sciendo

Journal of Official Statistics, Vol. 38, No. 3, 2022, pp. 793–822, http://dx.doi.org/10.2478/JOS-2022-0035

Identifying Data Quality Challenges in Online Opt-In Panels Using Cognitive Interviews in English and Spanish

Yazmín García Trejo[1], *Mikelyn Meyers*[1], *Mandi Martinez*[1], *Angela O'Brien*[1], *Patricia Goerman*[1], *and Betsarí Otero Class*[1]

In this article, we evaluate how the analysis of open-ended probes in an online cognitive interview can serve as a metric to identify cases that should be excluded due to disingenuous responses by ineligible respondents. We analyze data collected in 2019 via an online opt-in panel in English and Spanish to pretest a public opinion questionnaire (n = 265 in English and 199 in Spanish). We find that analyzing open-ended probes allowed us to flag cases completed by respondents who demonstrated problematic behaviors (e.g., answering many probes with repetitive textual patterns, by typing random characters, etc.), as well as to identify cases completed by ineligible respondents posing as eligible respondents (i.e., non-Spanish-speakers posing as Spanish-speakers). These findings indicate that data collected for multilingual pretesting research using online opt-in panels likely require additional evaluations of data quality. We find that open-ended probes can help determine which cases should be replaced when conducting pretesting using opt-in panels. We argue that open-ended probes in online cognitive interviews, while more time consuming and expensive to analyze than close-ended questions, serve as a valuable method of verifying response quality and respondent eligibility, particularly for researchers conducting multilingual surveys with online opt-in panels.

Key words: Data quality; cognitive interviews, open-ended probes; online opt-in panels.

1. Introduction

Online opt-in panels are becoming increasingly prevalent in survey research as a method for quick data collection (Chmielewski and Kucker 2019; Edgar et al. 2016; Kennedy et al. 2021). In recent years researchers have even used opt-in panels with convenience samples to pretest early versions of survey instruments in place of or in addition to traditional pretesting methods such as in-person cognitive interviews (Murphy et al. 2014; Lenzner and Neuert 2017; Neuert and Lenzner 2019). Having an easily accessible group of respondents who have signed up to fill out multiple surveys sent to them over time can be a

[1]U.S. Census Bureau, 4600 Silver Hill Road, Suitland-Silver Hill, MD 20746, U.S.A. Emails: yazmin.argen.garcia.trejo@census.gov, mikelyn.v.meyers@census.gov, mandi.martinez@census.gov, angela.c.obrien@census.gov, patricia.l.goerman@census.gov and betsari.otero.class@census.gov

Acknowledgments: The authors would like to thank both U.S. Census Bureau reviewers as well as this journal's reviewers for their time and thoughtful feedback. We also thank our collaborators at U.S. Census Bureau and our sponsors for making this research possible. Finally, we are thankful for our families, who encourage us and keep us motivated.

This article is released to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not those of the U.S. Census Bureau. The U.S. Census Bureau reviewed this data product for unauthorized disclosure of confidential information and approved the disclosure avoidance practices applied to this release [Approved for release under CBDRB-FY20-189 and CBDRB-FY19-546].

© Statistics Sweden

quick and efficient way for researchers to collect data on a whole host of topics (Matthijsse et al. 2015; Porter et al. 2019). However, questions remain about the quality and comparability of data that can be obtained from online opt-in panels in general and across languages in particular. It can be challenging to put together an online opt-in panel of members of hard-to-count populations, such as non-English speaking immigrants in the U.S., who have demographic characteristics that are representative of their population in the U.S. (Barreto et al. 2018; Ortega-Santos 2019). Moreover, researchers who purchase non-English convenience samples from online opt-in panel vendors can encounter challenges evaluating data quality, such as ensuring respondents' proficiency in non-English languages. Researchers need to be able to identify problematic answers and response patterns in order to differentiate between legitimate and non-legitimate responses across languages, which may need to be discarded from analysis.

In this article, we examine data from cognitive interview probes administered through online opt-in panels in English and Spanish to pretest survey questions from the 2020 U.S. Census Attitudes Survey (Team Y&R 2020). We were particularly interested in understanding how the inclusion of open-ended probes in online opt-in panels might help us identify data quality issues that indicate that a panel respondent was providing low quality responses. We refer to this problem as "response quality." An example of an indicator of a response quality problem is respondent behavior like copying and pasting text from the question stem rather than providing their own answer. Notably, this type of behavior would be highly unusual in a face-to-face cognitive interview or survey, and seems to be unique to the online, asynchronous format of this project.

Open-ended probes, such as those used in online cognitive interviews, require respondents to think about and write in their own response to a survey question, rather than choosing from a pre-set list of response options (Behr et al. 2017; Meitinger and Behr 2016). Such probes are common in face-to-face survey pretesting methods like cognitive interviewing, which typically involves the use of probing questions to gauge respondent understanding of survey questions and concepts (Neuert and Lenzner 2019; Willis 2005). Open-ended probes allow respondents to share what they are thinking without imposing the restrictions of categorical questions with predetermined response options (Willis 2005). Probing questions can detect cognitive issues with comprehension, retrieval, and the decision-making process when answering survey questions (Behr et al. 2017; Edgar 2013). For example, a typical probe may ask a participant to explain with their own words what a concept like "confidentiality" means to them.

