey, 2025 RED Internet Test, DRB #########
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. An asterisk (*) indicates a statistically significant result. Significance was tested based on a two tailed t-test at the α=0.1 level.

Table 6. Sample Table for Edit Message Rate

Question	Control	Treatment 1	Difference	P-value
Question 1	%%.%	%%.%	%%.% (#.#)	#.##
Question 2…	%%.%	%%.%	%%.% (#.#)	#.##
…Question Y	%%.%	%%.%	%%.% (#.#)	#.##
Overall	%%.%	%%.%	%%.% (#.#)	#.##

Source: U.S. Census Bureau, American Community Survey, 2025 RED Internet Test, DRB #########
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. An asterisk (*) indicates a statistically significant result. Significance was tested based on a two tailed t-test at the α=0.1 level.

Edit Message Format Table Shells

Table 7. Sample Table for Average Time on Screen

Question	Control	Treatment 2	Difference	P-value
Question 1	xx.x	xx.x	xx.x	#.##
Question 2…	xx.x	xx.x	xx.x	#.##
…Question Y	xx.x	xx.x	xx.x	#.##

Source: U.S. Census Bureau, American Community Survey, 2025 RED Internet Test, DRB #########
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. An asterisk (*) indicates a statistically significant result. Significance was tested based on a one tailed t-test at the α=0.1 level.

Table 8. Sample Table for Percent of Edit Messages Corrected

Question	Control	Treatment 2	Difference	P-value
Question 1	%%.%	%%.%	%%.% (#.#)	#.##
Question 2…	%%.%	%%.%	%%.% (#.#)	#.##
…Question Y	%%.%	%%.%	%%.% (#.#)	#.##

Source: U.S. Census Bureau, American Community Survey, 2025 RED Internet Test, DRB #########
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. An asterisk (*) indicates a statistically significant result. Significance was tested based on a one tailed t-test at the α=0.1 level.

Combined Effect Format Table Shells

Table 9. Sample Table for Write-in Item Nonresponse Rate

Question	Control	Treatment 3	Difference	P-value
Ethnicity	%%.%	%%.%	%%.% (#.#)	#.##
Race	%%.%	%%.%	%%.% (#.#)	#.##
Place of Birth	%%.%	%%.%	%%.% (#.#)	#.##
Year Built	%%.%	%%.%	%%.% (#.#)	#.##
Device	%%.%	%%.%	%%.% (#.#)	#.##

Source: U.S. Census Bureau, American Community Survey, 2025 RED Internet Test, DRB #########
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. An asterisk (*) indicates a statistically significant result. Significance was tested based on a two tailed t-test at the α=0.1 level.

Table 10. Sample Table for Percentage Selecting Multiple Options

Question	Control	Treatment 3	Difference	P-value
Ethnicity	%%.%	%%.%	%%.% (#.#)	#.##
Race	%%.%	%%.%	%%.% (#.#)	#.##
Device	%%.%	%%.%	%%.% (#.#)	#.##

Source: U.S. Census Bureau, American Community Survey, 2025 RED Internet Test, DRB #########
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. An asterisk (*) indicates a statistically significant result. Significance was tested based on a two tailed t-test at the α=0.1 level.

Table 11. Sample Table for Average Time on Screen

Question	Control	Treatment 3	Difference	P-value
Question 1	xx.x	xx.x	xx.x	#.##
Question 2…	xx.x	xx.x	xx.x	#.##
…Question Y	xx.x	xx.x	xx.x	#.##
Overall	xx.x	xx.x	xx.x	#.##

Source: U.S. Census Bureau, American Community Survey, 2025 RED Internet Test, DRB #########
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. An asterisk (*) indicates a statistically significant result. Significance was tested based on a two tailed t-test at the α=0.1 level.

Table 12. Sample Table for Breakoff Rate

Question	Control	Treatment 3	Difference	P-value
Question 1	%%.%	%%.%	%%.% (#.#)	#.##
Question 2…	%%.%	%%.%	%%.% (#.#)	#.##
…Question Y	%%.%	%%.%	%%.% (#.#)	#.##
Overall	%%.%	%%.%	%%.% (#.#)	#.##

Source: U.S. Census Bureau, American Community Survey, 2025 RED Internet Test, DRB #########
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. An asterisk (*) indicates a statistically significant result. Significance was tested based on a two tailed t-test at the α=0.1 level.

Table 13. Sample Table for Edit Message Rate

Question	Control	Treatment 3	Difference	P-value
Question 1	%%.%	%%.%	%%.% (#.#)	#.##
Question 2…	%%.%	%%.%	%%.% (#.#)	#.##
…Question Y	%%.%	%%.%	%%.% (#.#)	#.##
Overall	%%.%	%%.%	%%.% (#.#)	#.##

Source: U.S. Census Bureau, American Community Survey, 2025 RED Internet Test, DRB #########
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. An asterisk (*) indicates a statistically significant result. Significance was tested based on a two tailed t-test at the α=0.1 level.

Table 14. Sample Table for Percent of Edit Messages Corrected

Question	Control	Treatment 3	Difference	P-value
Question 1	%%.%	%%.%	%%.% (#.#)	#.##
Question 2…	%%.%	%%.%	%%.% (#.#)	#.##
…Question Y	%%.%	%%.%	%%.% (#.#)	#.##

Source: U.S. Census Bureau, American Community Survey, 2025 RED Internet Test, DRB #########
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. An asterisk (*) indicates a statistically significant result. Significance was tested based on a two tailed t-test at the α=0.1 level.

Table 15. Sample Table for Percentage of Edit Messages Corrected

Question	Control	Treatment 3	Difference	P-value
Question 1	%%.%	%%.%	%%.% (#.#)	#.##
Question 2…	%%.%	%%.%	%%.% (#.#)	#.##
…Question Y	%%.%	%%.%	%%.% (#.#)	#.##

Source: U.S. Census Bureau, American Community Survey, 2025 RED Internet Test, DRB #########
Note: Minor additive discrepancies are due to rounding. Standard errors are in parentheses. An asterisk (*) indicates a statistically significant result. Significance was tested based on a two tailed t-test at the α=0.1 level.

POTENTIAL CHANGES TO ACS

This test could result in a change to the ACS internet instrument and all internet instruments at the Census Bureau if either or both of the treatments are incorporated into DICE web design standards. If the response buttons prove to yield higher data quality or similar data quality but less respondent burden, then they will be incorporated into the standards applicable to all surveys as they move to DICE. The USWDS edit message format will be incorporated into the DICE standards as long as they are not found to make data quality or respondent burden significantly worse.

REFERENCES

Antoun, C., Nichols, E., Olmsted-Hawala, E., & Wang, L. (2020). Using buttons as response options in mobile web surveys. Survey Practice, 13(1). https://doi.org/10.29115/SP-2020-0002.

Couper, M. P., Tourangeau, R., Conrad, F. G., & Zhang, C. (2013). The design of grids in web surveys. Social science computer review, 31(3), 322-345. https://doi.org/10.1177/0894439312469865

Holland, J. L., & Christian, L. M. (2009). The influence of topic interest and interactive probing on responses to open-ended questions in web surveys. Social Science Computer Review, 27(2), 196-212. https://doi.org/10.1177/0894439308327481

Horwitz, R., Nichols, E. M., Katz, J., & Davis, M. (2021). Use of response buttons. Presentation available upon request.

Horwitz, R., Tancreto, J. G., Zelenak, M. F., & Davis, M. (2013). Use of paradata to assess the quality and functionality of the American Community Survey internet instrument. United States Census Bureau. Retrieved November 7, 2024 from https://www.census.gov/content/dam/Census/library/working-papers/2013/acs/2013_Horwitz_01.pdf

U.S. Census Bureau. (2022). American Community Survey and Puerto Rico Community Survey Design and Methodology, Version 3.0. https://www.census.gov/programs-surveys/acs/methodology/design-and-methodology.html

U.S. Census Bureau. (2024). Design guidelines for U.S. Census Bureau web surveys and censuses. Retrieved November 14, 2024 from uscensus.sharepoint.com/sites/DICE/Baselined Documents/Forms/AllItems.aspx?id=%2Fsites%2FDICE%2FBaselined Documents%2FISR - Web Survey Design Guidelines%2Epdf&parent=%2Fsites%2FDICE%2FBaselined Documents

U.S. Web Design System (USWDS). General Services Administration. Retrieved November 6, 2024, from https://designsystem.digital.gov/

Appendix A: Materials for the Experiment

American Community Survey

Internet Instrument Response Option and Error Message Design Test

2025

How to Use This Guide

This document contains copies of screens respondents will see in the ACS Internet Questionnaire[12]. These serve as examples of the types of questions impacted by the RED test. We will show the same impacted questions throughout the control and treatments.

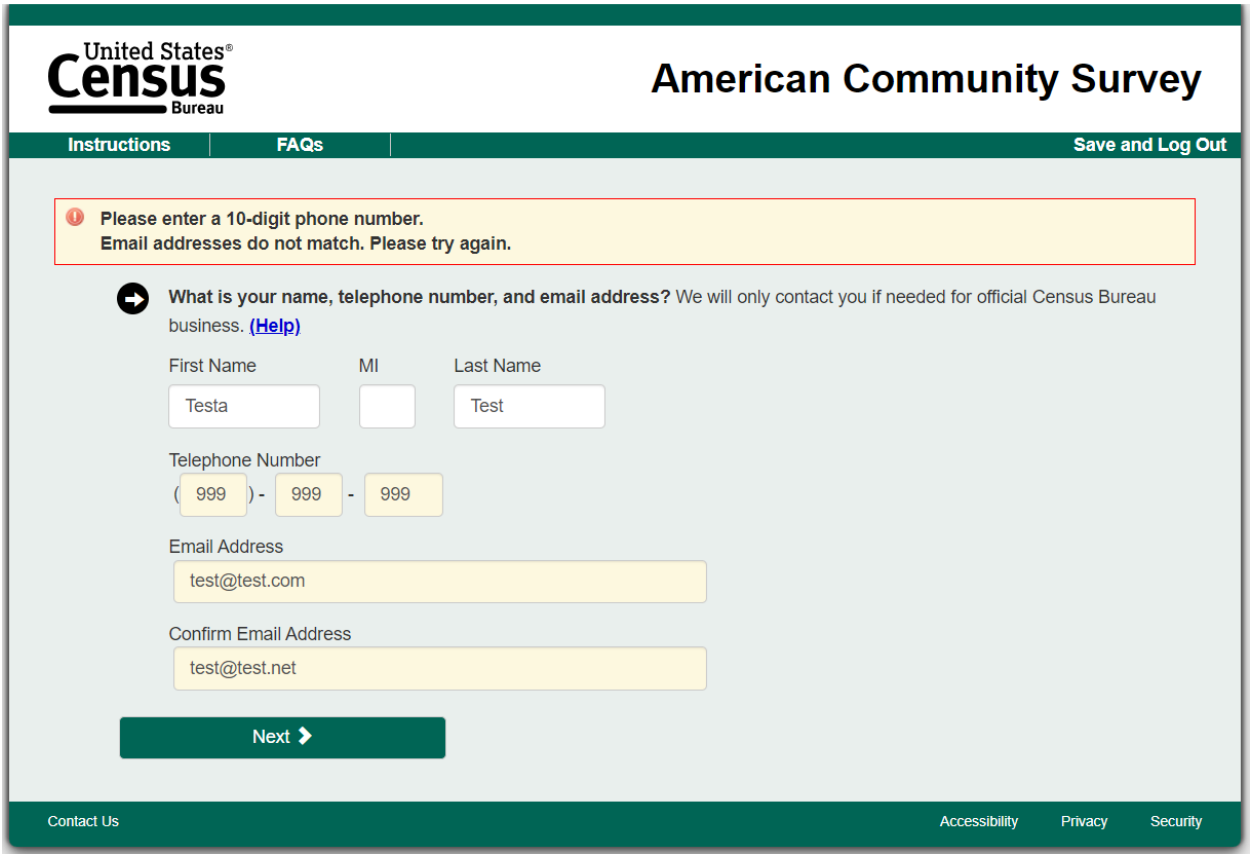
Note: in the 2016 data collection year a mobile optimized view was introduced for the Internet Data Collection instrument. If respondents are viewing the online instrument via a mobile device the screen layout will appear different based upon screen size. In addition, navigation buttons containing “Previous” and “Next” are replaced with forward and backward arrows. Instructions, FAQs and Save and Logout text are removed and replaced with links on the right-hand side of the header. Questions with a large amount of text will not be displayed on one screen and users will have to scroll to view the entire questions text.

[12] This screen capture guide does not contain any Title 13 data or other personally identifiable information (PII). All data are fictitious and any resemblance to actual data is coincidental.

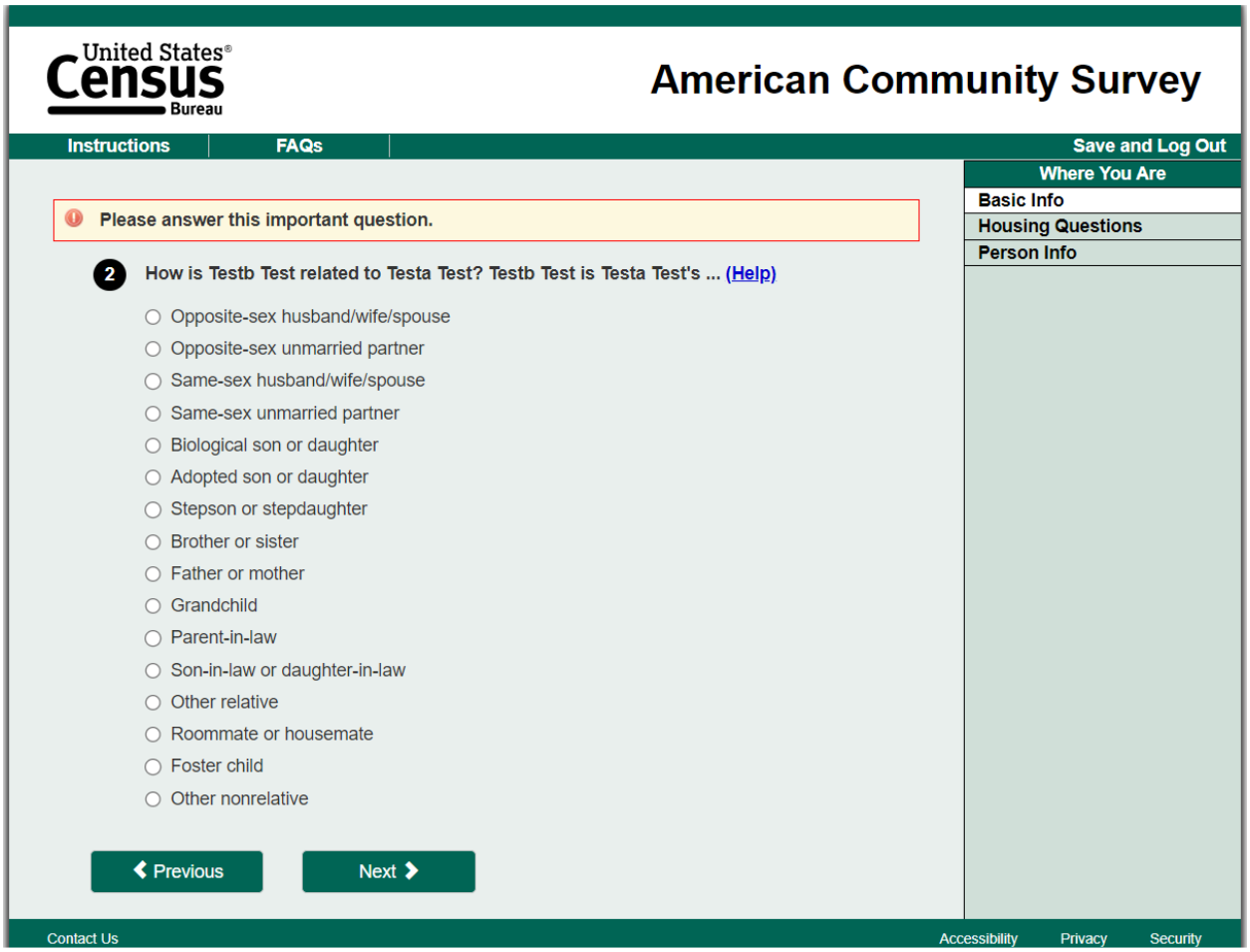
Control (C):

One fourth of respondents will receive the control, which is the current internet production instrument. No changes will be made. We have included example questions impacted (with errors) in the treatments as a baseline for comparison.

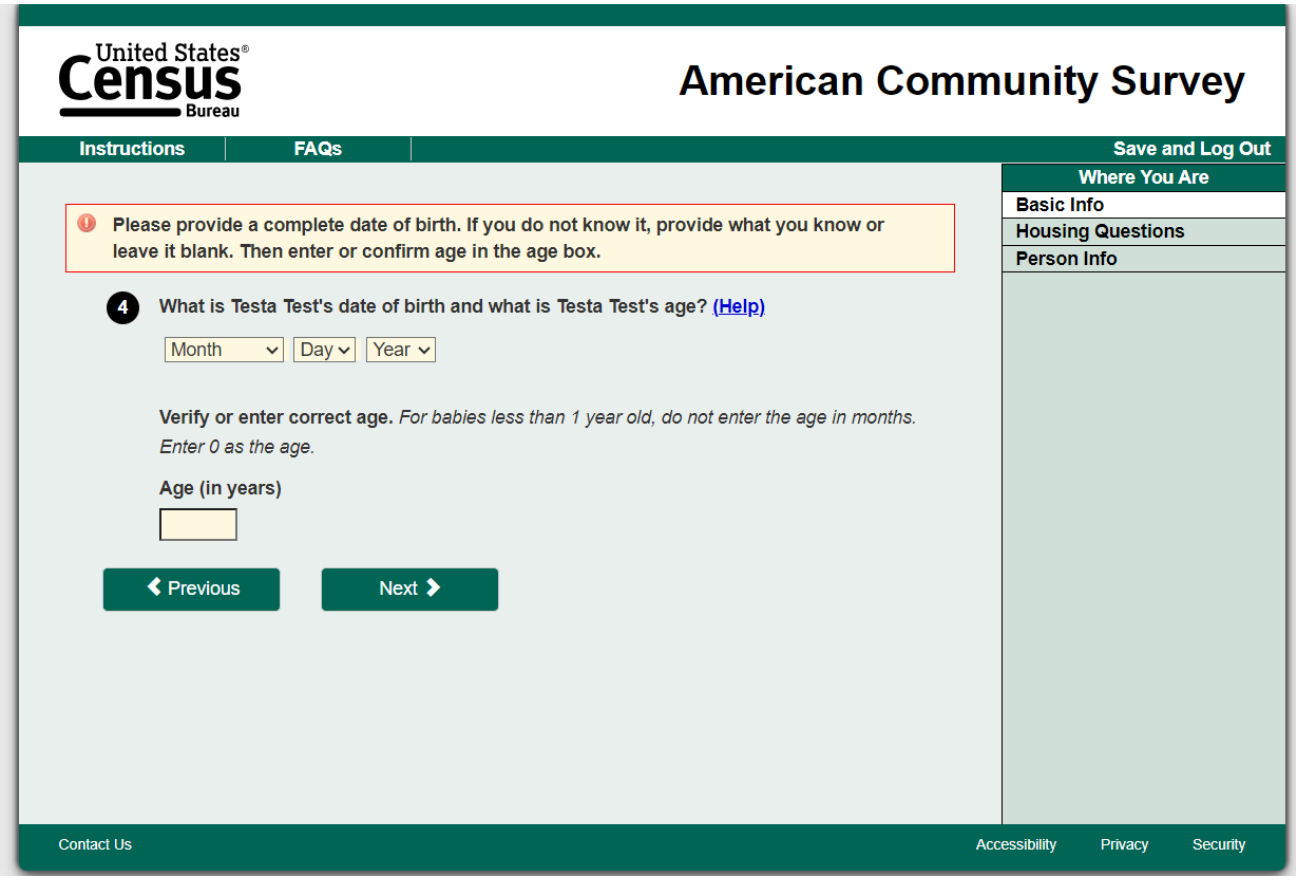
Respondent Name (C)



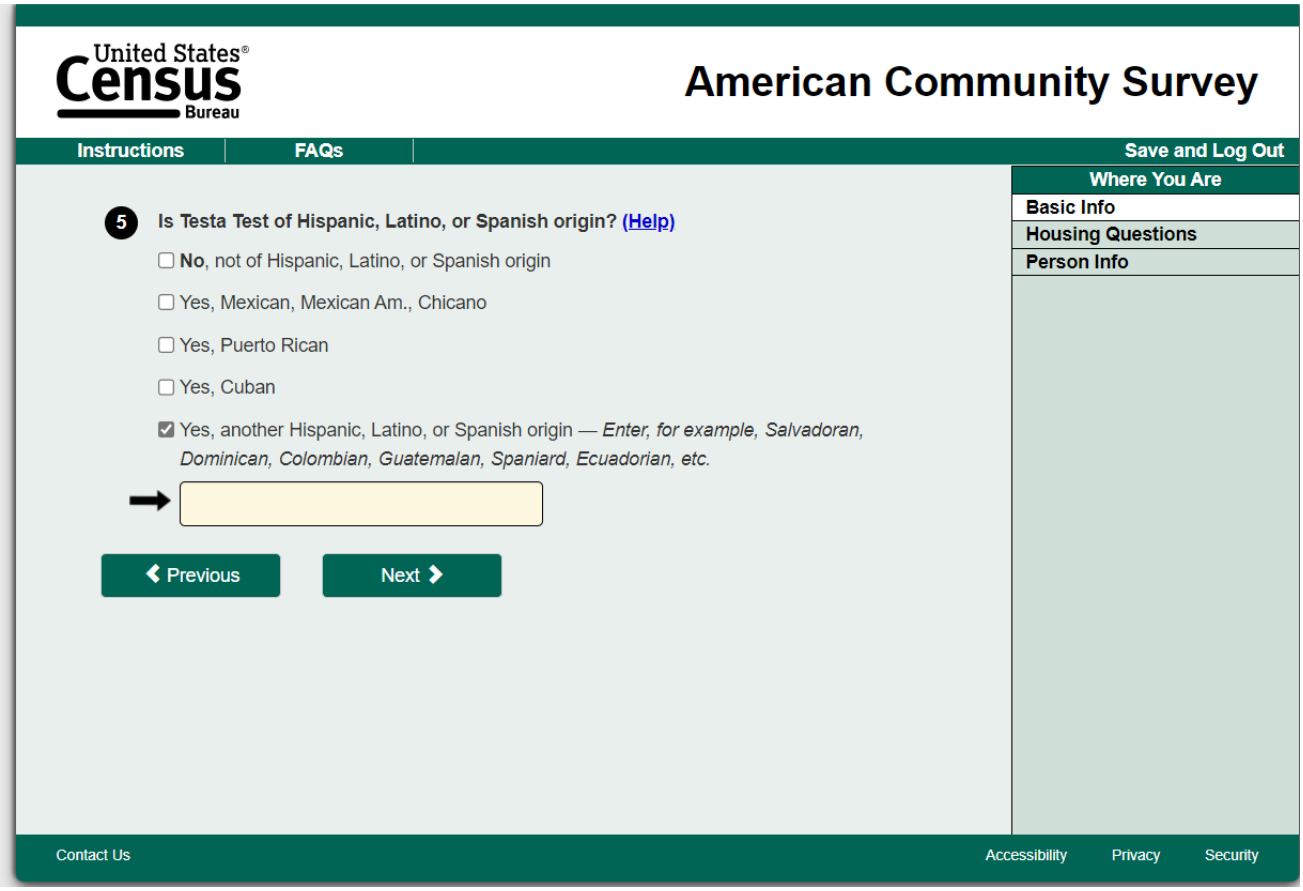
Relationship (C)



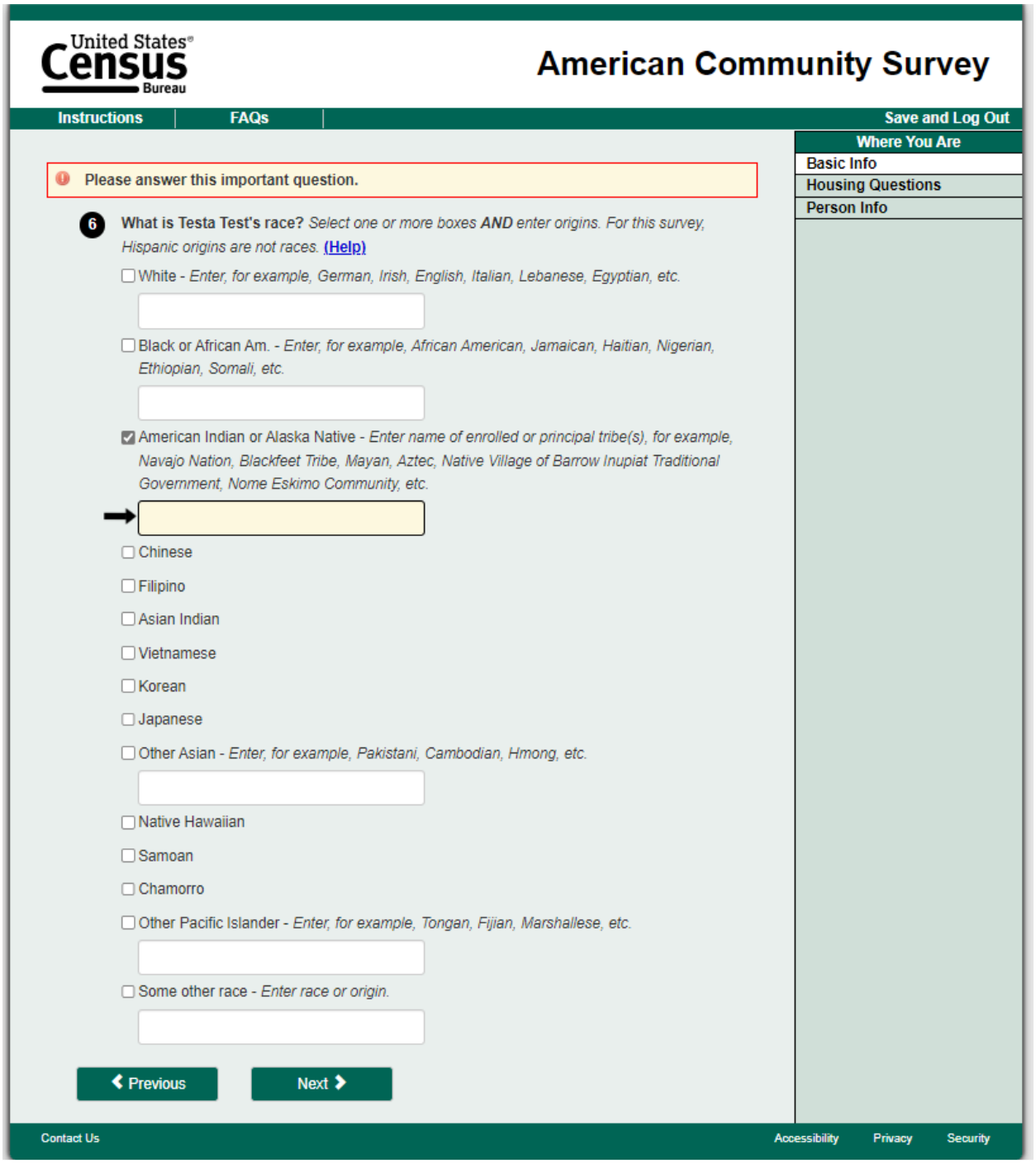
Date of Birth (C)



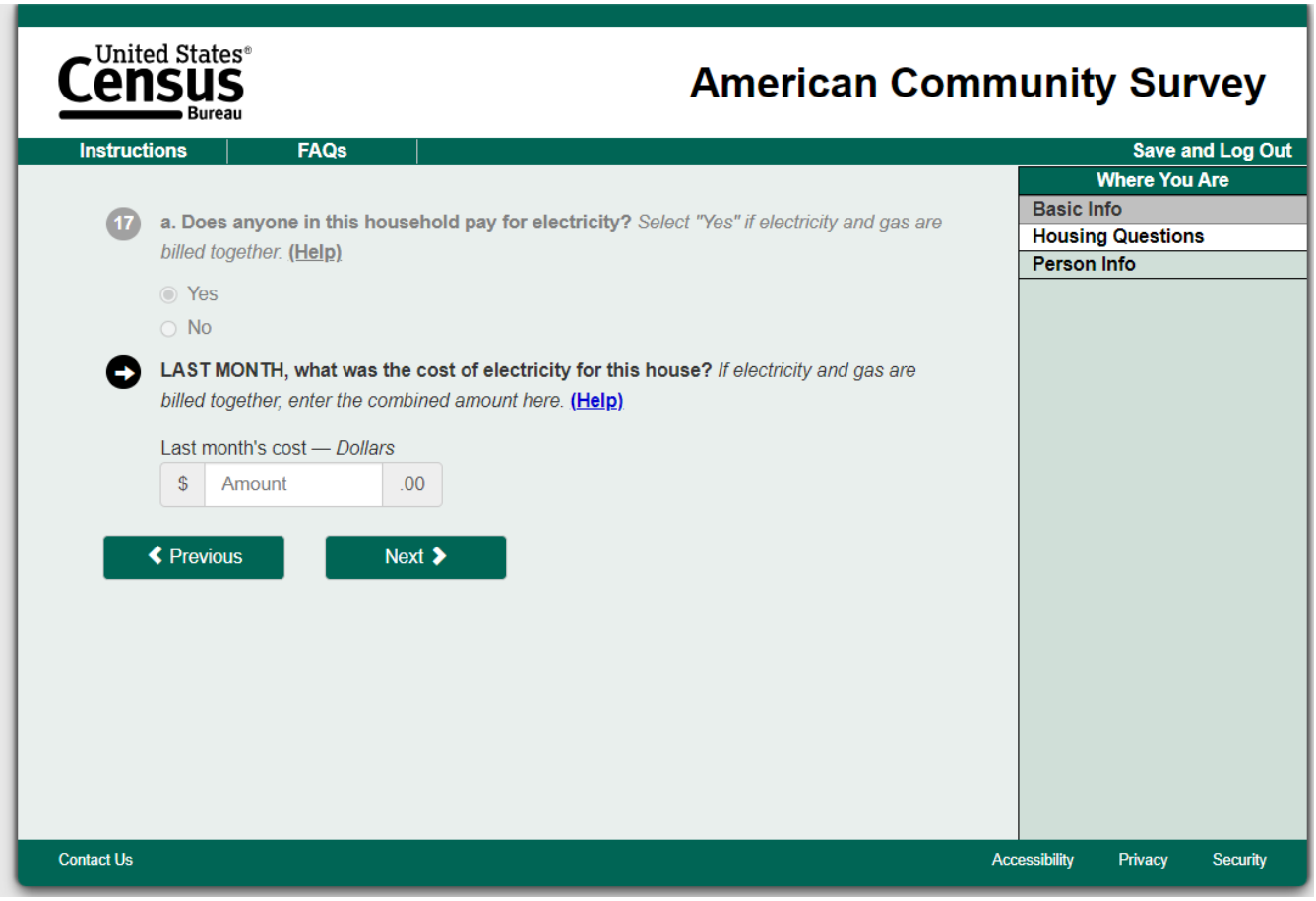
Hispanic (C)


United States Census Bureau
American Community Survey
Instructions
FAQs
Save and Log Out
Where You Are
Basic Info
Housing Questions
Person Info
5
Is Testa Test of Hispanic, Latino, or Spanish origin? (Help)
No, not of Hispanic, Latino, or Spanish origin
Yes, Mexican, Mexican Am., Chicano
Yes, Puerto Rican
Yes, Cuban
Yes, another Hispanic, Latino, or Spanish origin — Enter, for example, Salvadoran, Dominican, Colombian, Guatemalan, Spaniard, Ecuadorian, etc.
Previous
Next
Contact Us
Accessibility
Privacy
Security

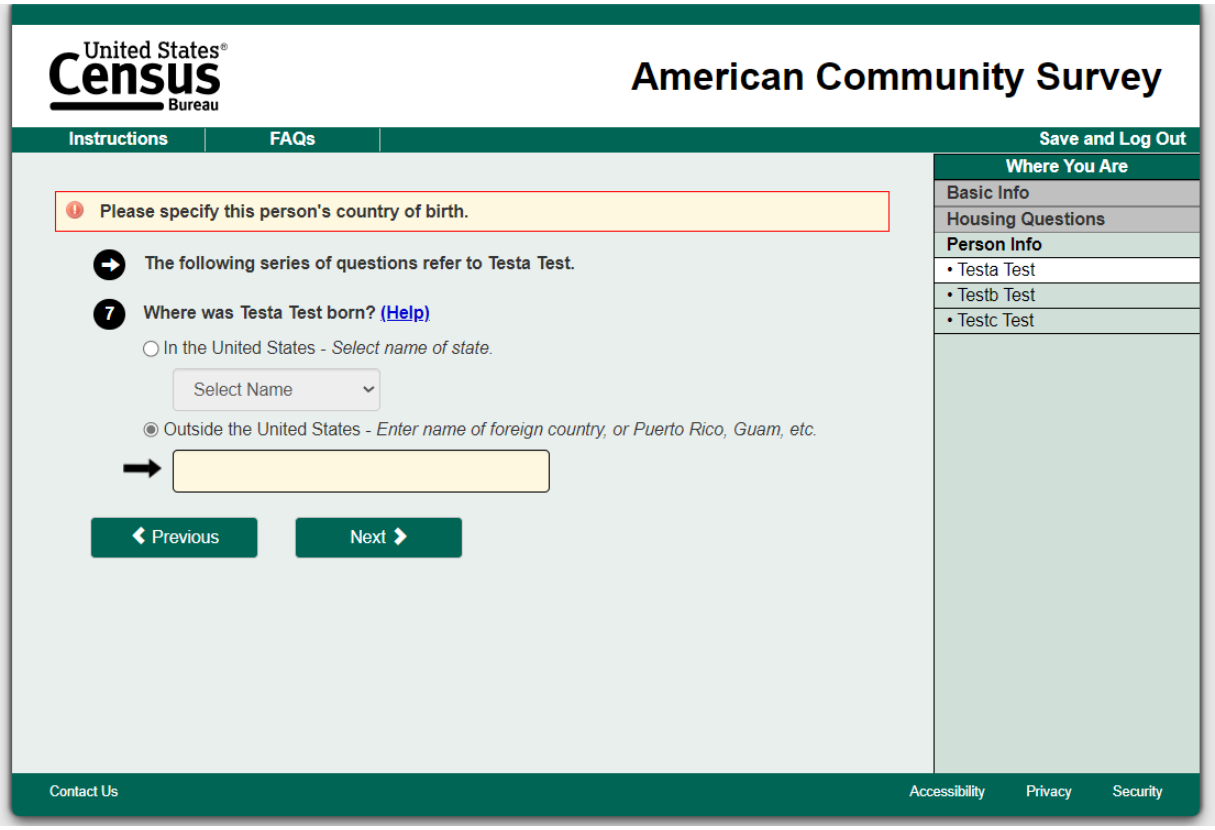
Race (C)


United States Census Bureau
American Community Survey
Instructions
FAQs
Save and Log Out
Where You Are
Basic Info
Housing Questions
Person Info
Please answer this important question.
6
What is Testa Test's race? Select one or more boxes AND enter origins. For this survey, Hispanic origins are not races. (Help)
White - Enter, for example, German, Irish, English, Italian, Lebanese, Egyptian, etc.
Black or African Am. - Enter, for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, etc.
American Indian or Alaska Native - Enter name of enrolled or principal tribe(s), for example, Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Native Village of Barrow Inupiat Traditional Government, Nome Eskimo Community, etc.
Chinese
Filipino
Asian Indian
Vietnamese
Korean
Japanese
Other Asian - Enter, for example, Pakistani, Cambodian, Hmong, etc.
Native Hawaiian
Samoan
Chamorro
Other Pacific Islander - Enter, for example, Tongan, Fijian, Marshallese, etc.
Some other race - Enter race or origin.
Previous
Next
Contact Us
Accessibility
Privacy
Security

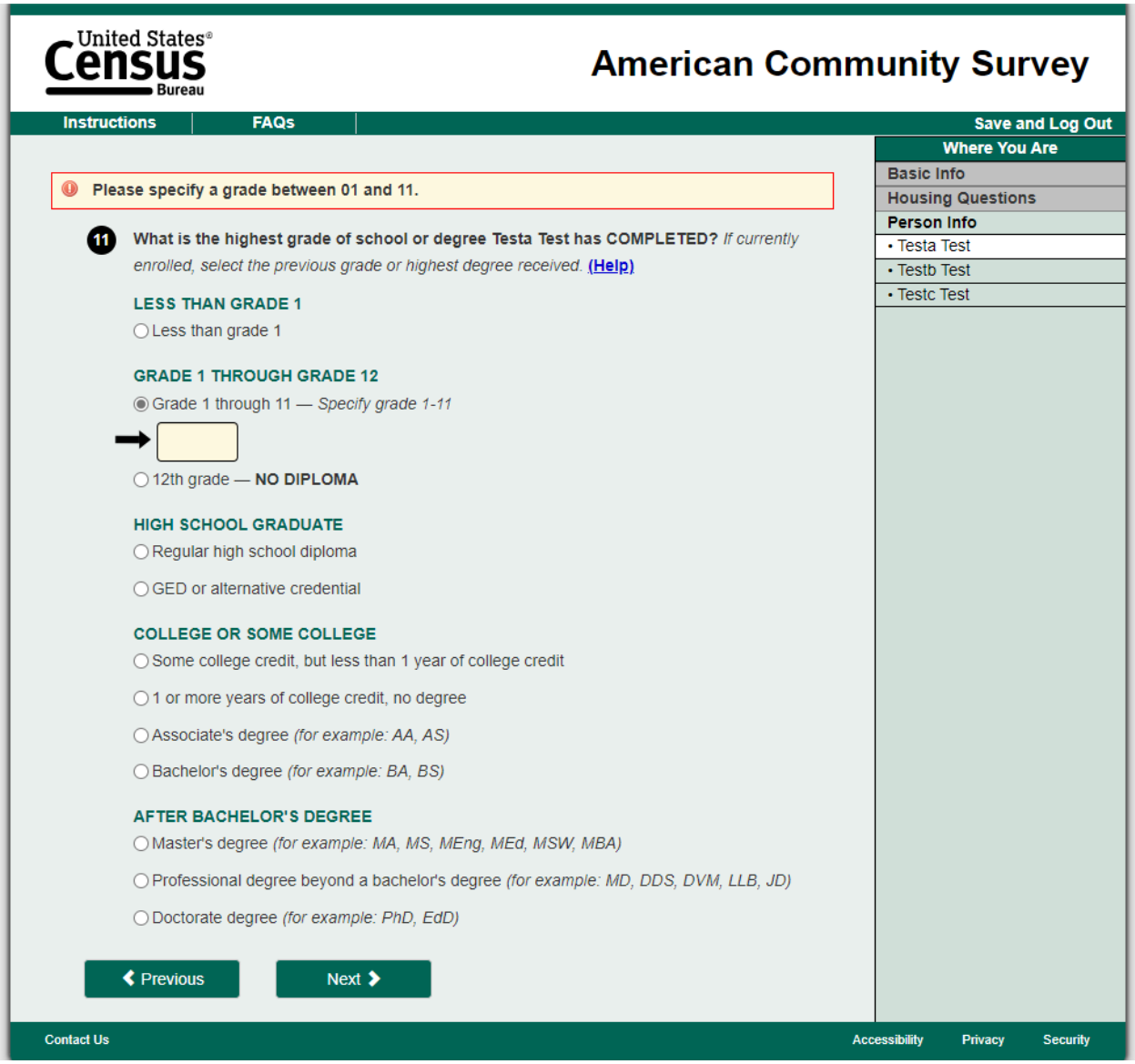
Electric Amount (C)