In recent years, researchers have increasingly experimented with online pretesting using opt-in panels as a supplement to or a replacement for traditional in-person cognitive interviewing (Yu et al. 2019; Lenzner and Neuert 2017; Meitinger and Behr 2016; Edgar et al. 2016; Behr et al. 2012). Although scholars have studied the efficacy of many online quality checks in detecting bogus respondents using methods like evaluating open-ended responses (Buchanan and Scofield 2018; Chandler and Paolacci 2017; C. Kennedy et al. 2020; R. Kennedy et al. 2020; Kennedy et al. 2021), little is known about the role that open-ended probes in English and Spanish can play in data quality checks for online cognitive interviews. We argue that open-ended probes in online pretesting can help to detect potential response quality problems when using opt-in panels to pretest survey questions in English and Spanish. Notably, online cognitive testing has become even more

common in recent times due to safety restrictions on face-to-face interactions during the COVID-19 pandemic (Shepperd et al. 2021; Upadhyay and Lipkovich 2020). Thus, scholarship needs to identify challenges associated with data quality issues using online opt-in panels. This type of analysis could also help vendors to think of new ways to improve non-English online opt-in panel data quality as the demand for this type of panels increases. For this article, we focus on three research questions:

1. How can open-ended probes serve to identify response quality issues in online cognitive interviews using online opt-in panels in English and Spanish?
2. Are response quality issues identified in open-ended probes different (or similar) across English and Spanish responses in online cognitive interviews using a large U.S. based online opt-in panel?
3. If poor quality cases (i.e., disingenuous respondents) are identified, are open-ended probes helpful in determining which responses should be eliminated and/or replaced?

2. Literature Review

Online opt-in panels provide a relatively affordable and efficient way to pretest surveys in comparison to the funding and time needed for in person cognitive interviewing (Callegaro et al. 2014; Edgar et al. 2016). For example, researchers do not need to worry about in person scheduling logistics when using online opt-in panels for survey pretesting (Barreto et al. 2018; Edgar et al. 2016). Additionally, online pretesting provides for greater geographic dispersion and increased diversity among research participants as compared with traditional cognitive testing (Meitinger et al. 2019; Meitinger and Behr 2016), which typically takes place in fewer geographic locations due to time and resource constraints (Chandler et al. 2019; Edgar et al. 2016; Yarrish et al. 2019).

At the same time, there are drawbacks to pretesting surveys online via opt-in panels. Due to the lack of an interviewer in self-administered online pretesting, it is impossible to include spontaneous, unscripted probes that might arise based on topics introduced by respondents (Behr et al. 2012; Behr et al. 2017). Self-response data collected via online opt-in panels lacks interviewer notes, which can often include insights about respondent affect or difficulty exhibited during in-person pretesting (Edgar et al. 2016). Consequently, online pretesting results may be more difficult to interpret and may result in less detailed findings (Behr et al. 2012; Behr et al. 2017; Edgar et al. 2016). An additional drawback is the presence of bots and/or humans purposely misrepresenting themselves, which pose a threat to data quality in online research (Chmielewski and Kucker 2019; Hillygus et al. 2014). Given these challenges, it is important to understand the data quality issues that might arise when using online opt-in panels for cognitive interviews.

While research on deceptive responses to pretesting surveys is still in its infancy, research to date indicates that disingenuous respondents and bots can have a negative impact on data quality and results (Chmielewski and Kucker 2019; Kennedy et al. 2021; Simone 2019a; Yarrish et al. 2019). Kennedy et al. (2021, 3) defines bots in online surveys as "computer algorithms designed to complete online surveys automatically." Additionally, researchers have pointed out that people can misrepresent themselves in order to qualify for studies with financial incentives for participants (Chandler and Paolacci 2017; Chmielewski and Kucker 2019; Kennedy et al. 2021). This type of misrepresentation can be particularly rampant in

online surveys that seek to recruit vulnerable populations (Chandler and Paolacci 2017). For example, Chandler and Paolacci (2017) determined that their study on the LGB community had a fraud rate of 62% when no measures were taken to curtail cheating. In other words, 62% of respondents falsely claimed to qualify for the study based on their sexual orientation when researchers did not try to prevent respondent misrepresentation of their sexuality. Kennedy et al. (2021) has found that bots tend to be particularly effective with multiple-choice, close-ended survey questions, but they can be programmed to complete other types of questions as well (Kennedy et al. 2021; Simone 2019a).

There are several ways to spot misrepresentation (both electronic and human) within a survey (Kennedy et al. 2021; Prince et al. 2012; Yarrish et al. 2019). First, researchers should look to see if the prevalence of a required specific characteristic is realistic within a sample (Chandler and Paolacci, 2017). For example, Ortega-Santos (2019) finds that although M Turk provides a large pool of Spanish-speakers, there are limitations on the diversity of bilingual and non-bilingual respondents. Ortega-Santos (2019) notes that in her M Turk study, most bilingual Spanish respondents resided in the United States, while the majority of monolingual Spanish-speaking respondents came from countries outside of the United States. Another potential problem with online surveys is the presence of bogus responses. Kennedy et al. (2021, 11–12) notes that when asked open-ended questions such as "When you were growing up, what was the big city nearest where you lived?", many participants responded with text appropriate for reviewing a product, but not answering that question, such as "ALL SERVICES SOUND VERY GOOD" or responding with the name of a state or other non sequitur responses. Moreover, some of the responses were word-for-word identical, such as a multitude of respondents who provided the identical response that their favorite thing to do in a new city is to, "Explore by Foot. VirtualTourist members agree that the best way to see a new destination is to experience it by foot" (C. Kennedy et al. 2020, 67).

This finding raised questions about the integrity of the participants and their subsequent responses to other questions in the survey. Another potential indicator of poor-quality responses is quick completion time. Although quick completion can be an indication of a satisficer, or a respondent who chooses the easiest possible response rather than putting much thought into their answer, it may also be an indication of a bot, since bots often complete surveys with inhuman quickness (C. Kennedy et al. 2020; Kennedy et al. 2021).