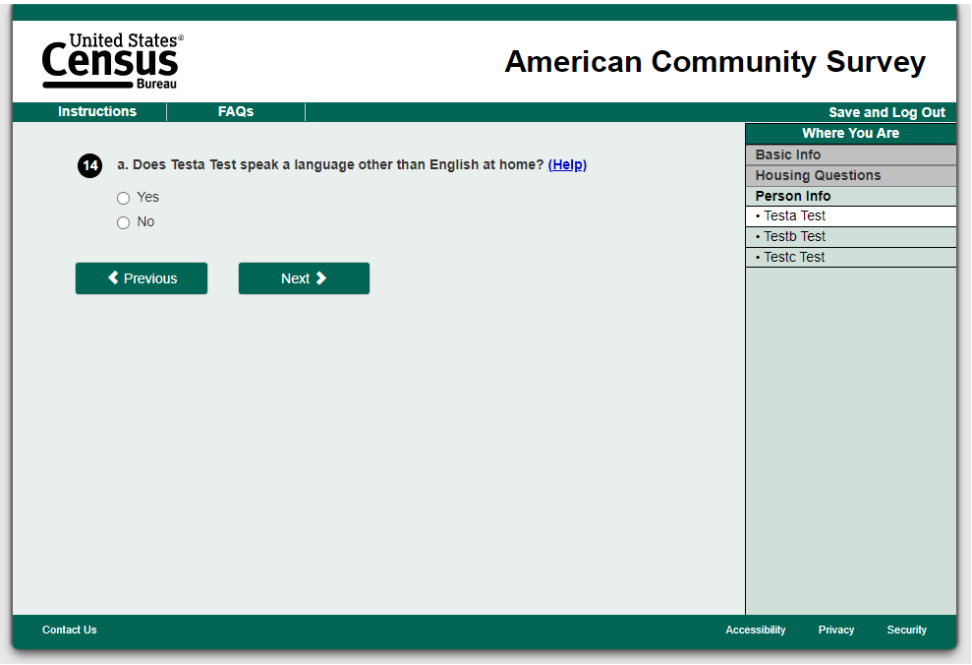
Place of Birth (C)



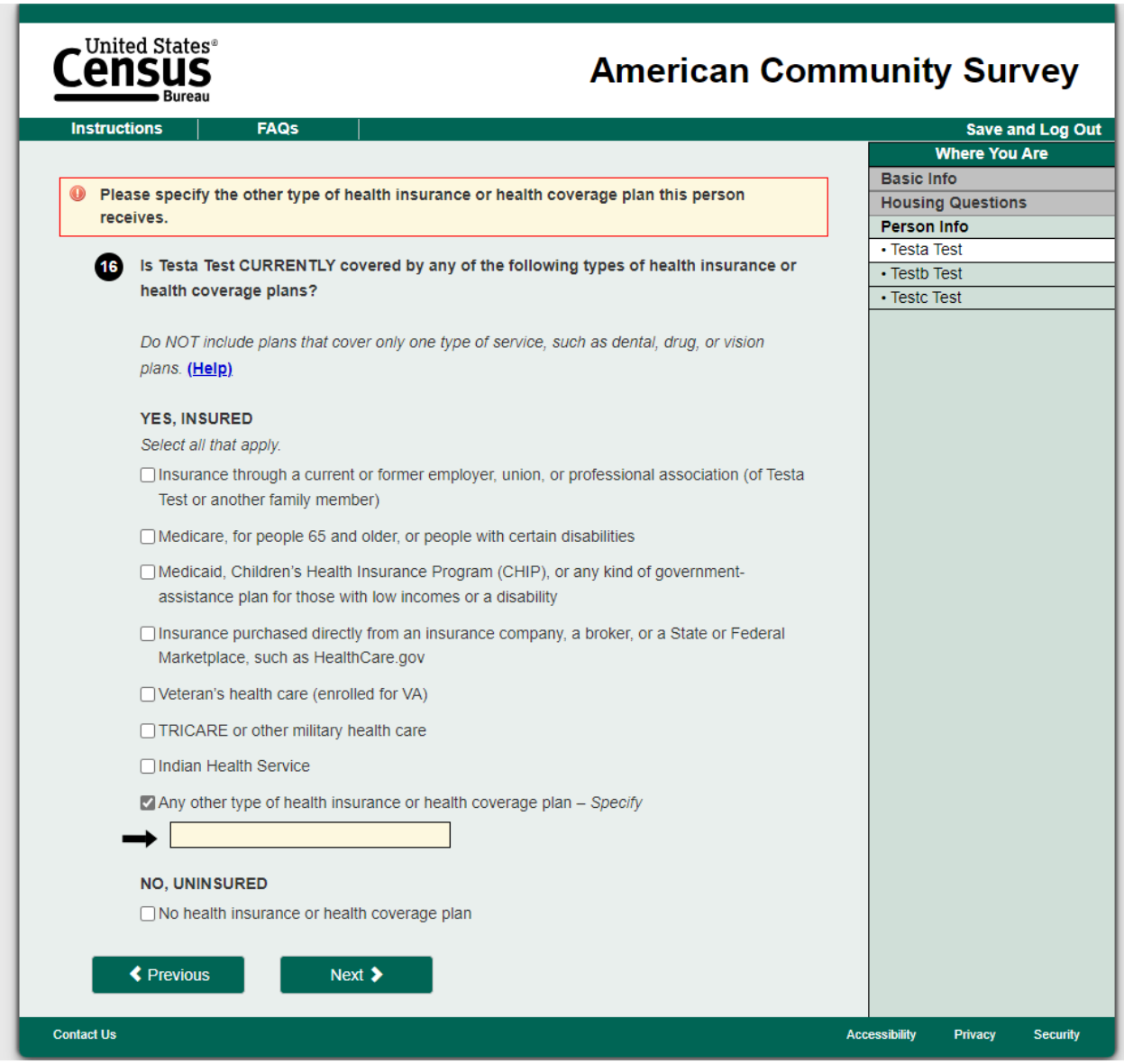
Highest Level (C)



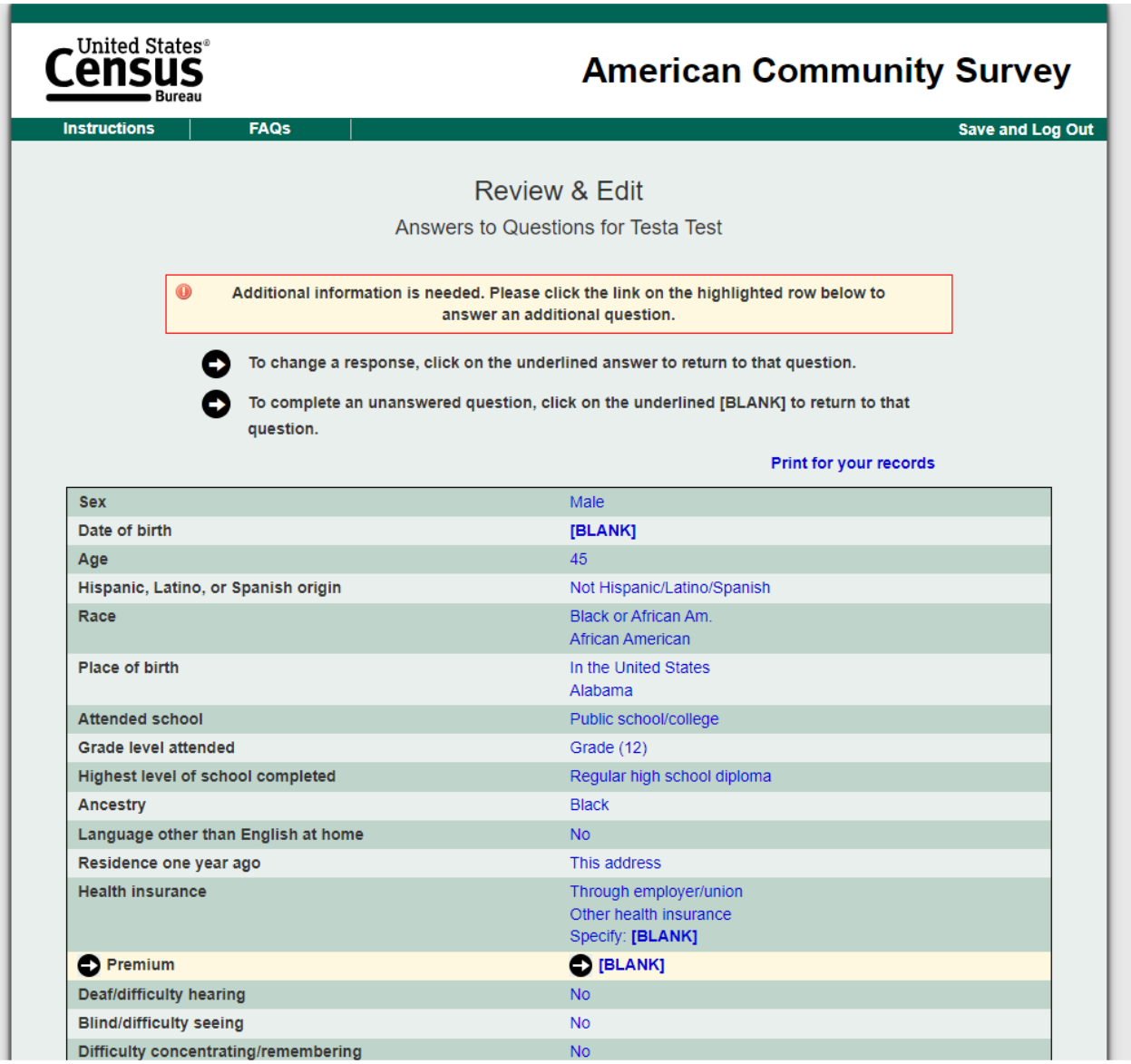
Language (C)


United States Census Bureau
American Community Survey
Instructions
FAQs
Save and Log Out
Where You Are
Basic Info
Housing Questions
Person Info
• Testa Test
• Testb Test
• Testc Test
14
a. Does Testa Test speak a language other than English at home? (Help)
Yes
No
Previous
Next
Contact Us
Accessibility
Privacy
Security

Insurance (C)


United States Census Bureau
American Community Survey
Instructions
FAQs
Save and Log Out
Where You Are
Basic Info
Housing Questions
Person Info
• Testa Test
• Testb Test
• Testc Test
Please specify the other type of health insurance or health coverage plan this person receives.
16
Is Testa Test CURRENTLY covered by any of the following types of health insurance or health coverage plans?
Do NOT include plans that cover only one type of service, such as dental, drug, or vision plans. (Help)
YES, INSURED
Select all that apply.
Insurance through a current or former employer, union, or professional association (of Testa Test or another family member)
Medicare, for people 65 and older, or people with certain disabilities
Medicaid, Children's Health Insurance Program (CHIP), or any kind of government-assistance plan for those with low incomes or a disability
Insurance purchased directly from an insurance company, a broker, or a State or Federal Marketplace, such as HealthCare.gov
Veteran's health care (enrolled for VA)
TRICARE or other military health care
Indian Health Service
Any other type of health insurance or health coverage plan – Specify
NO, UNINSURED
No health insurance or health coverage plan
Previous
Next
Contact Us
Accessibility
Privacy
Security

Review (C)



United States Census Bureau

American Community Survey

Instructions | FAQs | Save and Log Out

Review & Edit

Answers to Questions for Testa Test

Additional information is needed. Please click the link on the highlighted row below to answer an additional question.

To change a response, click on the underlined answer to return to that question.

To complete an unanswered question, click on the underlined [BLANK] to return to that question.

Print for your records

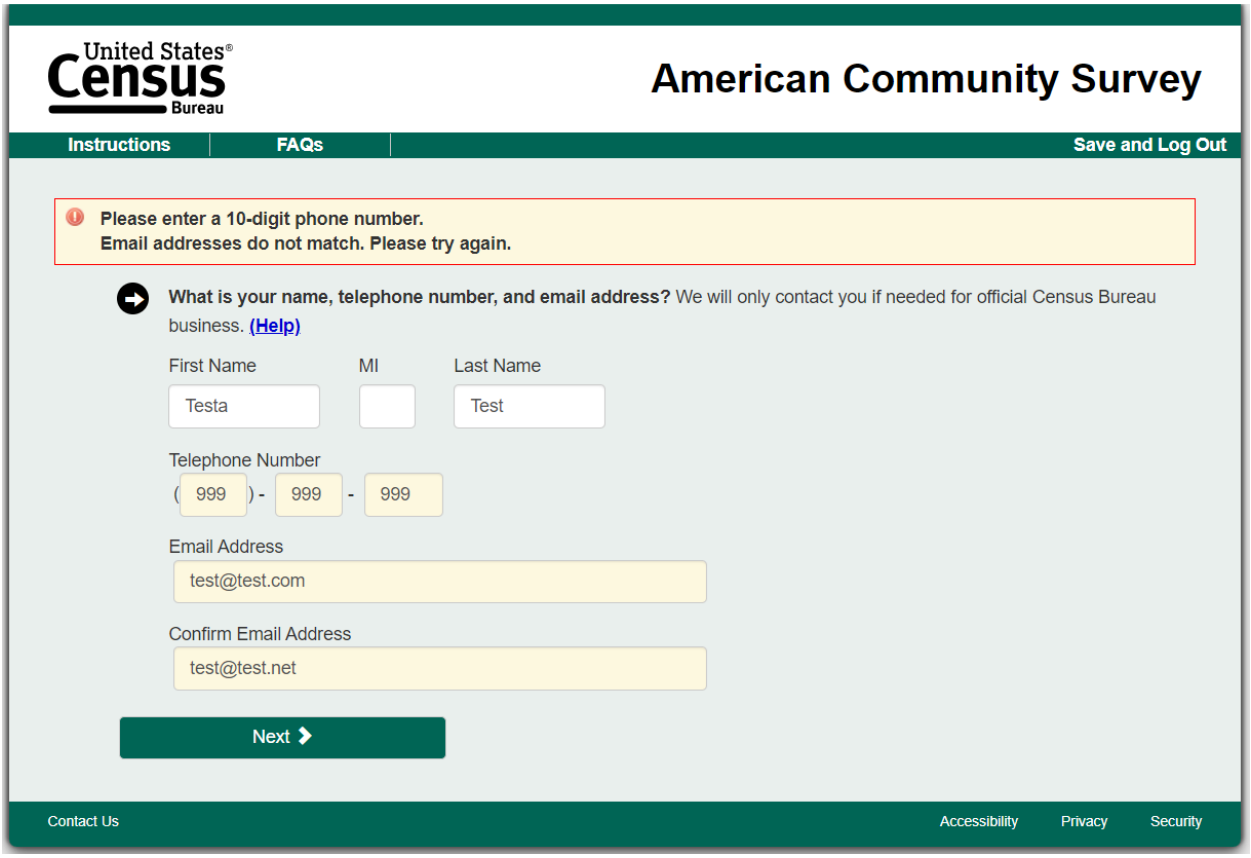
Sex	Male
Date of birth	[BLANK]
Age	45
Hispanic, Latino, or Spanish origin	Not Hispanic/Latino/Spanish
Race	Black or African Am. African American
Place of birth	In the United States Alabama
Attended school	Public school/college
Grade level attended	Grade (12)
Highest level of school completed	Regular high school diploma
Ancestry	Black
Language other than English at home	No
Residence one year ago	This address
Health insurance	Through employer/union Other health insurance Specify: [BLANK]
Premium	[BLANK]
Deaf/difficulty hearing	No
Blind/difficulty seeing	No
Difficulty concentrating/remembering	No

Treatment 1 (T1):

T1 will include changes to the answer fields. Answer fields include questions with multiple choice selections, including select-all-that-apply options. Changes include: Elongated box encompassing the answer options, a dark green outline, a light green shading once a respondent hovers over or selects, and entered radio buttons vertically when an option wraps around to more than one line. 130 screens/questions will be impacted.

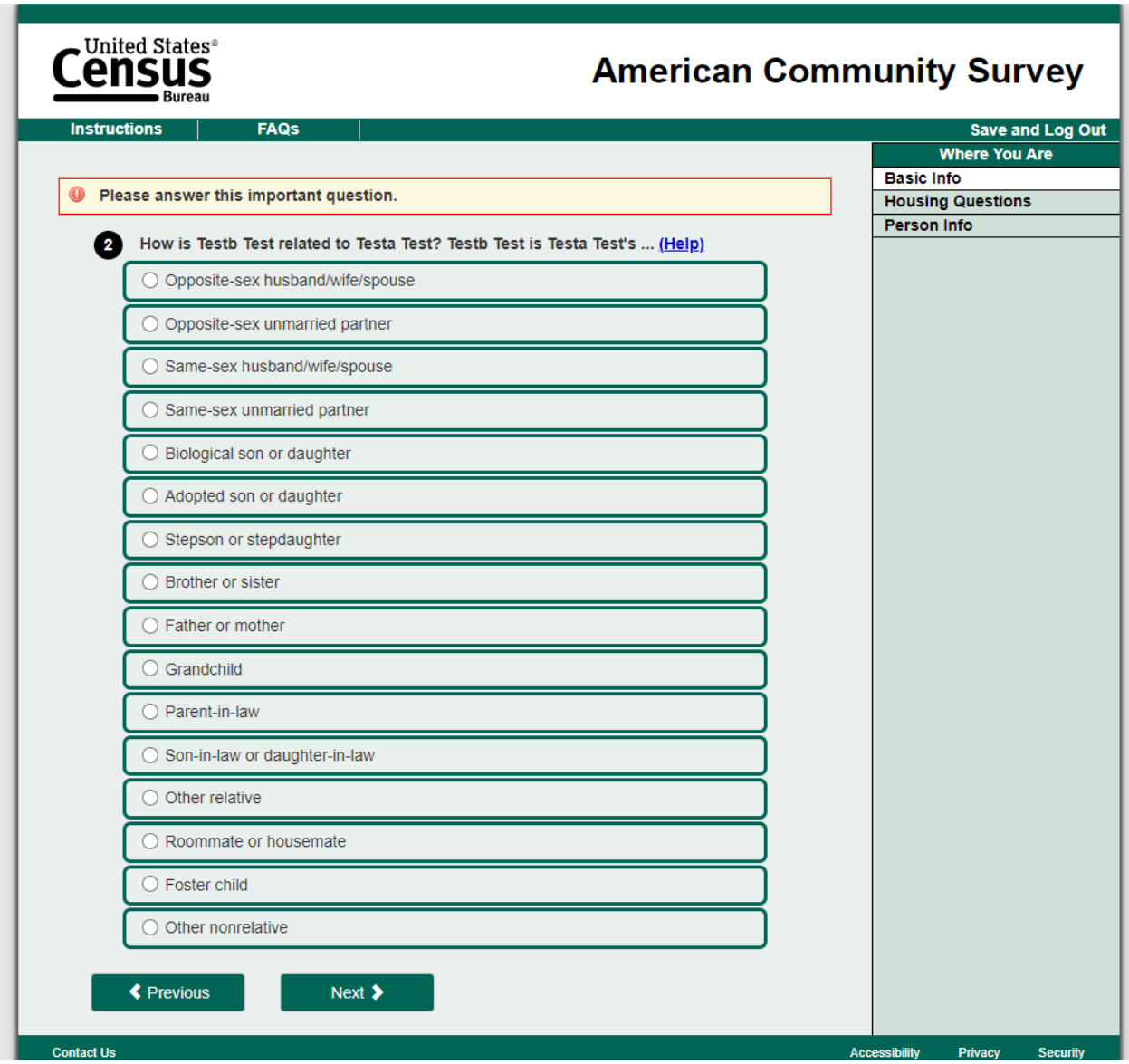
Respondent Name (T1)

No change



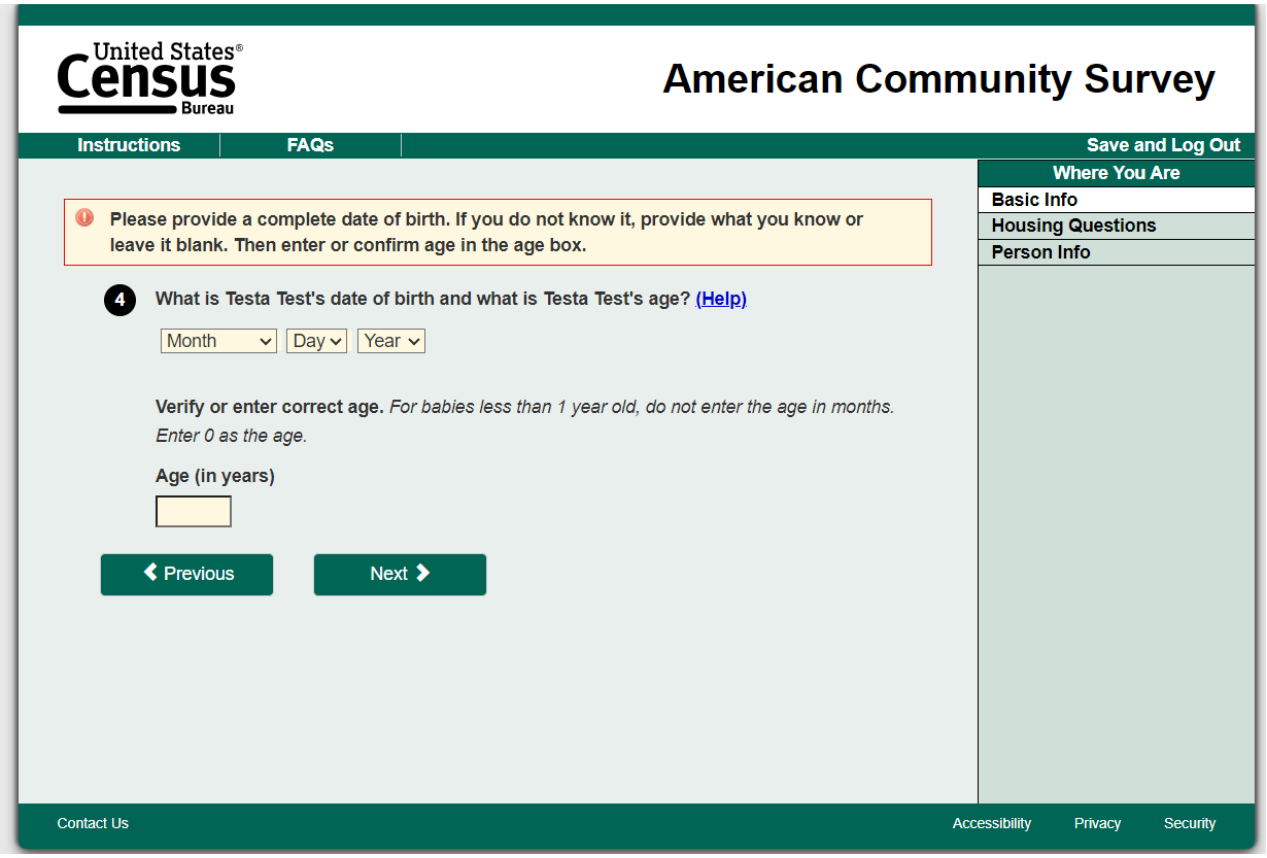
Relationship (T1)

Multiple Choice



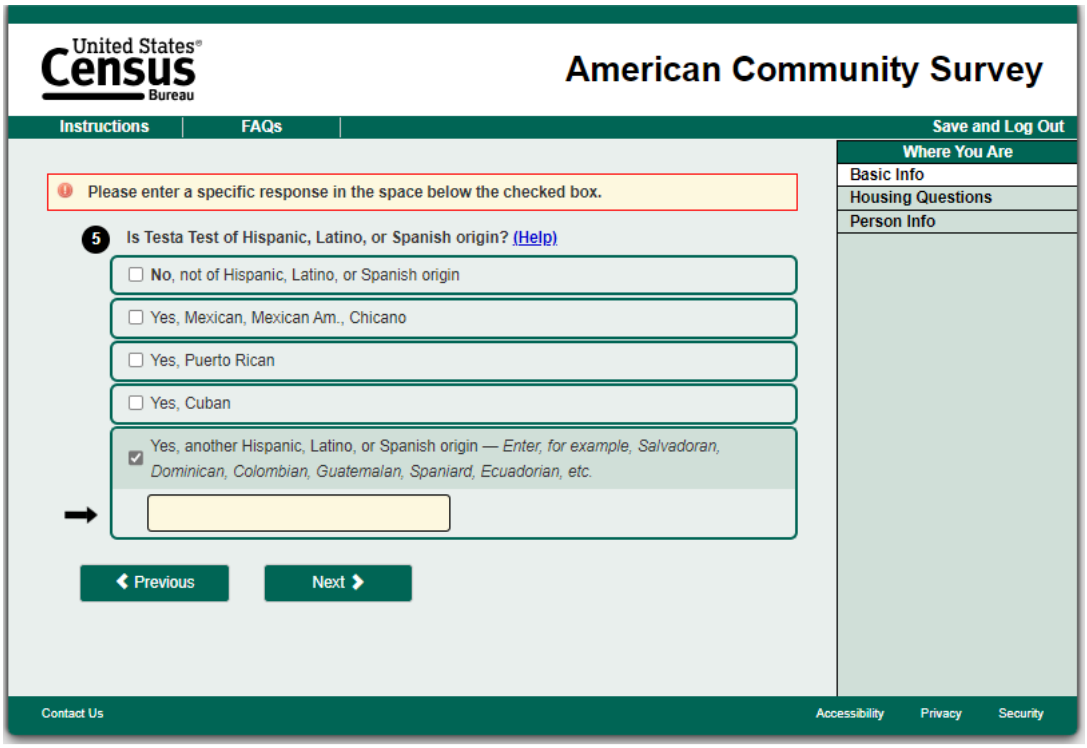
Date of Birth(T1)

No change



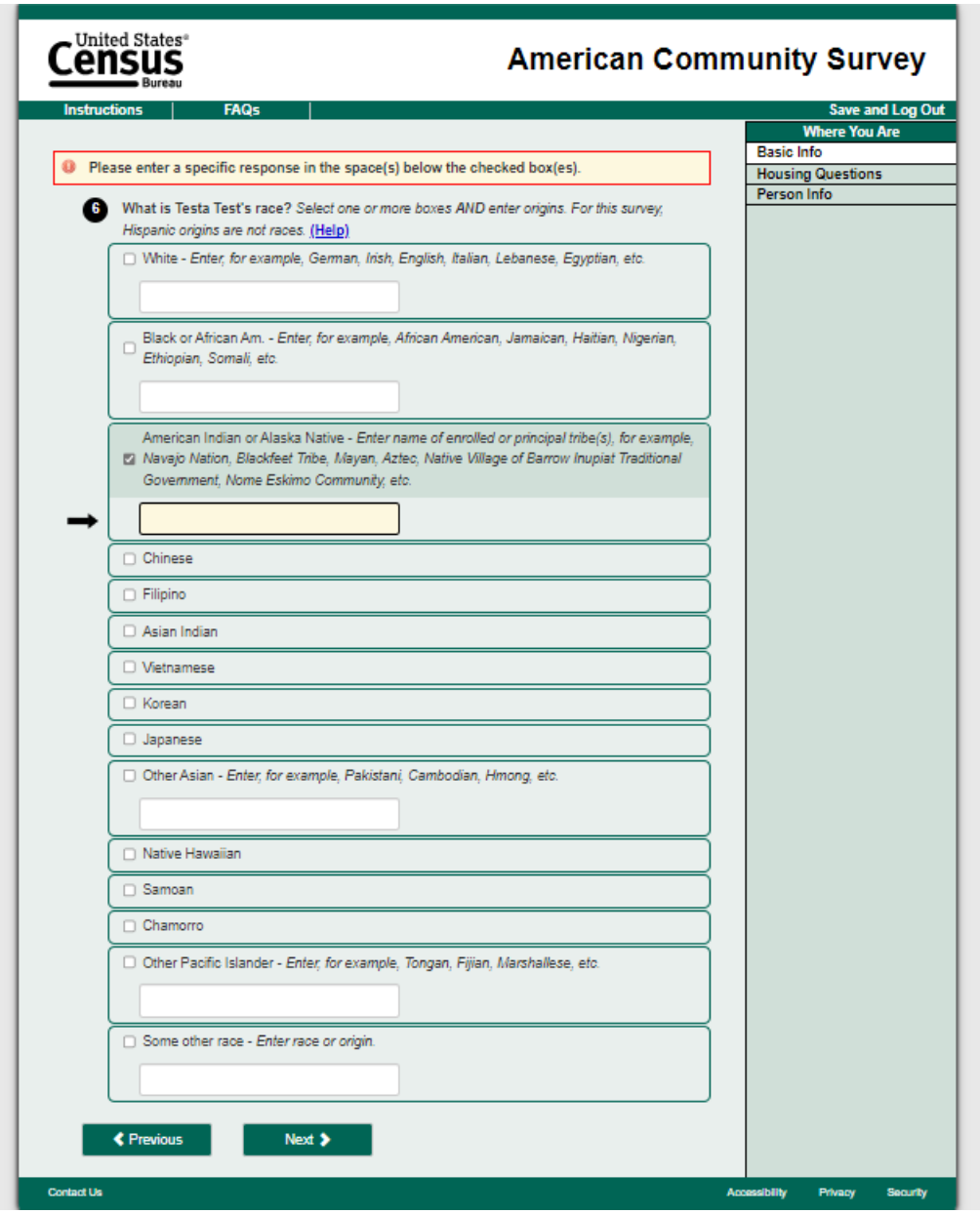
Hispanic (T1)

Select multiple, textbox, centered radio button



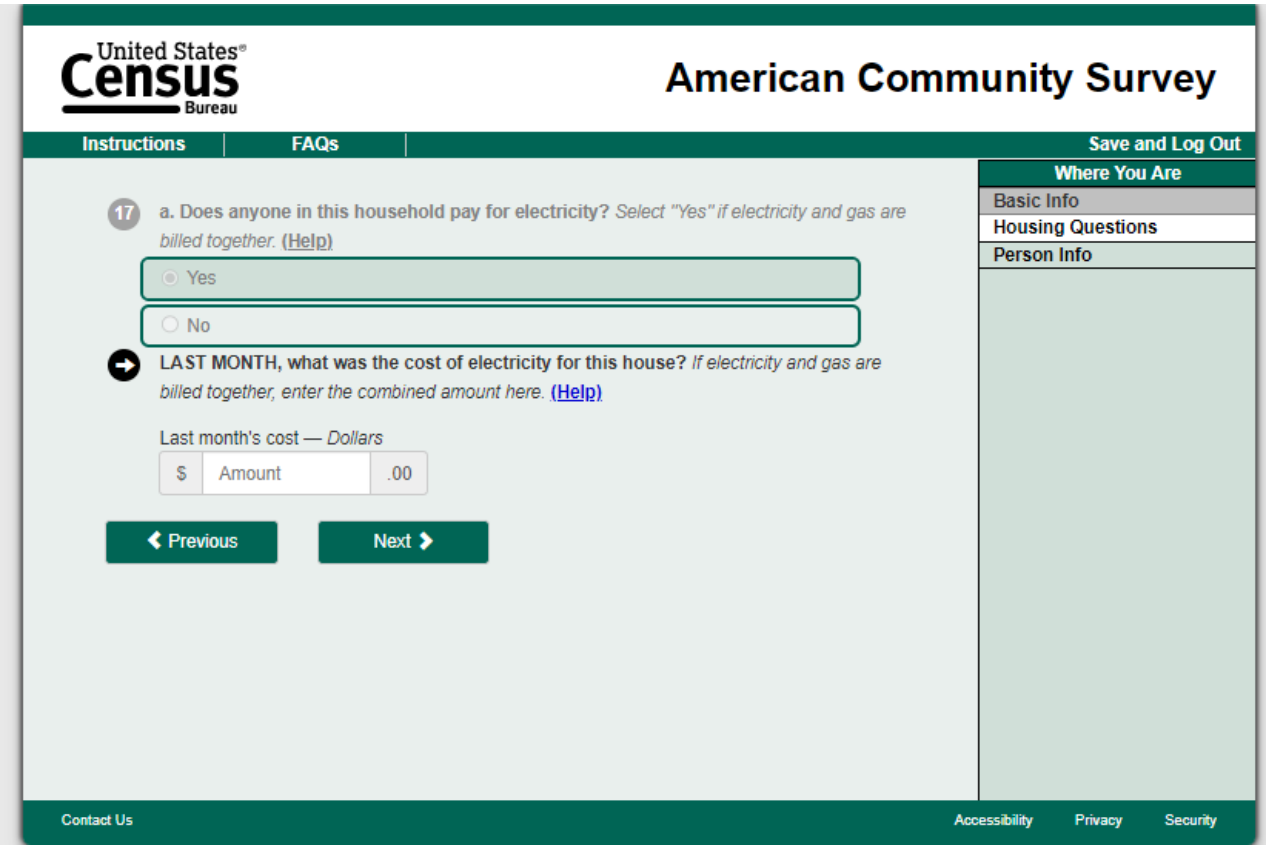
Race (T1)

Select multiple, textbox, centered radio button



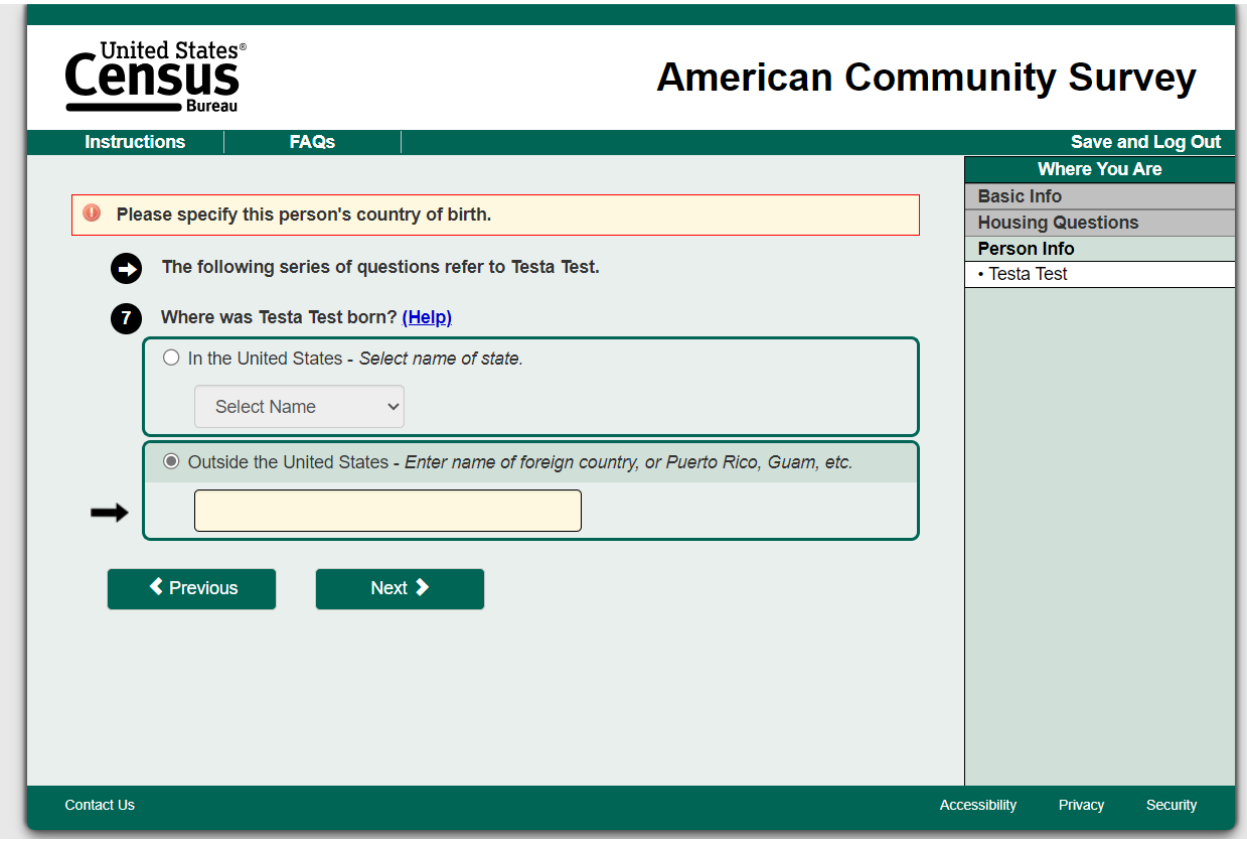
Electric Amount (T1)

Unfolded expanded answer field



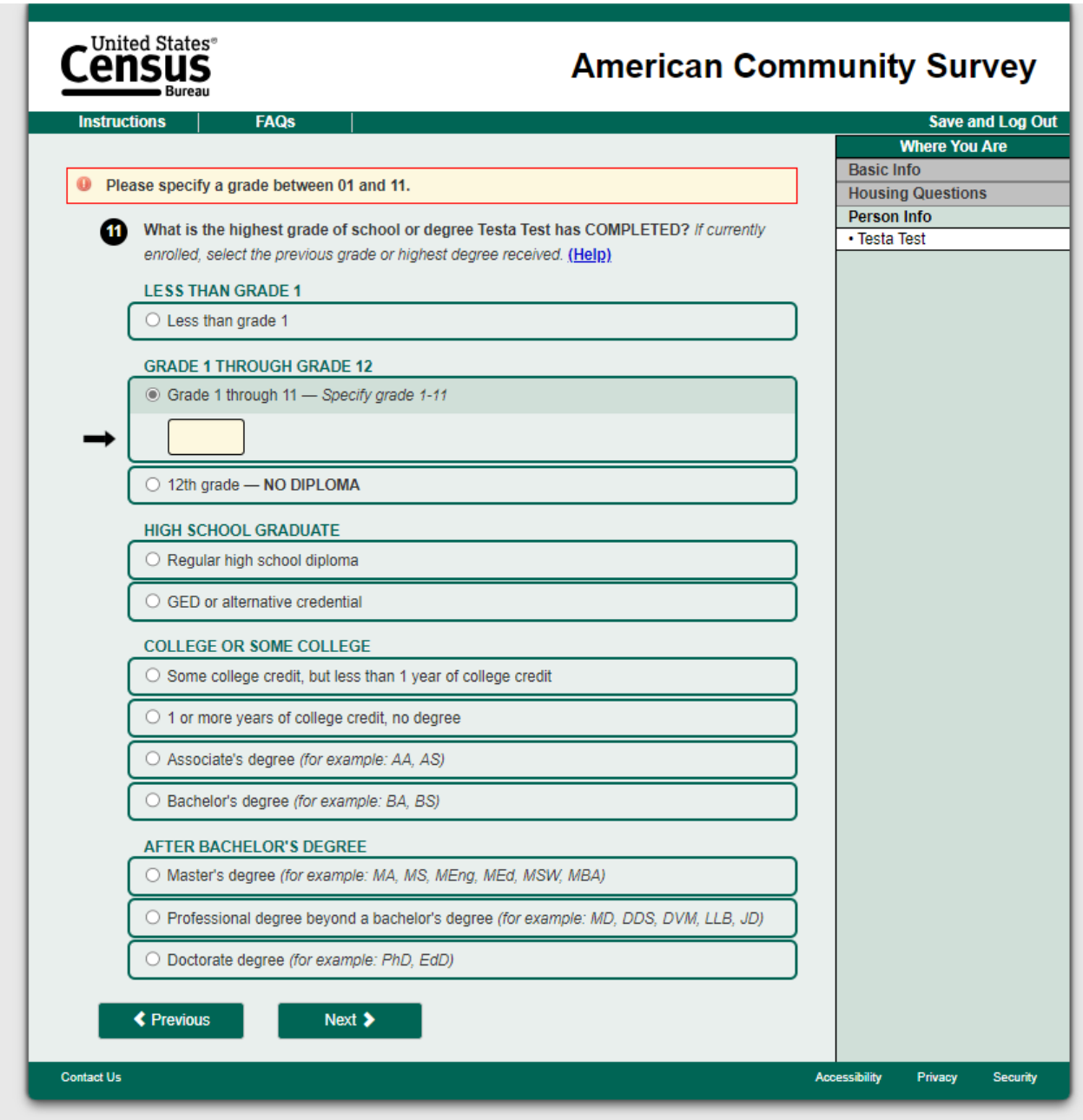
Place of Birth (T1)