There are several steps researchers can take to prevent poor quality in online opt-in panel data. Simone (2019b) recommends including at least two or three open-ended questions as a quality check. Although electronic bots are becoming increasingly adept, open-ended questions remain a reliable way of detecting and flagging unusual data. Kennedy et al. (2021) recommends adding homonyms, or words that have identical spellings and/or sounds, but different definitions, to open-ended questions to make problematic answers more detectable. For example, Kennedy et al. (2021) asked respondents, "What would you like to see elected leaders in Washington get done during the next few years? Please give as much detail as you can." Although some of her respondents responded appropriately with political issues, many respondents replied by detailing the life of George Washington or talking about Washington state. Particularly suspicious answers featured reordered text clearly taken from somewhere else on the Internet.

Nearly all research on detecting data quality issues in online opt-in panels has focused on English-speakers. To our knowledge, the only research focusing on detecting data quality issues with Spanish-speakers comes from studies where researchers were trying to

crowdsource translations. Crowdsourcing is a method by which researchers acquire services by asking for contributions from a large number of people, often from an online community (Chandler et al. 2019; Chmielewski and Kucker 2019; Ortega-Santos 2019). Some crowdsourced language studies have detected professional respondents in all languages, including Spanish (Aranberri et al. 2017; Ortega-Santos 2019). Our article contributes to a gap in the literature relating to Spanish-language responses to online opt-in panels as compared to English-speaking respondents. We hope that this article will start a methodological and ethics conversation on data quality issues for English- and Spanish-language online opt-in panels, particularly for those that aim to include vulnerable populations. Advancing this research topic will benefit both researchers and vendors as both need to understand the challenges in data quality that their research efforts may encounter.

3. Data

The data we discuss here were collected via an online opt-in panel in both English and Spanish. These data were collected from a convenience sample as part of a round of online cognitive interviews to pretest questions for the 2020 Census Attitudes Survey, a public opinion survey about people's attitudes and planned participation in the 2020 Census. The survey was also pretested via two rounds of traditional in-person cognitive interviews in English and Spanish prior to the online testing.

The online pretesting occurred in July of 2019 and focused on unresolved issues with the question wording from the previous in-person testing. In the survey introduction we explained that participants needed to first answer a survey questionnaire and then respond to follow-up items about the questions previously asked. Eighteen survey questions and thirteen probes were included in a forced-choice design (see Appendix, Subsection 8.1), and there were 463 completed cases (265 in English and 198 in Spanish). These numbers included unedited cases with signs of problematic behavior; for example, "speeders" (i.e., people who answered the survey below the 25th percentile respondent completion time, which is less than four minutes in both languages) and respondents in the Spanish-language sample who did not self-identify as Hispanic or Latino. While the online survey platform we used typically delivers cleaned data (e.g., with "speeders" deleted), we had a unique opportunity to examine the initial un-scrubbed data because our agency has regulations that only permit individuals with U.S. Census Bureau Special Sworn Status to review unedited responses in order to protect respondents' confidentiality. Thus, the typical vendor data scrub procedure was not used before we received the data. Rather, we received unedited responses and had the opportunity to review them ourselves.

For the pretesting study, respondents first answered a series of survey questions. Next, they answered a series of retrospective probes about a subset of the survey questions (see Appendix, Subsection 8.2). This retrospective probing approach was consistent with the prior in-person cognitive interview approach, with minor adaptations to the probe wording for an online audience. Half of the sample in both languages was randomly assigned to receive one of two alternate versions of two of the survey questions.

The survey platform we used typically includes a language toggle that allows respondents to switch between Spanish and English at any point while completing a survey. We asked the vendor to disable this toggle for our study to better separate our sample of English-speakers

from Spanish-speakers. We wanted to discourage ineligible English-speakers from posing as Spanish-speakers and filling out the survey more than once. In addition, a separate link to access the survey in each language was created to distribute invitations to Spanish- and English-speaking respondents directly in the appropriate language.

When designing respondent samples, researchers typically set goals or quotas for how many participants they would like to include with specific characteristics. In face-to-face cognitive interviews, it is common to seek a small, diverse sample of respondents with demographic characteristics of interest for a given study. In most online opt-in panels, researchers can also set quotas for a specific number of completed cases of any given type. For instance, researchers can request that half the sample be comprised of men and half of women, or that a certain percentage of respondents have an advanced educational degree.

We wanted to include Spanish-speaking respondents in this study who reflected the variation in educational level, sex, birthplace, and region of the general population of Hispanics in the U.S. Monolingual Spanish-speaking respondents in the U.S. are more likely to have completed high school or less than the general population. Therefore, we requested overrepresentation of Spanish-dominant speakers who had completed high school or less (see Table 1). Of note, for this study, it took longer to fill the quotas for Spanish-dominant speakers than it did for English-speakers. This may be due to the challenges that researchers face in recruiting vulnerable or hard-to-reach populations across survey modes.

Table 2 shows the demographic characteristics of the respondents who completed the survey, which in some cases, differed from the quotas that we had requested in Table 1. For example, we ended up with a larger number of English-speakers (265) than we had originally requested (250). The Spanish-speaking sample fell short of the originally requested 250 responses, with only 198 completed cases. Educational attainment was another variable for which quotas were not fulfilled in both languages. Spanish-speakers with less than high school education were difficult to find among the online opt-in panel respondents, and they only comprised about 20% of the respondents. Spanish-speaking women were over-represented compared with men. We did not impose quotas by race and ethnicity, so any respondent in the vendor's contact frame who was identified as a Spanish-speaker was eligible to answer the survey. One unexpected finding, shown in Table 2, was that about 48% of the sample of Spanish-speaking participants did not identify as Hispanic.

Table 1. Sample quotas requested for the study.