Textbox and dropdown menu



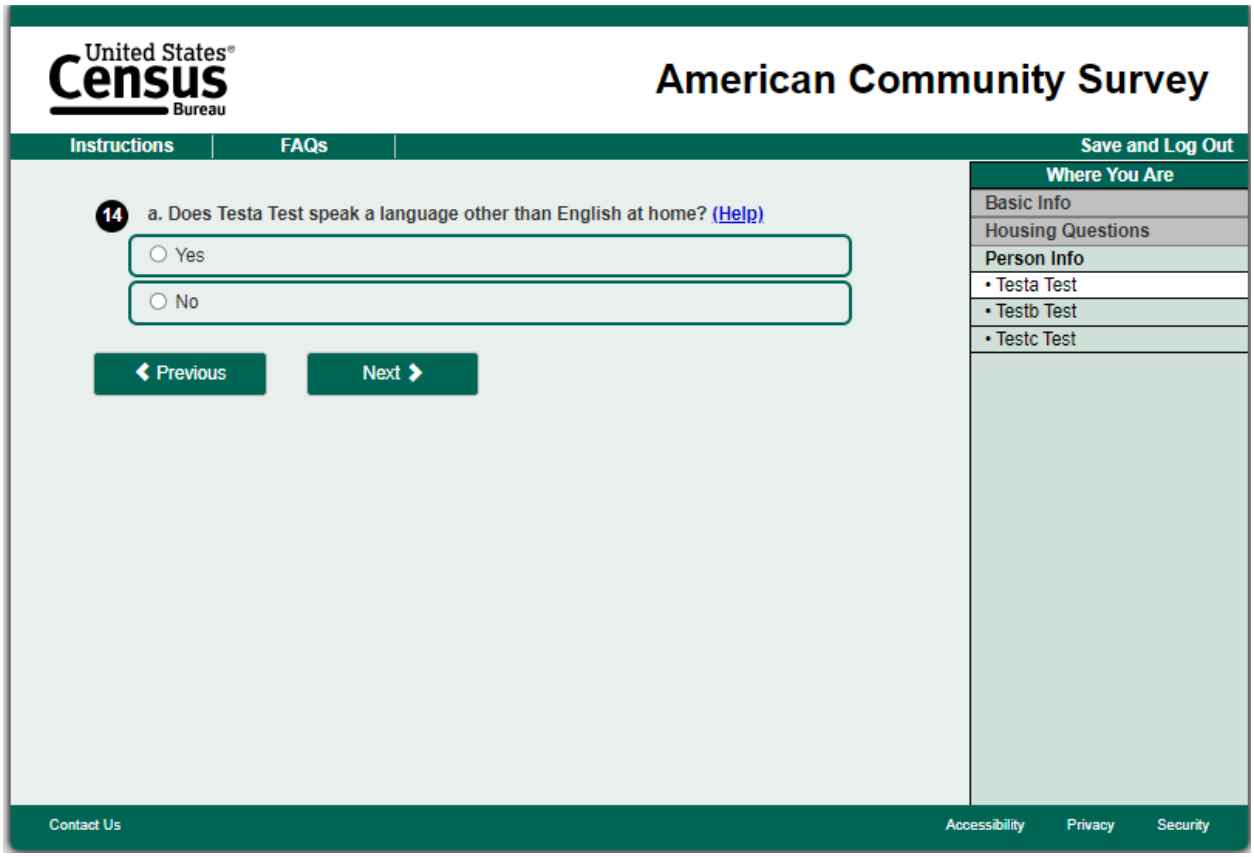
Highest Level (T1)

Multiple choice and textbox



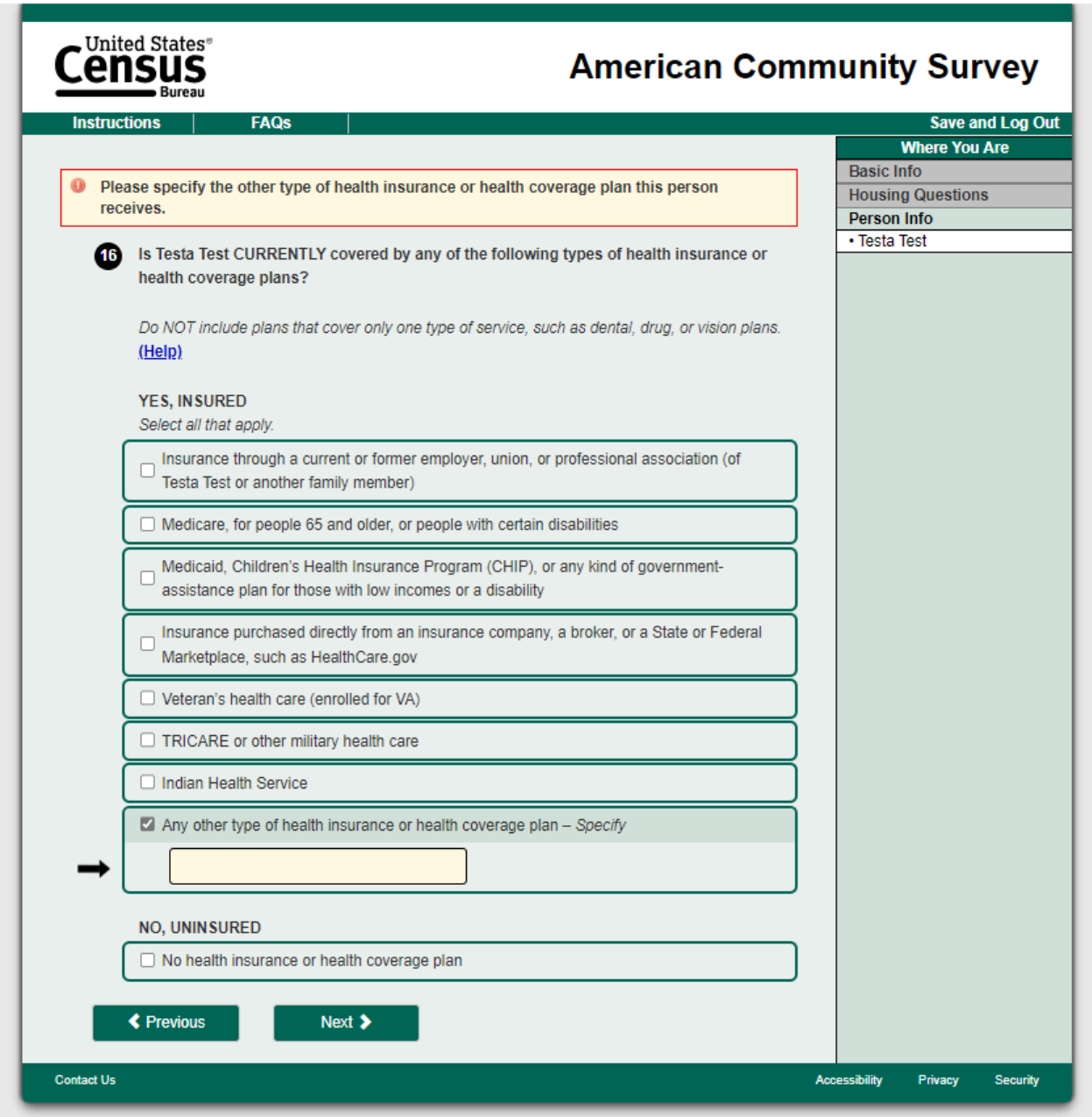
Language (T1)

Two options



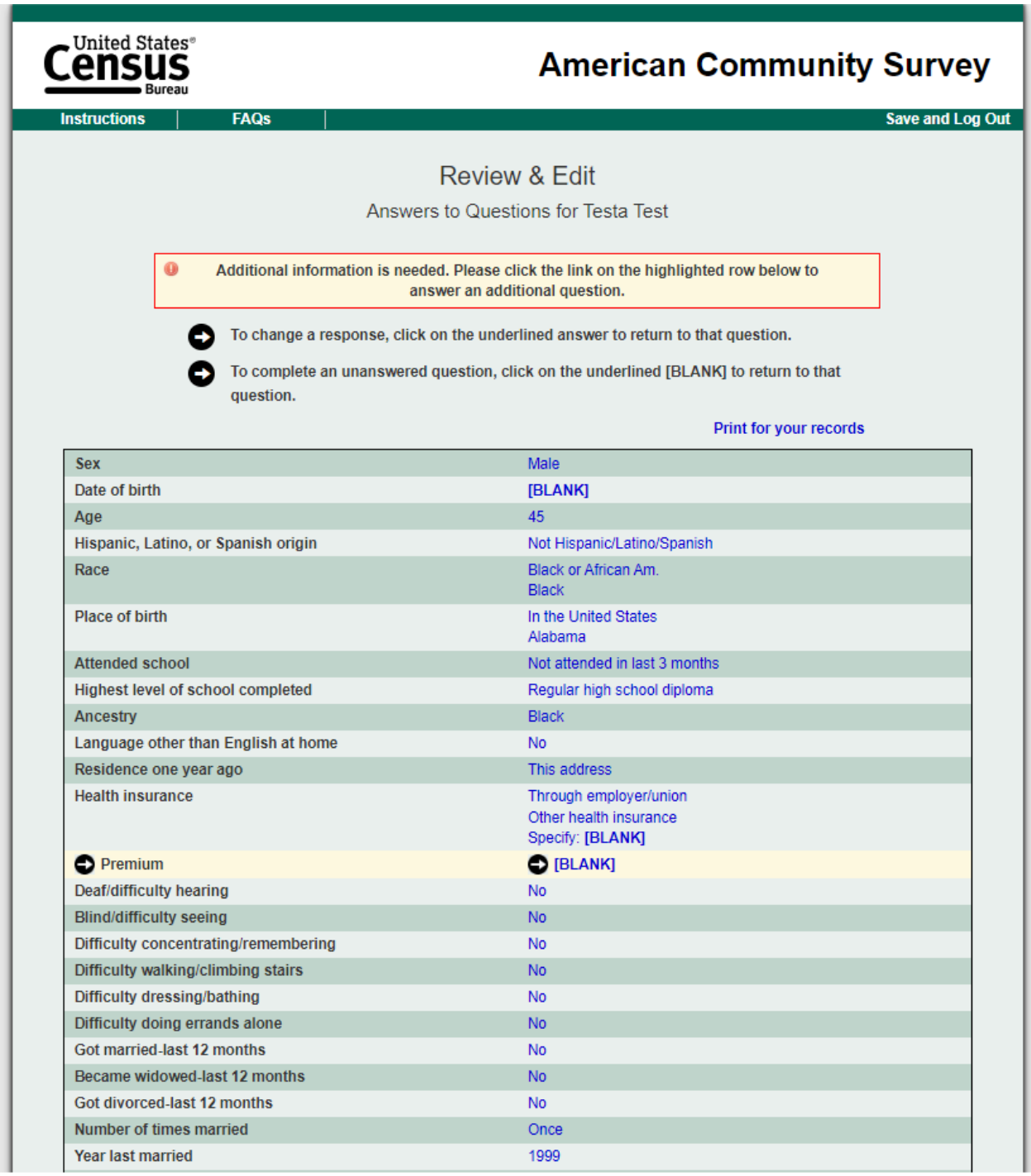
Insurance (T1)

Select multiple, textbox, centered radio button



Review (T1)

No change



United States Census Bureau

American Community Survey

Instructions | FAQs | Save and Log Out

Review & Edit

Answers to Questions for Testa Test

Additional information is needed. Please click the link on the highlighted row below to answer an additional question.

To change a response, click on the underlined answer to return to that question.

To complete an unanswered question, click on the underlined [BLANK] to return to that question.

Print for your records

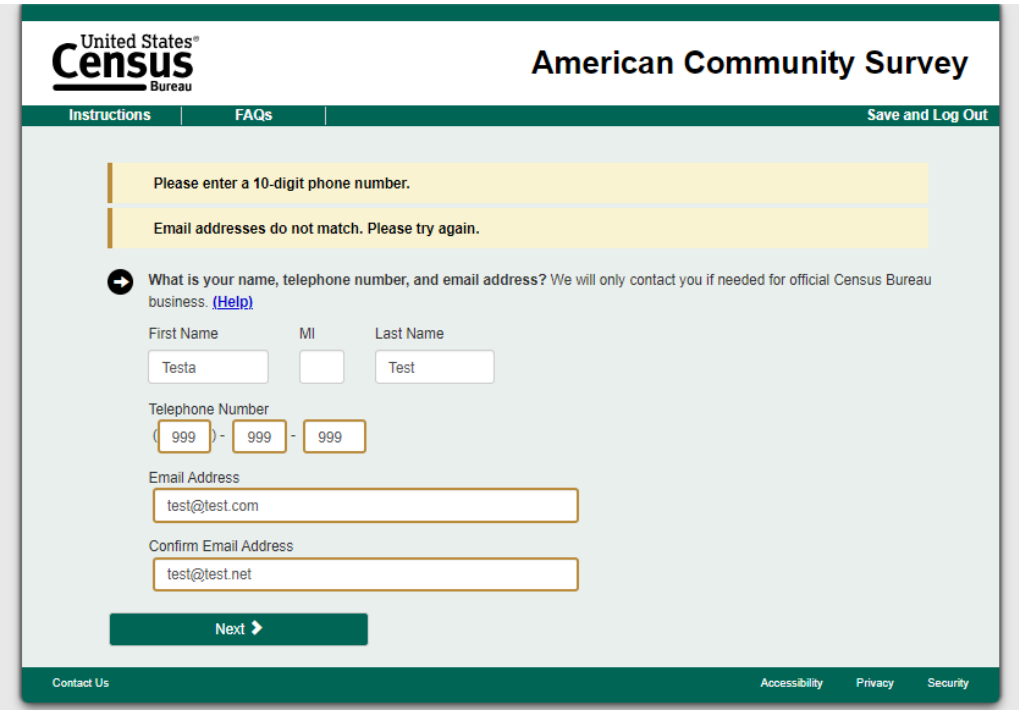
Sex	Male
Date of birth	[BLANK]
Age	45
Hispanic, Latino, or Spanish origin	Not Hispanic/Latino/Spanish
Race	Black or African Am. Black
Place of birth	In the United States Alabama
Attended school	Not attended in last 3 months
Highest level of school completed	Regular high school diploma
Ancestry	Black
Language other than English at home	No
Residence one year ago	This address
Health insurance	Through employer/union Other health insurance Specify: [BLANK]
Premium	[BLANK]
Deaf/difficulty hearing	No
Blind/difficulty seeing	No
Difficulty concentrating/remembering	No
Difficulty walking/climbing stairs	No
Difficulty dressing/bathing	No
Difficulty doing errands alone	No
Got married-last 12 months	No
Became widowed-last 12 months	No
Got divorced-last 12 months	No
Number of times married	Once
Year last married	1999

Treatment 2 (T2):

T2 will include changes all edit messages. This includes both edit messages and highlighted boxes when a write-in option is selected. Changes include: Banners featuring a yellow background, with a border only on the lefthand vertical edge, the arrows removed, the boxes highlighted with a dark yellow outline, and multiple edit messages broken up. 62 screens/questions will be impacted.

Respondent Name (T2)

Banner, box, multiple edit messages

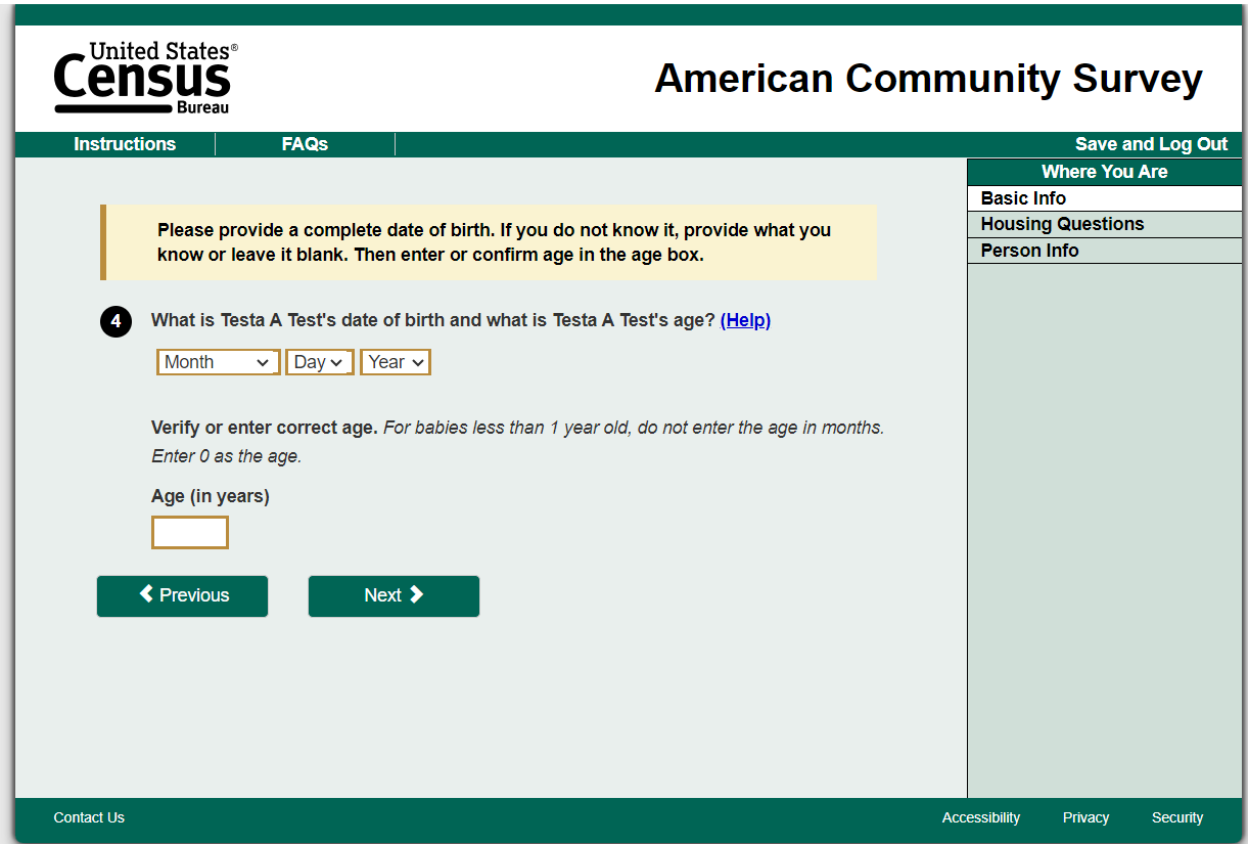


Relationship (T2)

Banner

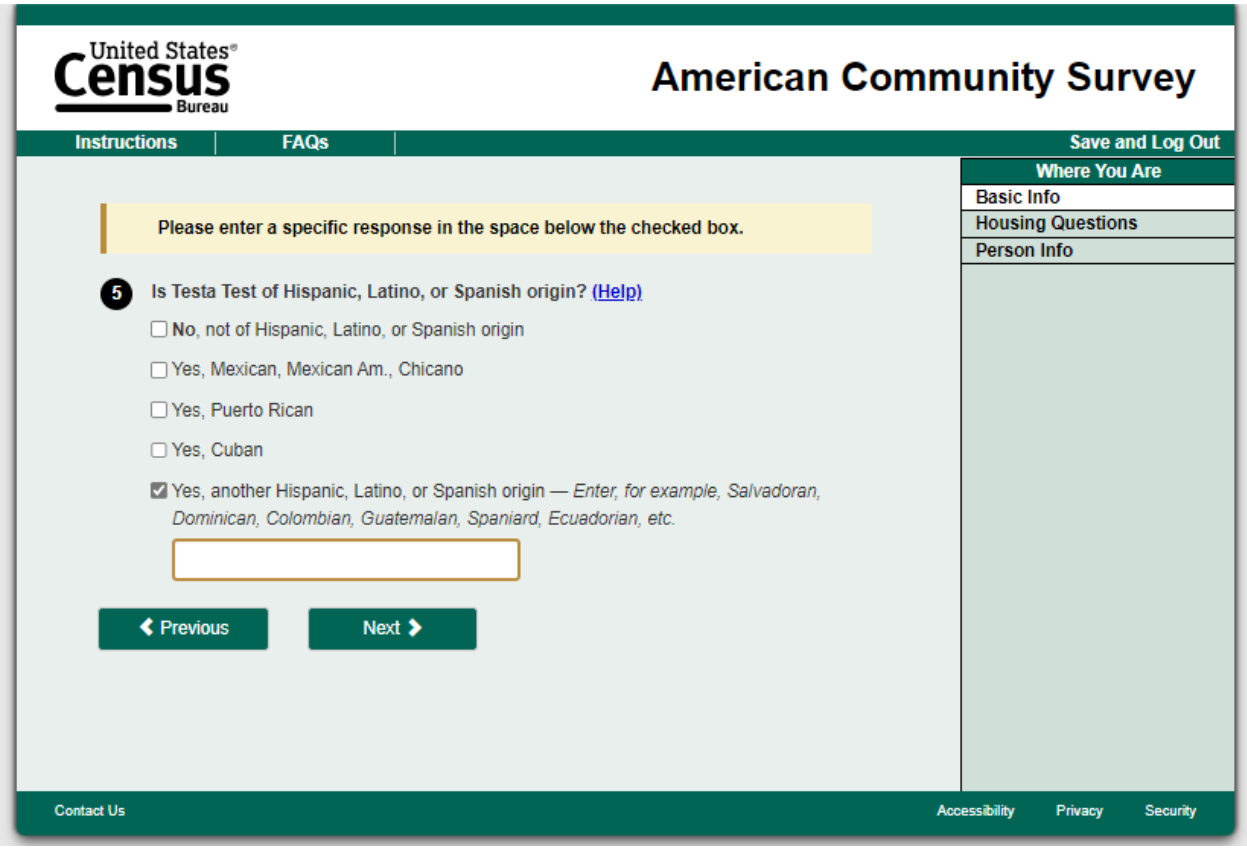
Date of Birth (T2)

Banner, boxes



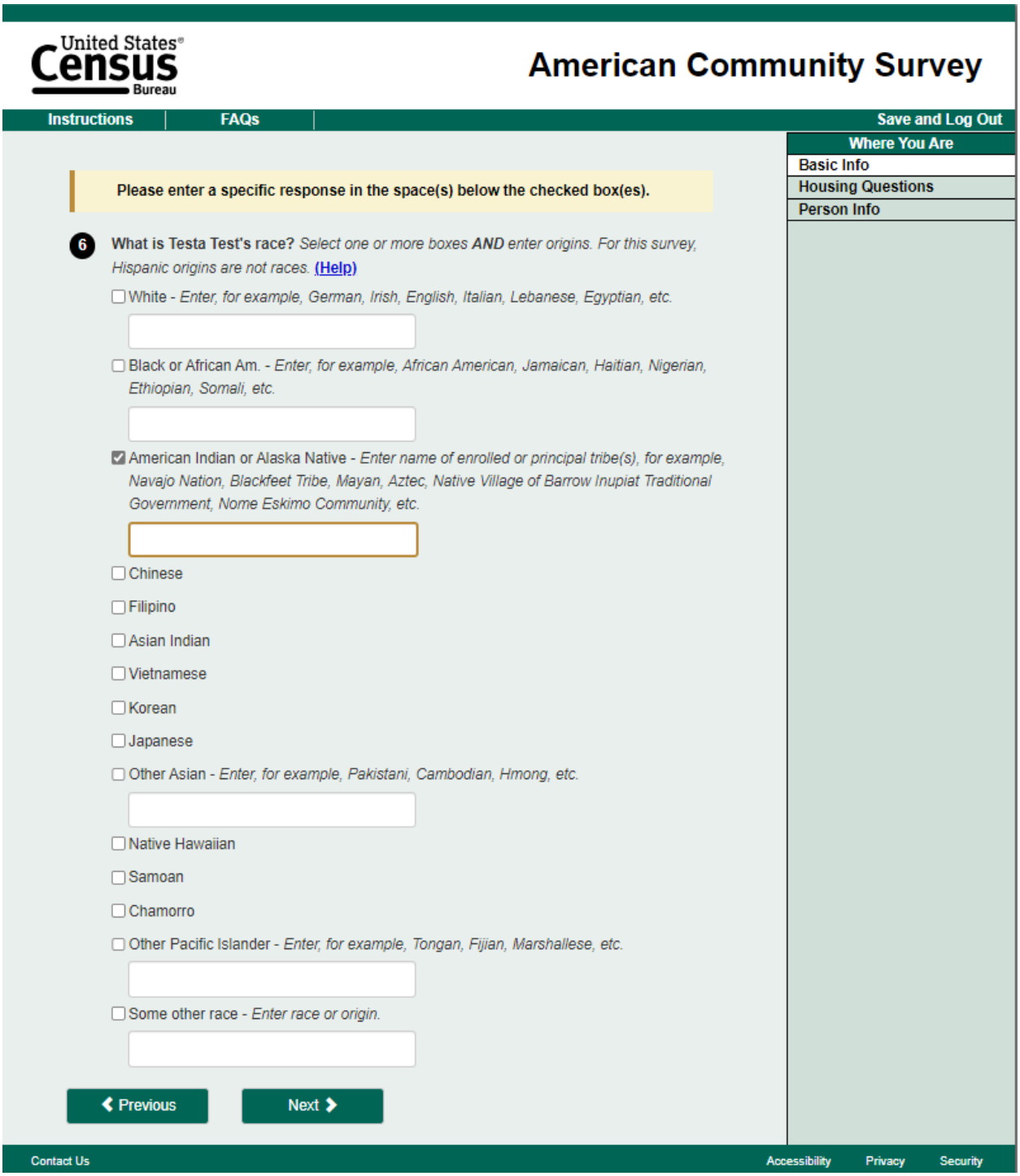
Hispanic (T2)

Banner, box, arrow removal



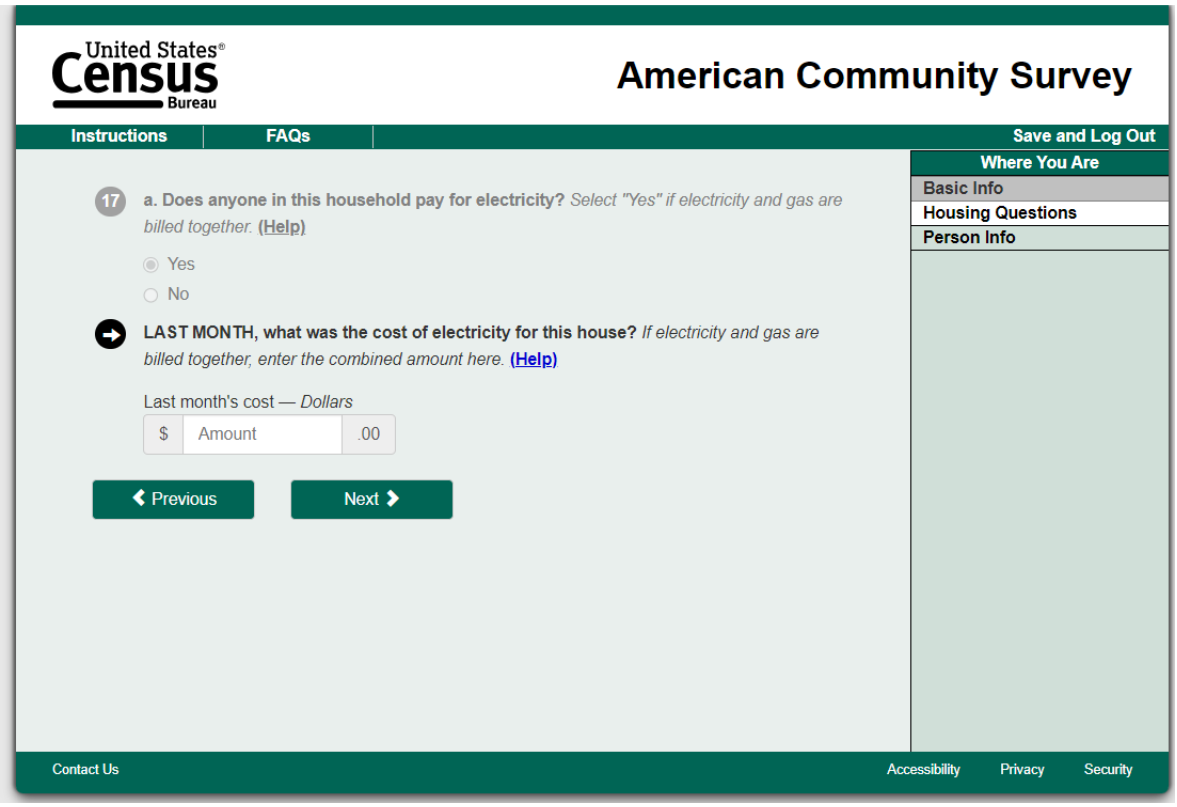
Race (T2)

Banner, box, arrow removal



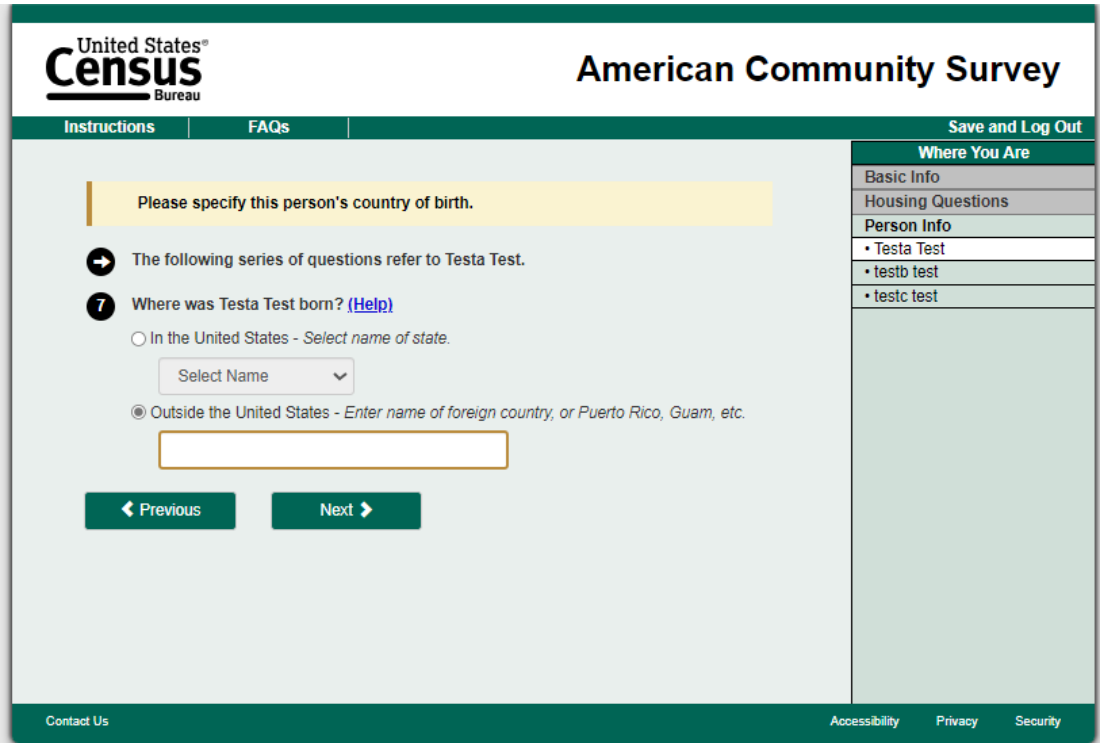
Electric Amount (T2)

No change



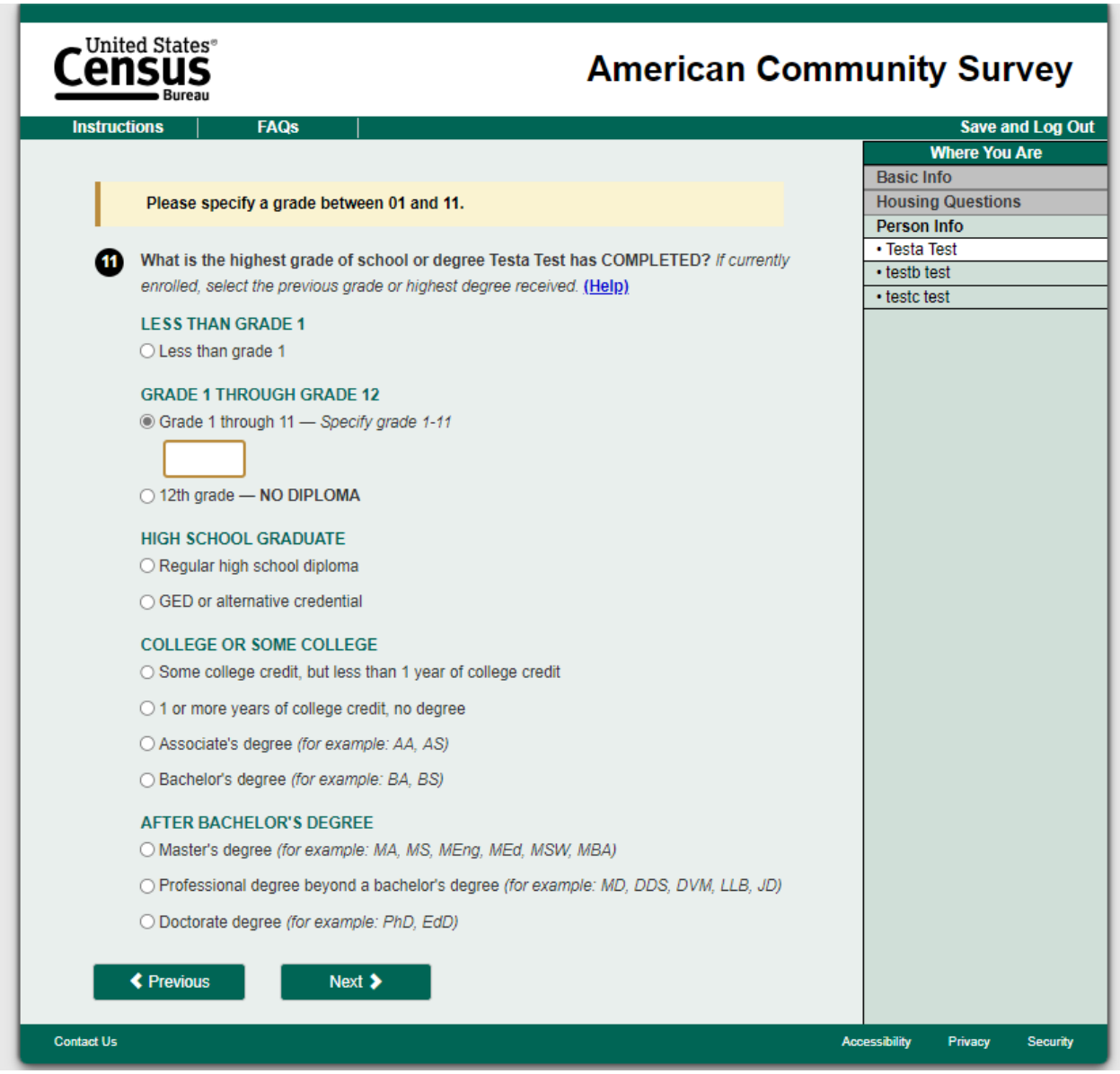
Place of Birth(T2)

Banner, box, arrow removal



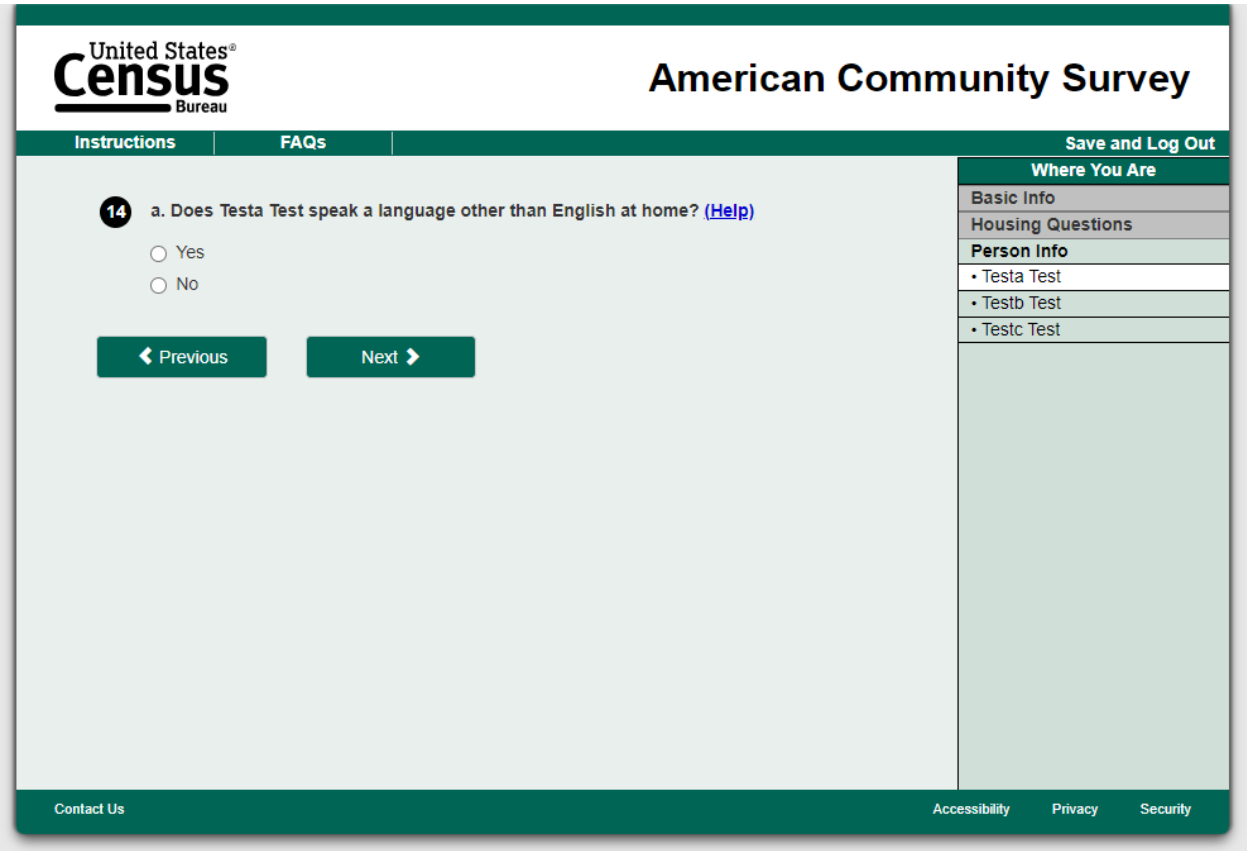
Highest Level (T2)

Banner, box, arrow removal



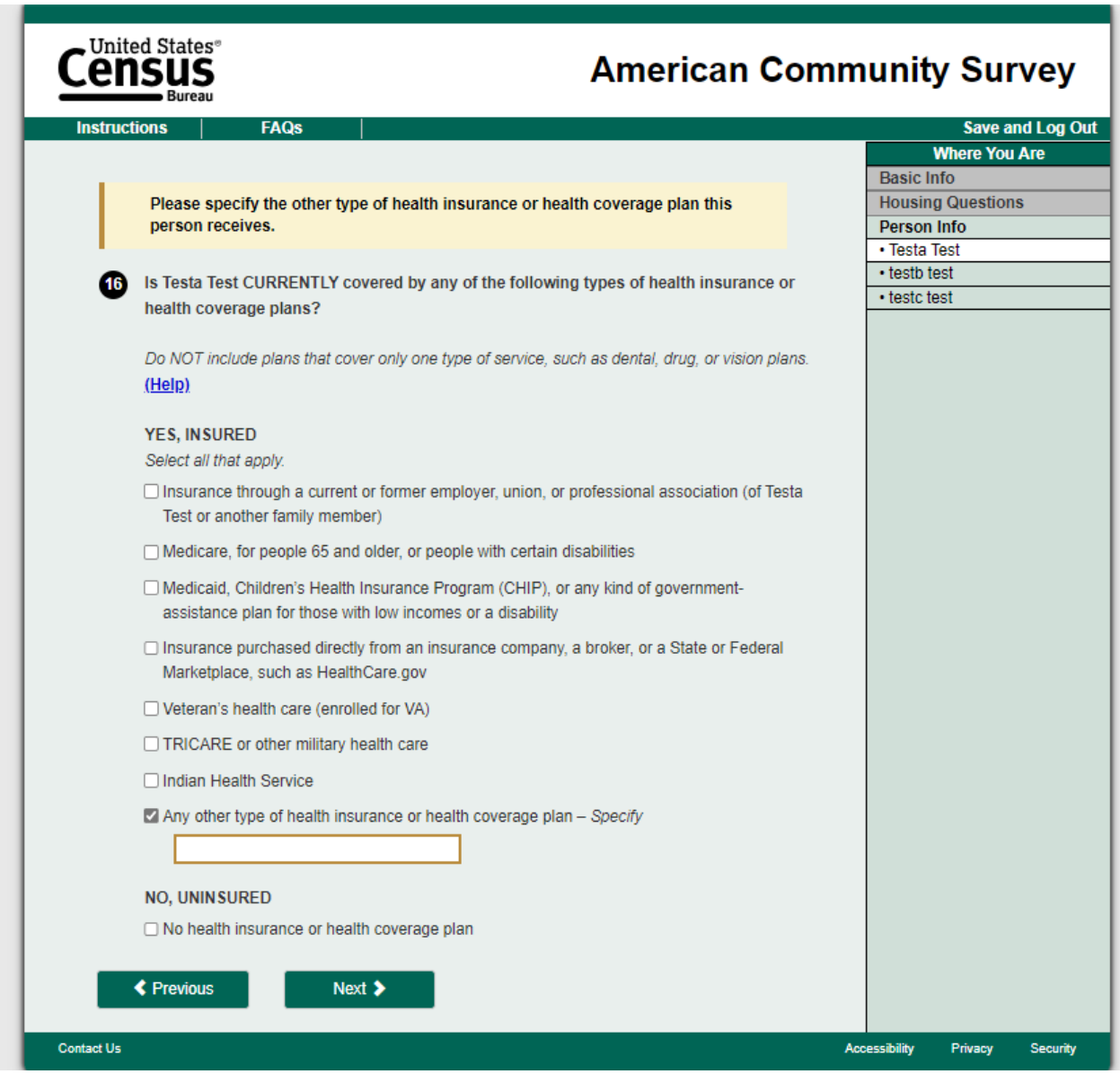
Language (T2)

No change



Insurance (T2)

Banner, box, arrow removal



Review (T2)

Banner and row

United States Census Bureau

American Community Survey

Instructions | FAQs | Save and Log Out

Review & Edit

Answers to Questions for Testa A Test

Additional information is needed. Please click the link on the highlighted row below to answer an additional question.