Variable	English	Spanish
Sex	Men (120) Women (130)	Men (120) Women (130)
Education	High school education or less (100) More than High School (150)	High school education or less (150) More than High School (100)
Birthplace	U.S. Born (250)	U.S. Born (150) Foreign born (100)
Region	Census regions (125 from the West and 125 from Northeast and South)	Census regions (150 from the West and 100 from Northeast and South)

Source: 2020 U.S. Census Attitudes Survey.

Table 2. Characteristics of the participants in the pretesting project.

		English	Spanish
Sex	Male	49.8%	36.2%
	Female	50.2%	63.8%
Ethnicity	Hispanic	18.1%	52.3%
	Non-Hispanic	81.9%	47.7%
Race	White	56.3%	58.5%
	Black or African American	27.1%	10.6%
	American Indian or Alaska Native	3.2%	3.4%
	Asian	2.8%	6.8%
	Native Hawaiian or Pacific Islander	1.8%	2.4%
	Some other race	8.8%	18.4%
Age	18-30	54.7%	44.2%
	31-45	22.3%	34.2%
	46-55	9.8%	11.6%
	56-65	6.0%	6.5%
	Over 65	7.2%	3.5%
Education	Less than High School	40%	19.6%
	Completed High School	29.8%	17.6%
	Some College, no degree	9.4%	15.6%
	Associate's degree (AA/AS)	6.0%	11.6%
	Bachelor's degree (BA/BS)	9.1%	27.6%
	Post Bachelor's degree	5.7%	8.0%
Total N		265	198

Source 2020 U.S. Census Attitudes Survey.

We requested some information from the vendor such as the number of invitations sent, language recruitment measurements used and fraud control mechanisms the vendor uses for their online opt-in panel participants. Disclosure of this type of information may vary from vendor to vendor for online opt-in panels. There were 326,272 invitations sent for this study, and a total of 8,439 respondents accessed the survey but did not complete it. Given our 463 completed cases, the completion rate for this study was 0.14%, with an additional 2.6% of the sample providing incomplete responses. Since we have not sought to compare data from other vendors, we do not know how common these extremely low response rates are for online opt-in panels, and we are unaware of any scholarly publications that compare this type of information or any other type of methodological information across online opt-in panels. The design of guidelines for participation rates and data quality in online opt-in panels is an area ripe for future research.

At our request, the vendor shared that they used the following question about language spoken at home for the recruitment of Spanish-speaking participants:

What best describes the combination of English and Spanish that you speak within your home?

1. Spanish only
2. Spanish more than English
3. Spanish/English equally,
4. English more than Spanish
5. English only

6. None of the above
7. Prefer not to answer.

We requested that the vendor included only people who responded "Spanish only" or "Spanish more than English" for our study. The vendor further told us that all demographic information collected during the recruitment of respondents for the opt-in panel was self-reported and no additional language or reading proficiency items were used. The vendor also shared that they typically use the following mechanisms to prevent fraud:

- Diverse recruitment partnerships with blue chip companies in retail, travel, entertainment, and business,
- Proprietary members who are known and trusted,
- Consistent source blending that is constantly monitored,
- Double opt-in "member join process" measuring thought consistency, unlikely correlations, low-probability answers, speeding and more,
- Close panelist monitoring for behavioral anomalies (e.g., abrupt location changes or abnormal qualification rates), and
- Phone/SMS validation for reward redemption 3rd party data validation, network of checks for browser, system and BIOS configurations.

The Spanish translation of the survey instrument was produced by certified translators with extensive experience translating survey materials in the U.S. Census Bureau's Decennial Translation Branch. The translations were also expert reviewed to verify that the phrasing of probes was built on findings from prior multilingual pretesting research using a process outlined in [Citation removed].

4. Methods

Through our analysis we sought to provide answers to the following hypotheses:

- *Hypotheses 1*(H1): Open-ended probes can serve to identify response quality issues in multilingual online cognitive interviews using online opt-in panels in English and Spanish,
- *Hypotheses 2* (H2): Open-ended probes in Spanish online opt-in panel pretesting will have more response quality issues compared to respondents answering in the language predominant for online opt-in panels (i.e., English in the U.S.), and
- *Hypotheses 3* (H3): Open-ended probes can help identify and remove disingenuous responses from online cognitive interviews when used in tandem with other data quality measures (e.g., removing "speeders").

To test whether open-ended probes can uncover data quality issues in online cognitive interviews, a team of researchers developed a coding scheme based on recommendations from the 2019 Washington Statistical Society annual address (Kennedy et al. 2021) to document seven types of problematic responses (i.e., responses provided by respondents who were not making a good faith effort to fill out the survey). We did not include valid survey responses in our analysis. Table 3 provides a definition for each of these types of problematic responses. The study included three different types of open-ended probes: (1) probes that asked about the meaning of concepts or words, (2) probes about a specific part

Table 3. Coding categories for valid and problematic responses.

Coding categories	Definition
Valid response.	Valid answers/No issues found
Problematic responses	Responses that do not answer the question adequately
1. Incomplete responses	Respondent uses language related to the topic of the question but provides an incomplete answer (e.g., "About United States")
2. Non sequitur	The answer has nothing to do with the topic of the question (e.g.,"Pink, "yes" or "no" as the response to a question asking for a definition or a preference)
3. Copied survey text	Text copied from the question stem
4. Random typing	Set of letters typed at random (e.g., "asdf;asdf;ljasf")
5. Profanity	Respondent uses inappropriate or offensive language (e.g., "stop asking the same fxxxxxx questions")
6. Volunteers that respondent is not a Spanish-speaker	Respondent writes that they do not know or speak Spanish (e.g., "I don't speak Spanish)
7. Repetitive textual pattern	Uses the same text across three or more responses (e.g., "no idea," "DK," "Yes," etc.)

of the question stem, and (3) probes about participants' question wording preference. In total, we analyzed data from thirteen open-ended probes, each of which respondents had to answer in order to advance to the next screen.

The codes enable us to measure the extent of response quality problems (to address H1), and to compare their prevalence across languages and probe types (to address H2). For our final hypothesis (H3), we conduct an exploratory analysis on flagging cases that might need to be removed due to data quality issues. In particular, we focus on how response quality issues identified in open-ended probes can be used in combination with other data quality measures, such as identifying "speeders," to make decisions about which cases should be removed.

As mentioned previously, when conducting in-person cognitive interviews, researchers rarely encounter these types of data quality problems. Typical data quality issues encountered during cognitive interviewing might include discovering a respondent does not have the expected demographic characteristics or interviewing a respondent who provides minimal responses to probing questions. However, response quality issues comparable to a respondent providing "random answers" to complete the survey more quickly and receive an incentive would not typically occur in face-to-face cognitive interviews. Moreover, for in-person cognitive interviews, which are considered a qualitative method, researchers do not typically need to interview a large number of respondents in order to obtain useful findings. Thus, even if researchers need to discard a relatively large number of the online cases when using an opt-in panel, the results could still be useful for pretesting purposes. This consideration distinguishes online pretesting efforts from both in person cognitive testing and regular production surveys.

Qualitative analysis of the open-ended responses was done using NVivo. Researchers first collaborated on creation of the coding scheme. Next, two bilingual researchers independently coded a portion of the open-ended responses in English and Spanish. After they had coded

four of the probes, we met as a team to come to a consensus on how to code the more complex cases. We conducted an evaluation of intercoder reliability based on a third of the probes coded, which resulted in a 92% intercoder reliability rate for Spanish and 94% for English. Researchers determined that the number of potential codes and the frequency that multiple codes were assigned to the same response made the probability of agreement relatively low. As a result, this intercoder reliability evaluation was completed using percent agreement. As a last data quality step, the coding was reviewed by a senior researcher. Once all coding was completed, we produced frequency tables of the seven indicators of poor-quality. Given that multiple codes could be applied to each response, the denominator used to calculate percentages corresponds to the number of references (or codes) for a given question, rather than the total number of respondents who answered the question.

5. Findings (Hypotheses 1 and Hypotheses 2)

5.1. *H1: Open-ended Probes Can Serve to Identify Response Quality Issues in Multilingual Online Cognitive Interviews Using Online Opt-in Panels in English and Spanish*

After coding the responses to the thirteen open-ended probes, we assessed our first hypothesis by evaluating how responses to the open-ended probes could serve as a metric to identify response quality issues in the pretesting. Table 4 summarizes the number of responses coded as valid (appropriate/helpful) responses and problematic responses for both English and Spanish. Results are presented by type of open-ended probe and by language. The rows add up to a hundred percent in each language because the denominator is the number of references coded for each row, showing the prevalence of the data quality issue.

Table 4. Prevalence of valid responses and problematic responses by probe and by language.

Probe types	Response Types (English)			Response Types (Spanish)		
	Valid response	Problematic responses	Total	Valid response	Problematic responses	Total
Probes about meaning of concepts or words	%	%		%	%	
Meaning of "questions on the census form"	49	51	267	32	68	184
Meaning of "confidential"	71	29	268	30	70	188
Meaning of "The Census Bureau will not keep answers to the 2020 Census confidential"	51	49	267	25	75	183
Meaning of "share individuals answers with other government agencies"	48	52	268	19	81	193
Meaning of the phrase "used against you"	55	45	266	27	73	185
Meaning of the "2020 Census" (Version 2 asked to half of the sample)	62	38	133	31	69	93
Conditional probe about the content of the question						
Concerns respondents had in mind when asked, "How concerned are you, if at all, that the answers you provide to the 2020 Census will be used against you?"	49	51	267	25	75	184
Concerns respondents had in mind when asked, "How concerned are you, if at all, that an unauthorized person from outside the Census Bureau will access your answers to the census?"	52	48	267	29	71	184
Government agencies respondent had in mind	58	42	267	25	75	185
Probes related to preference of the wording of a question						
Comparing two versions of the intention to participate in the census questions. Version A vs. Version B1 (Asked to half of the sample)	69	31	67	30	70	47
Comparing two versions of the intention to participate in the census questions. Version A vs. Version B2 (Asked to half of the sample)	52	48	73	40	60	43
Reasons for preference of Census awareness questions (Form 1)	69	31	77	43	57	46
Reasons for preference of Census awareness questions (Form 2)	74	26	76	31	69	55
Average across all questions	**58**	**42**		**30**	**70**	

Note: "Valid responses" refers to all the responses that did not present indications of response quality issues as measured in this paper. The columns labeled as "Problematic Responses" aggregated all quality issues defined in Table 3. The average across all questions was calculated adding the percentages reported in each row and divided by the number of rows. Source: 2020 U.S. Census Attitudes Survey.

Interestingly, although responses in both languages were coded as problematic, the Spanish-language responses had a higher incidence of problematic responses across all probe types. Table 4 shows that valid English-language responses range from 48 to 74% depending on the type of probe asked, while valid responses occurred far less often in Spanish, ranging from 19 to 43%. For example, when respondents were asked about the meaning of the term "confidential," about 71% of responses were coded as valid in English compared to only 30% of the Spanish-language responses.