To change a response, click on the underlined answer to return to that question.

To complete an unanswered question, click on the underlined [BLANK] to return to that question.

Print for your records

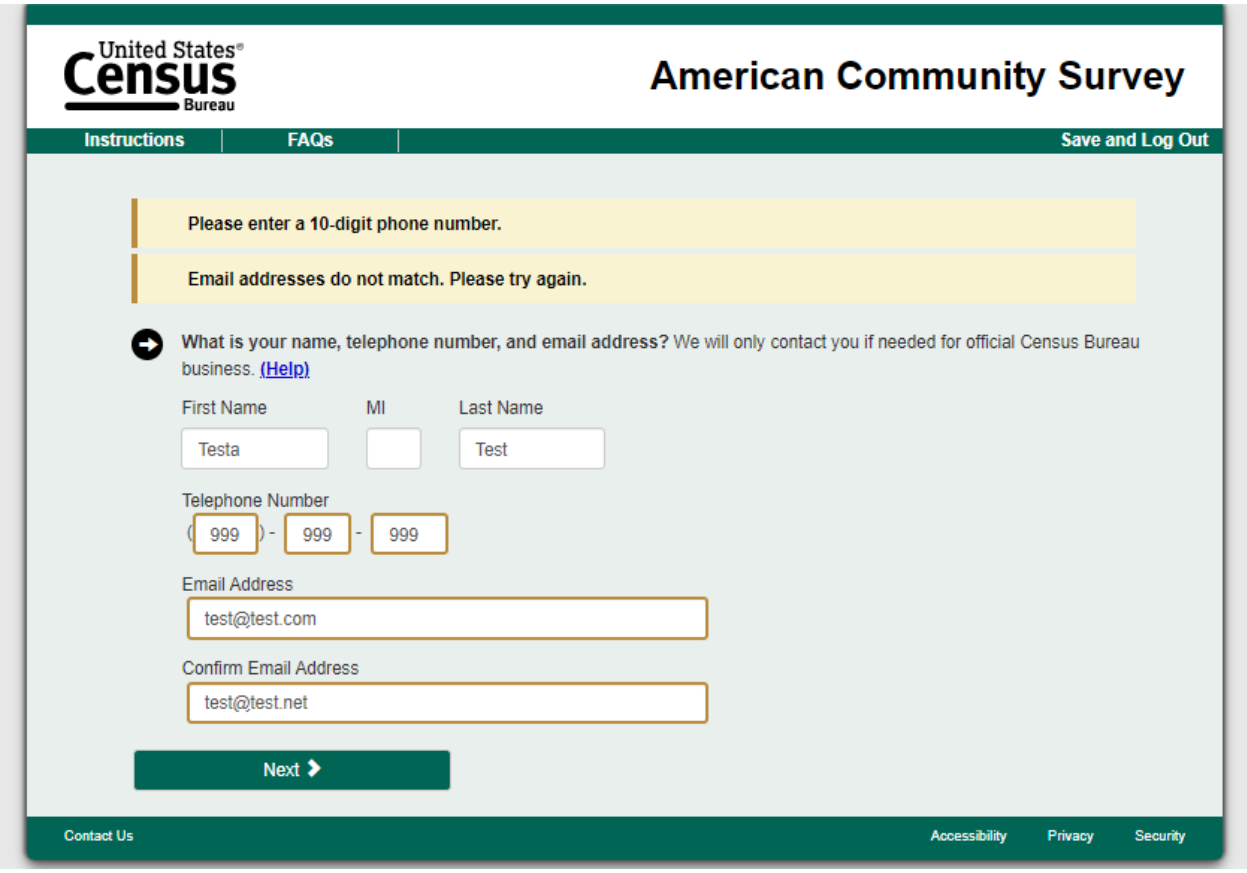
Sex	Male
Date of birth	06/09/2004
Age	20
Hispanic, Latino, or Spanish origin	Other Hispanic/Latino/Spanish
Race	Black or African Am.
Place of birth	Outside the United States dominican
Citizenship	Naturalized U.S. citizen (1999)
Year came to live in U.S.	2005
Attended school	Public school/college
Grade level attended	[BLANK]
Highest level of school completed	Some college-less than 1 year
Ancestry	[BLANK]
Language other than English at home	No
Residence one year ago	This address
Health insurance	Through employer/union
Deaf/difficulty hearing	[BLANK]
Blind/difficulty seeing	[BLANK]
Difficulty concentrating/remembering	[BLANK]
Difficulty walking/climbing stairs	[BLANK]
Difficulty dressing/bathing	[BLANK]
Difficulty doing errands alone	[BLANK]
Marital status	Now married
Got married-last 12 months	[BLANK]
Serve on active duty	Never served

Treatment (T3):

Treatment 3 will include both Treatment 1 and Treatment 2 changes. Some questions are impacted by both sets of changes and will include both answer field changes and error testing. 155 screens/questions will be impacted.

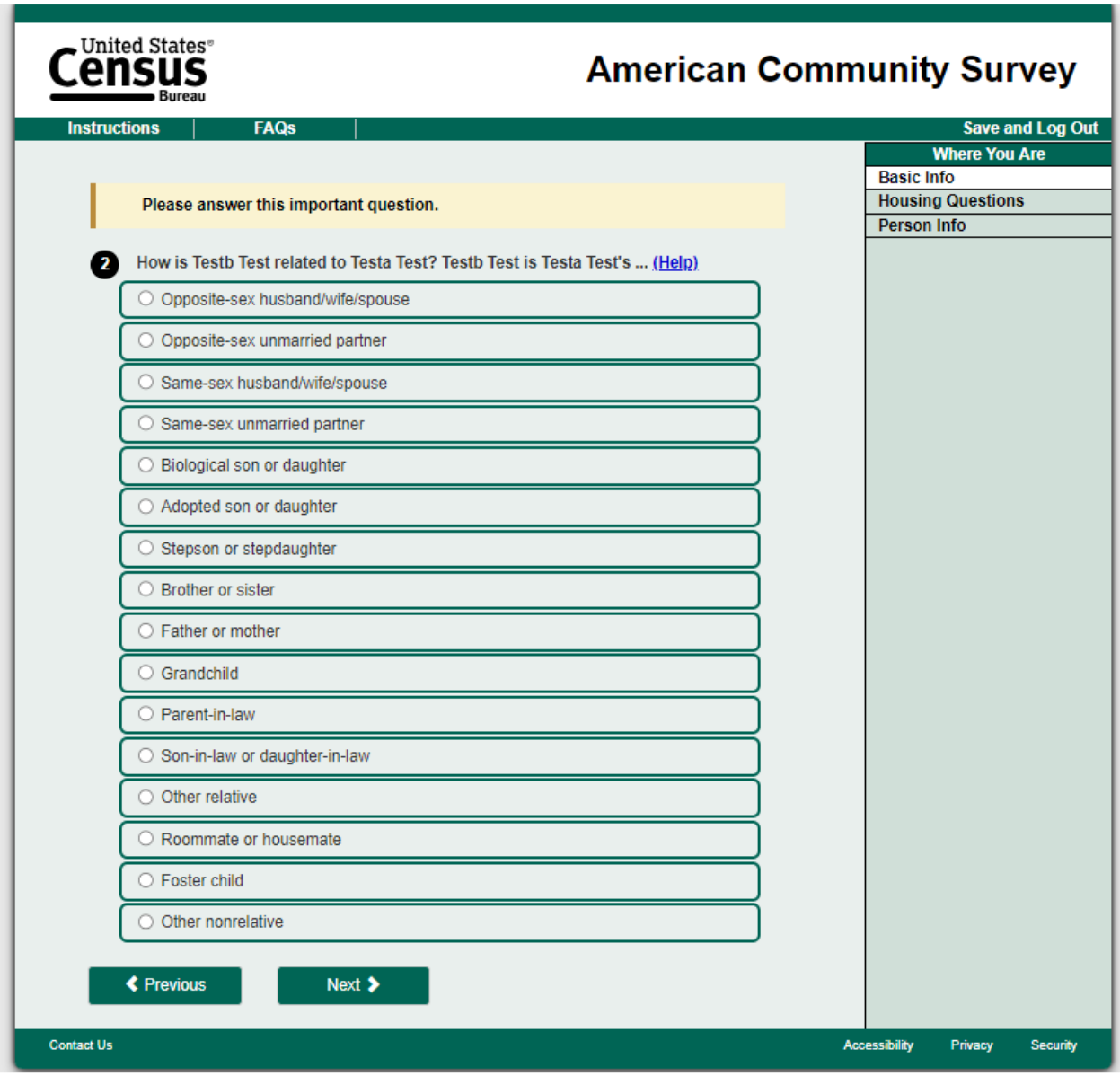
Respondent Name (T3)

Banner, box, multiple edit messages



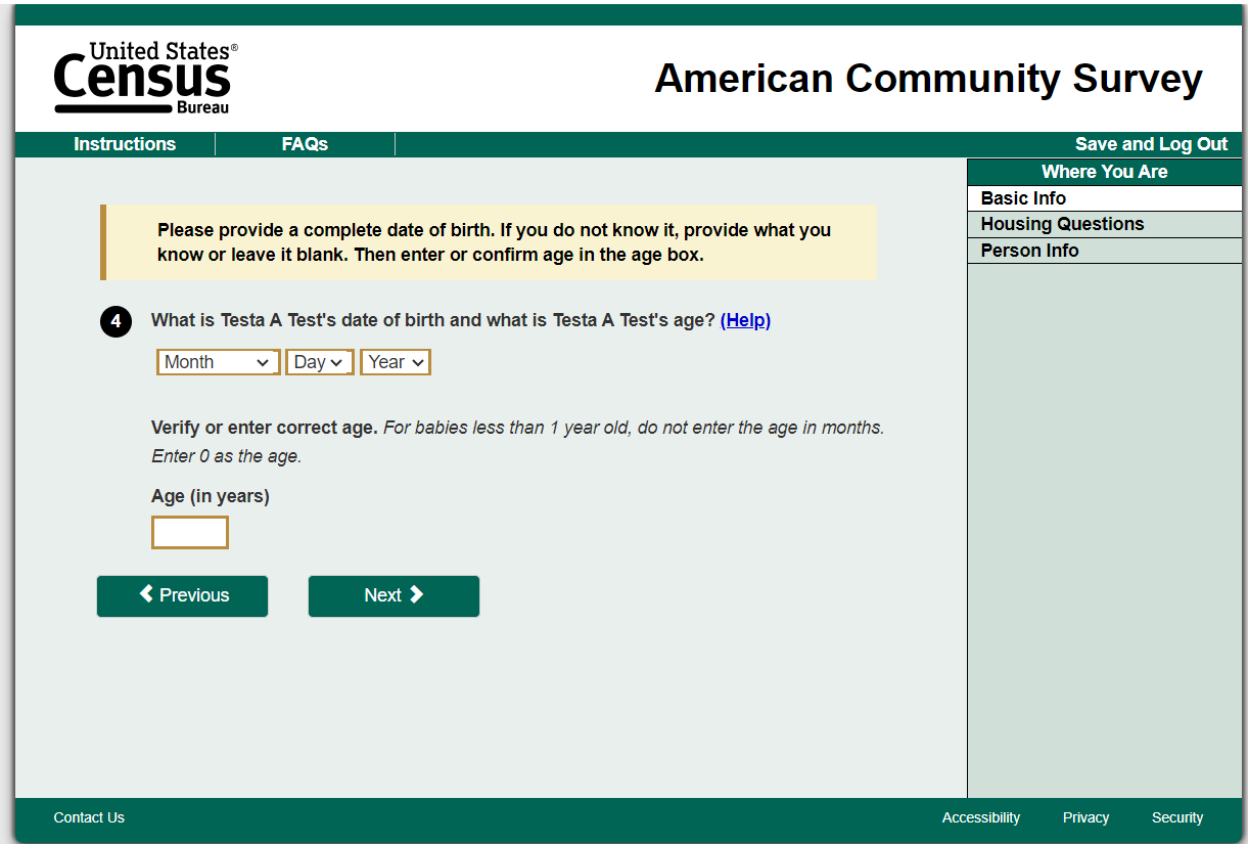
Relationship (T3)

Multiple choice, banner



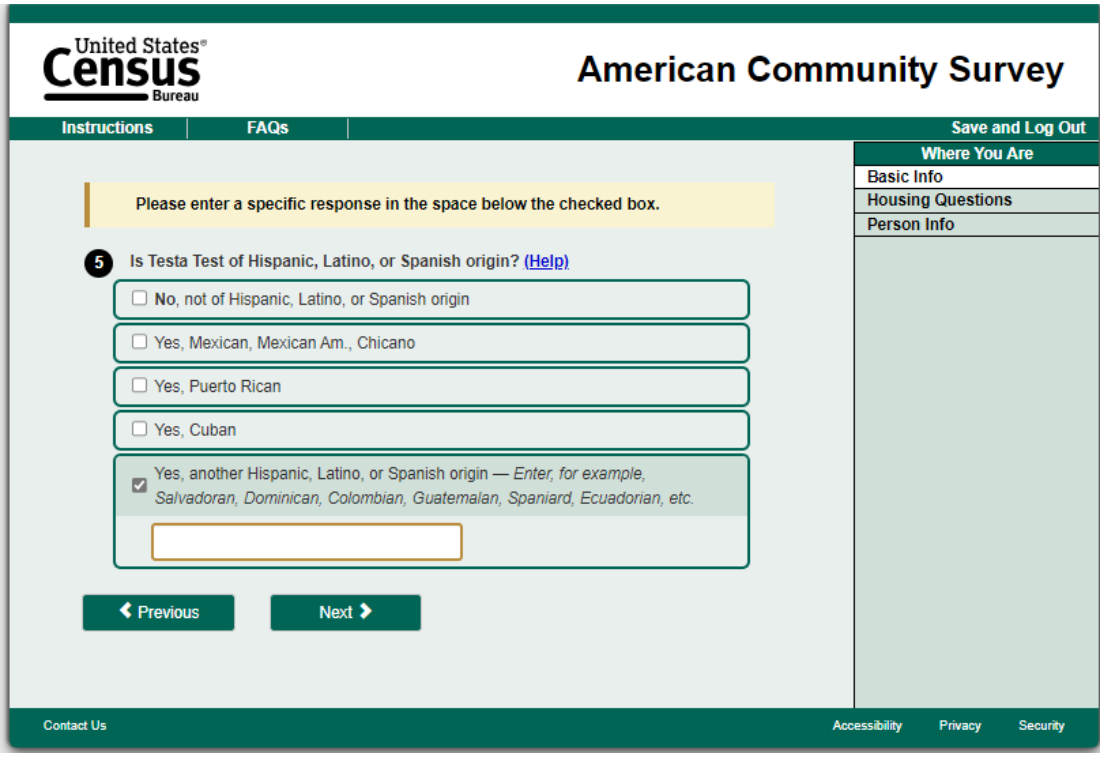
Date of Birth (T3)

Banner, boxes



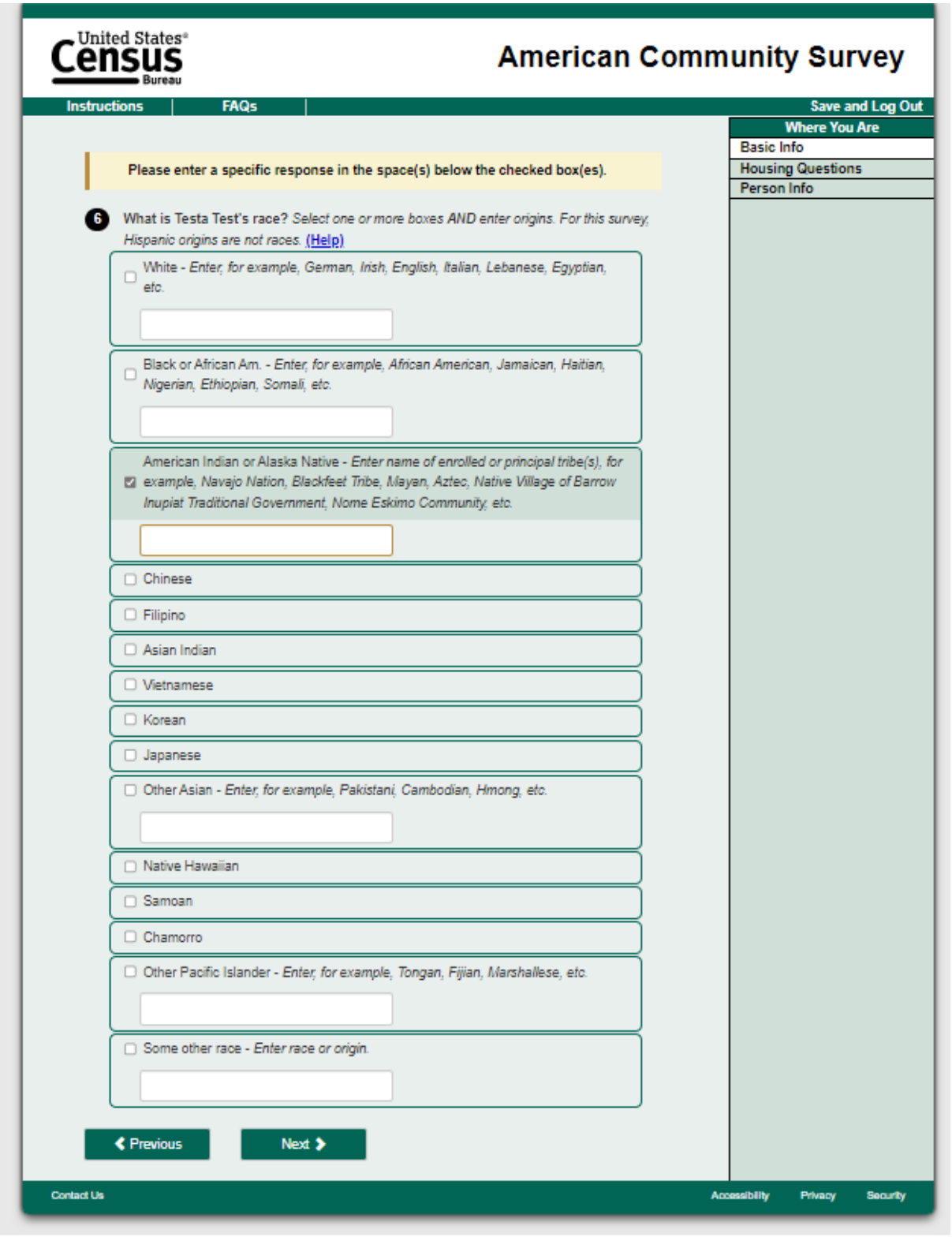
Hispanic (T3)