5.2. *H2: Open-ended Probes in Spanish Online Opt-in Panel Pretesting Will Have More Response Quality Issues Compared to Respondents Answering in the Predominant Language for Opt-in Panels (i.e., English in the U.S.)*

For our second hypothesis, we focus only on the problematic responses shown in Table 4 above. We expected more response quality issues with Spanish-language responses than English-language responses based on past findings from in-person cognitive interview studies. For example, research using face-to-face pretesting among Spanish-dominant Hispanics indicates that internet access does not necessarily correlate to the type of digital affinity necessary to complete online surveys (e.g., experience typing in URLs, familiarity with establishing accounts with a username and password, etc.) (Brown et al. 2016; Trejo and Schoua-Glusberg 2017). Moreover, research on survey pretesting also shows reliability measurement challenges for the Spanish-speaking population in the U.S. For instance, educational attainment is a well-established item used in many surveys in English. However, when pretested in Spanish, researchers have found that Spanish-speakers can interpret educational attainment response options differentially based on conventions for naming levels of education in their country of origin, which implies that further improvements to the translation of certain questions may be necessary to ensure reliability (Goerman et al. 2018).

Table 5 shows the frequency of six response quality issues for each probe type in English and seven response quality issues for Spanish. The code for respondents who volunteered that they were not Spanish-speakers was only applicable to Spanish-language cases, which results in one additional type of data quality issue possible in Spanish-language cases. One of the most salient themes that emerged was the prevalence of repetitive textual patterns. In our initial coding, we focused on coding one probe at a time, with each probe in its own column. As we worked, we noticed that contiguous responses by the same respondent across columns or different questions sometimes contained identical responses, indicating that a respondent was answering multiple open-ended probes the same way. Once we noticed this potential copying and pasting behavior, we began to code this as a repetitive textual pattern. This problem occurred on average in 42% of English-language responses, but the frequency of repetitive textual patterns on average was higher in Spanish (54%). These repetitive textual patterns have important implications for response quality and may be the result of respondents attempting to complete the survey as quickly as possible, or they may indicate the presence of a bot.

The second common code in both languages and across all types of probes was non sequiturs (26% on average in both English and Spanish), which is similar to what other researchers have discussed as a quality issue in the open-ended question literature (Kennedy et al. 2021). For instance, researchers have hypothesized that respondents may

Table 5. Prevalence of problematic responses and probe and by language.

Probe Types	Type of Problematic Responses																
	Incomplete responses	Incomplete responses	Non sequitur	Non sequitur	Copied survey text	Copied survey text	Random typing	Random typing	Profanity	Profanity	Volunteers that Respondent is not English speaker	Volunteers that Respondent is not Spanish speaker	Repetitive Textual Pattern	Repetitive Textual Pattern	Total	Total
Probes' language (EN= English, SP= Spanish)	EN	SP	EN	SP	EN	SP	EN	SP	EN	SP	EN	SP	EN	SP	EN	SP
Probes about meaning of concepts or words	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	
Meaning of "questions on the census form"	4	5	30	20	10	0	13	6	1	1	0	3	41	66	135	126
Meaning of "confidential"	1	17	17	17	0	0	27	8	4	2	0	4	51	54	78	132
Meaning of "The Census Bureau will not keep answers to the 2020 Census confidential"	19	1	22	30	0	0	17	7	2	2	0	5	39	54	132	138
Meaning of "share individuals' answers with other government agencies"	17	9	26	20	0	10	17	7	3	1	0	4	38	48	139	157
Meaning of the phrase "used against you"	5	10	24	19	4	1	26	6	2	1	0	1	39	61	120	135
Meaning of the "2020 Census" Version 2 (asked to half of the sample)	16	0	16	27	0	9	10	6	0	2	0	8	58	48	50	64
Conditional probe on the content of the question																
Concerns respondents had in mind when asked "How concerned are you, if at all, that the answers you provide to the 2020 Census will be used against you?"	6	1	33	28	1	0	19	7	1	1	0	3	40	59	136	138
Concerns respondents had in mind when asked "How concerned are you, if at all, that an unauthorized person from outside the Census Bureau will access your answers to the census?"	3	5	30	26	1	0	20	9	2	2	0	2	45	57	128	130
Government agencies respondent had in mind	5	2	21	32	0	1	24	7	2	1	0	3	47	53	112	139
Probes related to preference of the wording of a question																
Comparing two versions of the intention to participate in the census questions. Version A vs. Version B1 (Asked to half of the sample)	5	12	33	18	0	0	29	6	5	9	0	0	29	55	21	33
Comparing two versions of the intention to participate in the census questions. Version A vs. Version B2 (Asked to half of the sample)	3	4	37	31	0	0	29	15	0	0	0	4	31	46	35	26
Reasons for preference of Census awareness questions (Form 1)	4	4	17	42	0	0	33	4	4	0	0	0	42	50	24	26
Reasons for preference of Census awareness questions (Form 2)	5	11	30	21	0	0	10	8	5	3	0	3	50	55	20	38
Average across all questions	7	6	26	26	1	2	21	7	2	2	0	3	42	54		

Note: The rows add up to one hundred percent in each language. The total for each language in each row shows the number of problematic references coded for each probe (the sum of all seven quality issues defined in Table 3). The average across all questions was calculated adding the percentages reported in each row and divided by the number of rows. Source: 2020 U.S. Census Attitudes Survey.

be tired of answering too many open-ended probes, or they may type nonsensical responses in order to advance to the next question more quickly and complete the survey sooner (Neuert and Lenzner, 2019).

While many findings were common across languages, some codes were more commonly applied to responses in one language than another. Random typing, where respondents entered gibberish such as "adjkipsdjklfuid," was on average more common in English (21%) than in Spanish (7%) across all probes. One behavior only seen in the Spanish-language data was respondents volunteering that they were not Spanish-speakers (3% on average). While data quality issues like non sequiturs and random typing can be attributed to different causes and may not always be indicative of a deceptive respondent (e.g., respondents may have made a good faith effort to answer most of the survey but may demonstrate fatigue as the survey goes on), this is not the case for respondents who volunteer that they are not Spanish-speakers when filling out a Spanish-language survey. This behavior is much more clearly indicative of a deceptive respondent. In both languages, there were relatively few instances of text copied directly from the survey, with two percent on average or fewer of responses coded as "copied text" in both languages. Use of profanity was also relatively rarely observed (2% on average across languages). When this behavior occurred, it was a red flag in terms of respondents not making a good faith effort to fill out the survey.