Banner, box, arrow removal



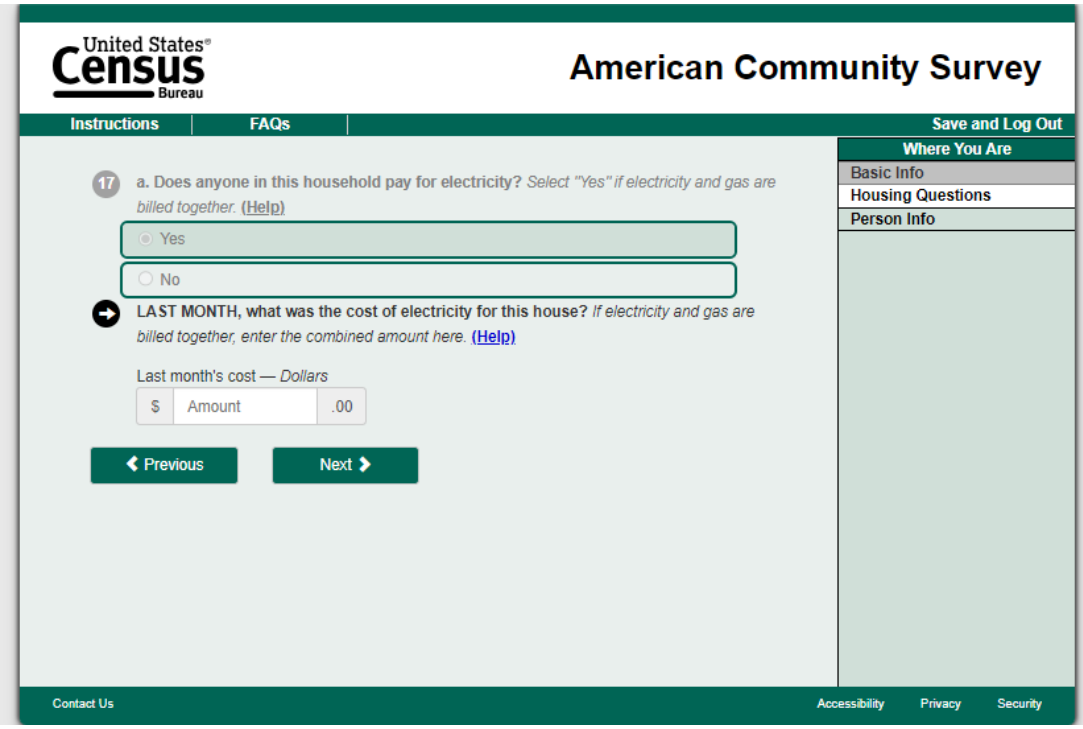
Race (T3)

Select multiple, textbox, centered radio button, banner, box, arrow removal


United States Census Bureau
American Community Survey
Instructions
FAQs
Save and Log Out
Where You Are
Basic Info
Housing Questions
Person Info
Please enter a specific response in the space(s) below the checked box(es).
6 What is Testa Test's race? Select one or more boxes AND enter origins. For this survey, Hispanic origins are not races. (Help)
White - Enter, for example, German, Irish, English, Italian, Lebanese, Egyptian, etc.
Black or African Am. - Enter, for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, etc.
American Indian or Alaska Native - Enter name of enrolled or principal tribe(s), for example, Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Native Village of Barrow Inupiat Traditional Government, Nome Eskimo Community, etc.
Chinese
Filipino
Asian Indian
Vietnamese
Korean
Japanese
Other Asian - Enter, for example, Pakistani, Cambodian, Hmong, etc.
Native Hawaiian
Samoan
Chamorro
Other Pacific Islander - Enter, for example, Tongan, Fijian, Marshallese, etc.
Some other race - Enter race or origin.
Previous
Next
Contact Us
Accessibility
Privacy
Security

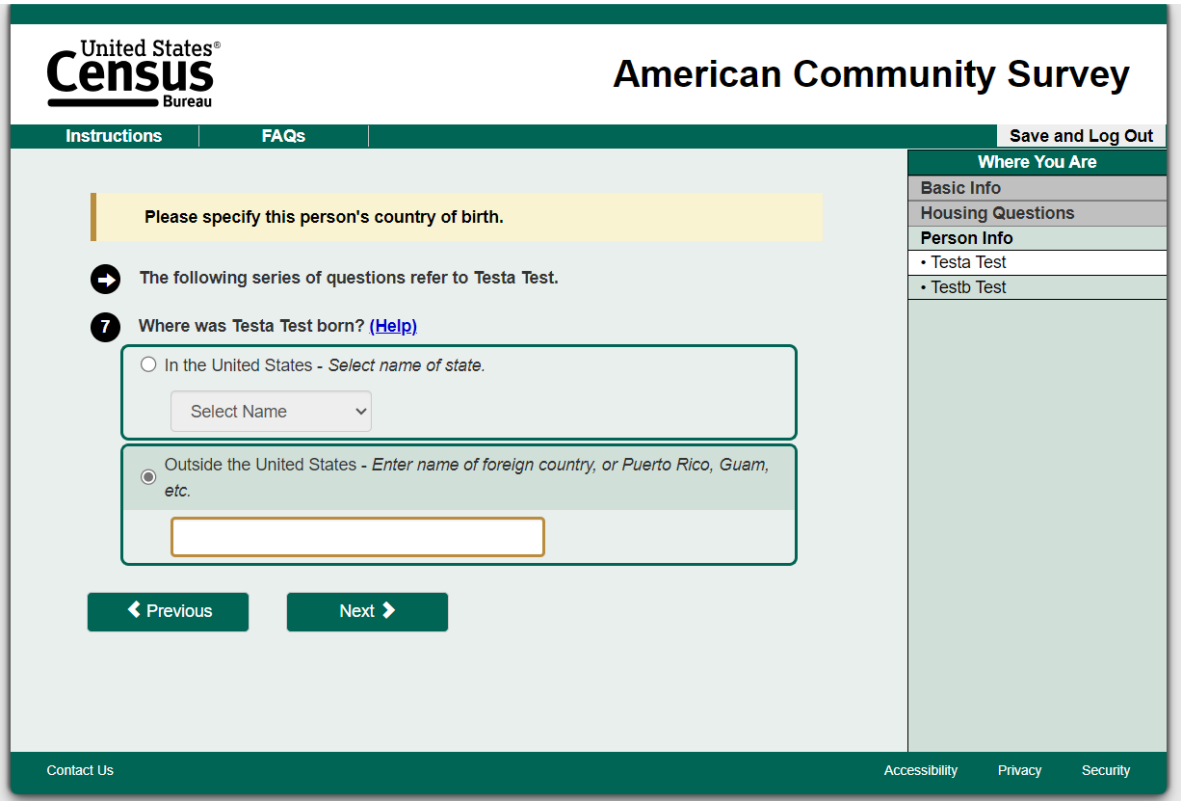
Electric Amount (T3)

Unfolded expanded answer field



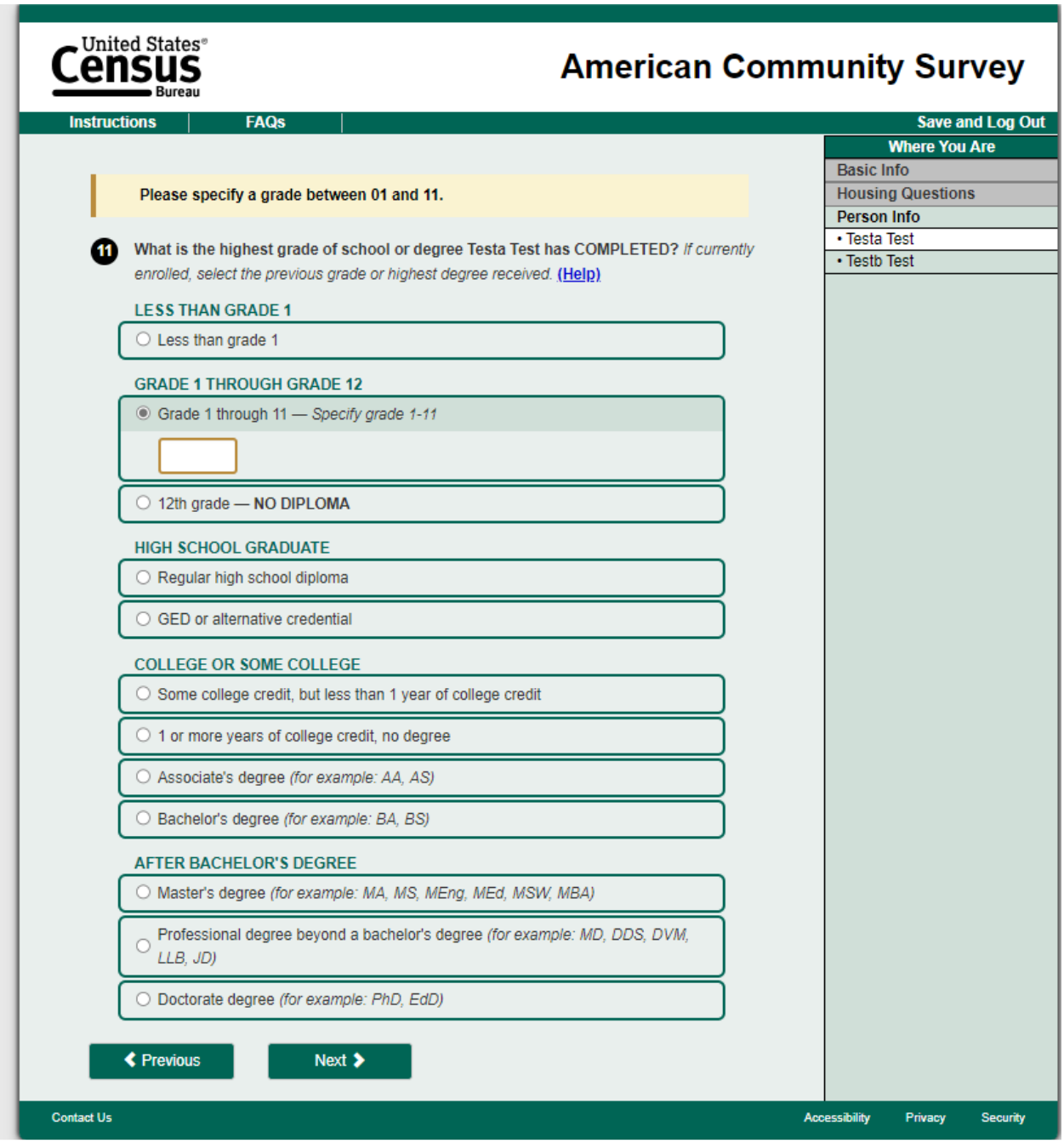
Place of Birth (T3)

Textbox, dropdown menu, banner, box, arrow removal



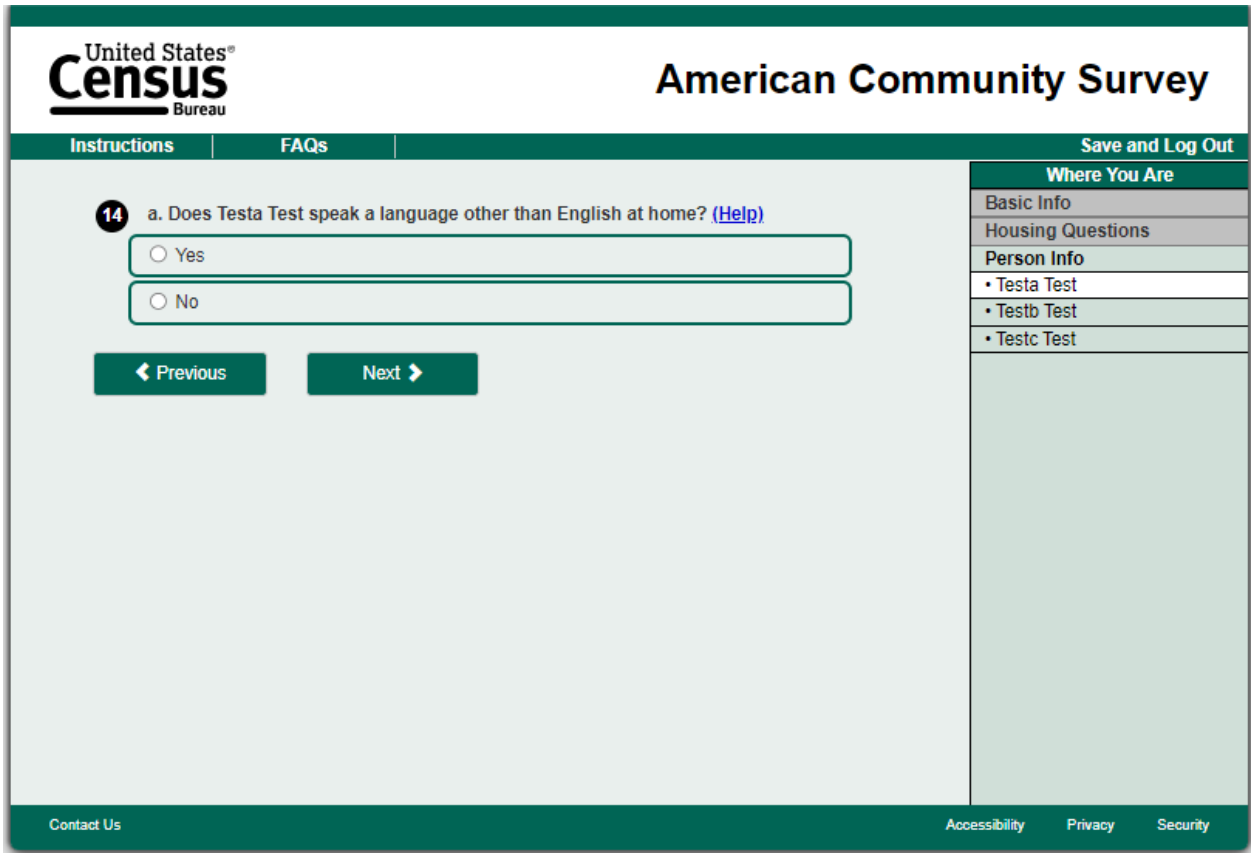
Highest Level (T3)

Multiple choice, textbox, banner, box, arrow removal



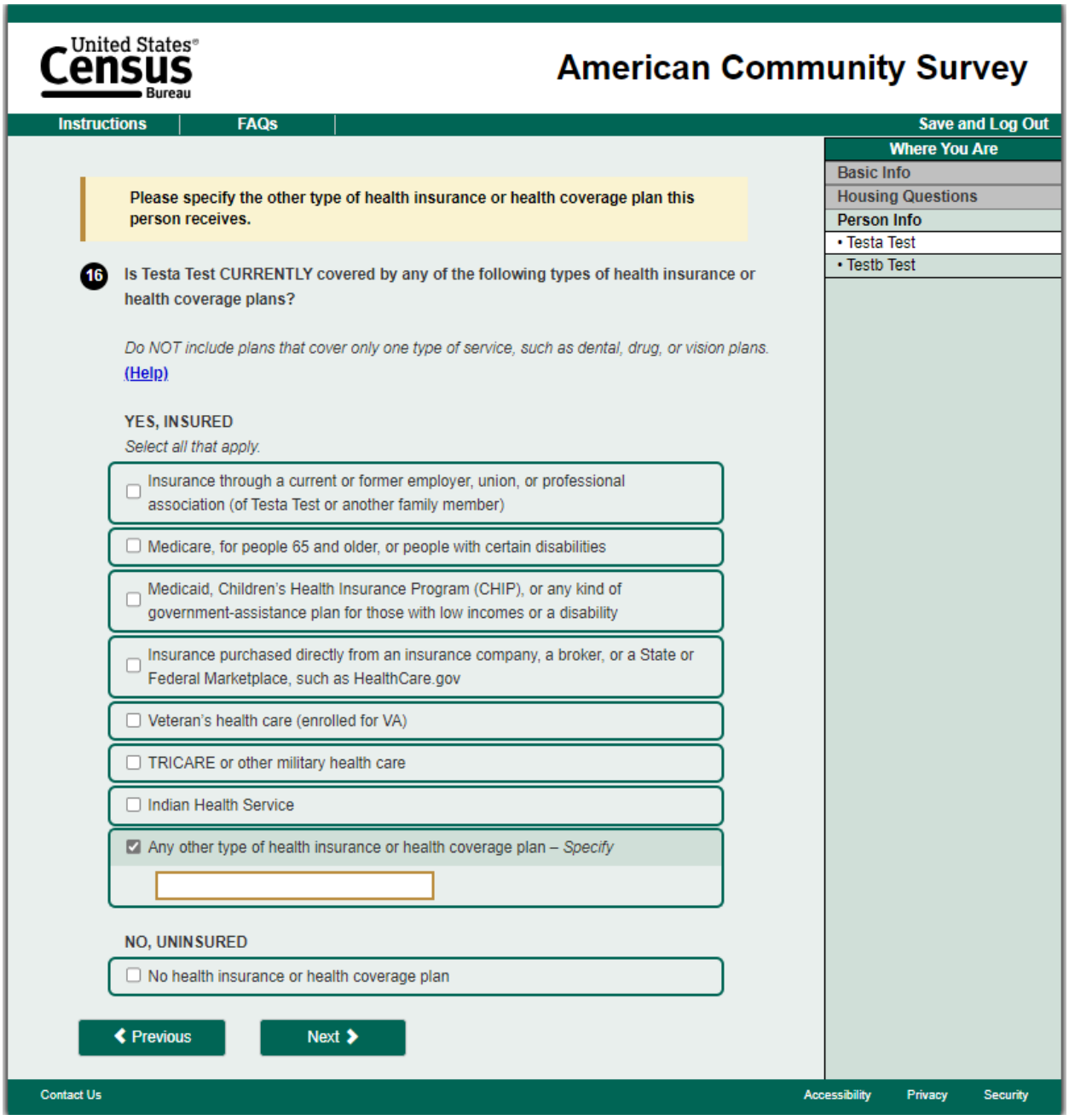
Language (T3)

Two options



Insurance (T3)

Select multiple, textbox, centered radio button, banner, box, arrow removal



Review (T3)

Banner and row

United States Census Bureau

American Community Survey

Instructions | FAQs | Save and Log Out

Review & Edit

Answers to Questions for Testa A Test

Additional information is needed. Please click the link on the highlighted row below to answer an additional question.

To change a response, click on the underlined answer to return to that question.

To complete an unanswered question, click on the underlined [BLANK] to return to that question.

Print for your records

Sex	Male
Date of birth	06/09/2004
Age	20
Hispanic, Latino, or Spanish origin	Other Hispanic/Latino/Spanish
Race	Black or African Am.
Place of birth	Outside the United States dominican
Citizenship	Naturalized U.S. citizen (1999)
Year came to live in U.S.	2005
Attended school	Public school/college
Grade level attended	[BLANK]
Highest level of school completed	Some college-less than 1 year
Ancestry	[BLANK]
Language other than English at home	No
Residence one year ago	This address
Health insurance	Through employer/union
Deaf/difficulty hearing	[BLANK]
Blind/difficulty seeing	[BLANK]
Difficulty concentrating/remembering	[BLANK]
Difficulty walking/climbing stairs	[BLANK]
Difficulty dressing/bathing	[BLANK]
Difficulty doing errands alone	[BLANK]
Marital status	Now married
Got married-last 12 months	[BLANK]
Serve on active duty	Never served