To our knowledge, there are no widely accepted standards in the survey industry for ensuring high quality data when including Spanish-dominant Spanish-speakers in online opt-in panels based in the U.S. However, as our experience shows, the onus is often on researchers to request the inclusion of project-specific language screening questions in order to ensure that the correct respondents are included in the final sample. As we found in this study, even when Spanish proficiency is a requirement for the study, the sample of respondents may erroneously include respondents who are not fluent in the language in question.

5.3. *Findings: Further Exploratory Analysis for Hypothesis 3 (H3): Open-Ended Probes Can Help Identify and Remove Disingenuous Responses from Online Cognitive Interviews when used in Tandem with other Data Quality Measures (i.e., Removing "Speeders")*

The occurrence of repetitive textual patterns in open-ended probes, as shown in the previous section in Table 5, allowed us to identify cases that might need to be eliminated from or replaced in the analysis. As an exploratory analysis for this article, we decided to compare these findings on problematic open-ended responses with a data quality metric used to evaluate close-ended responses: respondents who showed signs of being "speeders." Speeding is one of the behaviors commonly examined when evaluating the data quality of survey responses (Kennedy et al. 2021; Storozuk et al. 2020).

We defined "speeders" for the purposes of this exploratory research as participants whose time to complete the main survey and the probing questions fell on or below the 25th percentile (227 seconds for English and 215.5 seconds for Spanish). Respondents whose response time was above the 75th percentile (557 seconds for English and 566 seconds for Spanish) were considered to be respondents who took a substantial amount of time to respond to the online cognitive interview. There are two reasons we selected the 25th percentile to define speeders in our exploratory analysis. First, in analysis not shown in this article, we calculated the number of data quality issues on the 10th fastest percentile based on time of completion of the online cognitive interview. When we did this calculation, we found that there were no differences in the average number of data quality problems when comparing the 10th and 25th percentiles. For example, for the cases in Spanish, the 10th percentile had an average of ten data quality issues and the 25th percentile had 9.8 problems on average. Our second reason for defining speeders using the 25th percentile is that the 25th percentile has more than twice the number of cases, which improves comparability with other quartiles and the sample overall. In particular, the decile size is about 20 cases but the quartile size is about 50 cases. Once we identified "speeders," we calculated the average number of problematic codes for respondents in each quartile. The results for the average number of problems by topic and time to answer are shown in Table 6. This table shows that "speeders" in the survey also accounted for most, but not all, data quality problems identified in the analysis of the open-ended questions in English and Spanish. These results signal that the analysis of open-ended items may identify additional participants displaying problematic behaviors that are missed when only using more traditional methods for verifying data quality like identifying "speeders."

In an effort to better understand the data quality issues found in the Spanish-language sample, we completed an analysis not shown in this article in which we evaluated how different sociodemographic groups varied in their responses to open-ended questions. When analyzing gender and age, we did not find differences by age across response quality issues. However, the results by gender show that women had fewer response quality issues.

6. Limitations

Some limitations should be kept in mind regarding our study results. Online opt-in panel vendors typically use data cleaning procedures to weed out poor quality responses prior to sending data to clients. As mentioned before, due to our agency's data security

Table 6. Average Number of Problems found by quartile of the duration of the stuvey.

Quartile of the duration of the survey	Average number of problems	Type of Problematic Responses: Incomplete responses	Non sequitur	Copied survey text	Random typing	Profanity	Volunteers that respondent is not a Spanish sneaker	Repetitive textual pattern
English								
Q1 (0-25)	8.2	0.2	1.5	0.1	2.6	0.2	0.0	3.7
Q2 (25-50)	3.9	0.5	1.1	0.1	0.4	0.1	0.0	1.9
Q3 (50-75)	2.5	0.3	0.9	0.1	0.4	0.0	0.0	0.8
Q4 (75-100)	2.4	0.4	0.9	0.1	0.1	0.1	0.0	0.9
Spanish								
Q1 (0-25)	9.8	0.1	1.5	0.1	1.3	0.2	0.4	6.2
Q2 (25-50)	8.3	0.6	2.1	0.1	0.3	0.0	0.4	4.7
Q3 (50-75)	4.5	0.4	1.5	0.1	0.2	0.0	0.0	2.2
Q4 (75-100)	3.3	0.4	1.2	0.2	0.0	0.1	0.0	1.2

Note: Q = Quartile.
Source: 2020 U.S.Census Attitudes Survey.

requirements, the vendor's customary data cleaning procedures could not be implemented, and as such, the raw data that we started with was likely of poorer quality than what consumers might typically receive from online opt-in panels. However, we suspect that the ability of many online opt-in panel vendors in the U.S. to provide in depth data cleaning in non-English languages is probably somewhat limited. Some of the problems with the Spanish-language online opt-in panel may have persisted even if the typical data cleaning procedures had been followed prior to our receiving the dataset. For example, Kennedy et al. (2021) found similar data quality concerns to those we identified when analyzing responses from Hispanic respondents in edited datasets. They also found that bots often select the first response option of close-ended questions, which can lead to an overrepresentation of deceptive respondents reporting themselves as Hispanic, since this option appears first in most ethnicity questions (Kennedy et al. 2021). In addition, some language or cultural groups may be less likely to be "looped in" to participating in this type of online opt-in panel than others. This type of finding can help online opt-in panel vendors to refine their recruitment and data cleaning methods, particularly for non-English-speaking, U.S. based respondents.

We encourage online opt-in panel vendors and researchers to revise the guidelines on the recruitment of vulnerable populations (University of Maryland 2014), in particular for non-English-speaking populations, to encourage greater participation by these types of respondents. We recommend that the recruitment of participants who speak languages other than English be a two-part process, with quality control implemented not just by sample vendors but also by researchers who evaluate data quality to ensure that respondents meet eligibility criteria and provide useful responses.

Another limitation of this study is the increased respondent burden when answering open-ended probes as compared to closed-ended items. It seems likely that open-ended probes will impose some burden since they deal with comprehension issues and require in-depth thinking, and this burden may be compounded for participants who are responding to these questions on their smartphones. This leads to a potential dilemma: the burden of answering probing questions may motivate some of the respondent "shortcuts" that we aim to identify in this manuscript. This dilemma is an important limitation to consider for future analyses and may be more relevant in pretesting efforts that include many open-ended probes than in production surveys that contain fewer open-ended questions.

A further limitation of our study is the lack of quotas imposed for race and Hispanic ethnicity. We focused our analysis on participants who reported being Spanish-dominant Spanish-speakers, but we did not screen out respondents who selected that they did not self-identify as Hispanic. Interestingly, we found that 48% of the respondents who self-identified as Spanish-dominant speakers did not identify as Hispanics. This discrepancy may be an artifact of participants interpreting the recruitment question about language spoken at home in different ways, or it may be related to the issue of participants who posed as Spanish-speakers to receive the incentive but reported not speaking Spanish in open-ended questions. In their article about best practices for collecting data on racially diverse online panels, Barreto et al. (2018) stress the importance of including quotas relevant to race and ethnic groups that help researchers better represent the groups under study. These recommendations should be considered in future designs of quotas for online cognitive interviews using opt-in panels.

A final limitation we would like to acknowledge is that the survey that we administered was programmed with a "forced choice" design, in which respondents were required to enter a response to the open-ended questions before they could continue with the survey. This practice is often implemented to determine participant eligibility for subsequent questions, or to ensure that researchers have enough data to analyze for each question. However, a drawback to forced choice question designs is that some participants may be forced to answer a question they do not know the answer to, or to choose between inadequate response choices in order to complete the survey and get their incentive. Therefore, some respondents may have been forced to answer a question they would have left blank under other circumstances.

7. Summary and Discussion

This article provides evidence that open-ended probes used in online cognitive interviews can root out response quality issues in the emerging domain of online opt-in panels. While respondents taking part in live cognitive interviews may provide low quality responses by engaging in behaviors like giving vague or brief responses, in our experience, most respondents would be reluctant to completely flout conversational norms in a live interview (e.g., by stringing together nonsense syllables, or only repeating the question verbatim instead of providing a response to the probe). Thus, a unique artifact of conducting online cognitive interviews is that the lack of a live interviewer may make respondents feel more comfortable engaging in behaviors that result in low quality responses such as typing gibberish or copying and pasting the question text into the response field.

While this research highlights the benefits of using open-ended probes as a data quality validation for cognitive interviews conducted with online opt-in panels in English and Spanish, our findings also signal that analysis of open-ended items is an important data quality check for multilingual surveys beyond pretesting efforts. We argue that research about data quality and best practices for multilingual online opt-in panels should be carried out in earnest as online opt-in panels in non-English languages become more common. Online opt-in panels are an innovative medium that have great promise for survey research in general and for the inclusion of diverse multilingual and multicultural populations as well. Moreover, due to the COVID-19 pandemic, remote data collection methods are becoming increasingly common as researchers around the world adapt face-to-face data collection to other modes appropriate to the current context.

An interesting issue raised by this analysis is the quality of non-English opt-in panels in countries such as the United States. The analysis of Spanish-language open-ended probe responses uncovered response quality issues, including the use of repetitive textual patterns and participants in the Spanish-language survey who explicitly said that they were not Spanish-speakers. We also found that open-ended probes can help determine which cases should be replaced when conducting a survey using an online opt-in panel. These findings aligned with the exploratory analysis we conducted on respondents who were speeding, since speeders often provided low-quality open-ended responses.

Based on our findings, there are several areas that we identify here for future research. It would be valuable to compare responses to open-ended probes across other online opt-in panels to evaluate the type and frequency of problematic behaviors when vendors use

different approaches to recruitment. Further research could expand on our exploratory analysis of respondents who provide low quality data for open-ended items, and how they correspond (or not) to respondents identified as providing problematic responses using traditional measures for detecting disingenuous responses (e.g., speeders, straightlining, etc.). As researchers adapt to the COVID-19 pandemic, conducting online cognitive interviews may become increasingly common. While cognitive interviews are traditionally a qualitative method, the usage of online opt-in panels could result in larger sample sizes with potentially low response rates such as those that we encountered in this study. As researchers tailor cognitive interviews to an online format with larger samples, the study of non-response bias in online pretesting studies may become a necessary topic of discussion.

Additionally, further research to compare data from different vendors or in other languages beyond English and Spanish is needed. For example, experiments comparing online opt-in panels could help to evaluate the variables needed to recruit monolingual, bilingual and multilingual participants in Spanish and other languages, as well as the incidence of poor data quality. Another potential area of research could investigate whether some participants would feel more comfortable recording audio of their responses in non-English languages rather than typing their responses, especially if they are fluent speakers of the language but rarely use it to respond in writing to surveys.

While some of our findings regarding data quality for U.S. Spanish-language cases are troubling, we believe that efforts to improve the representation of non-English-speakers in survey research using emerging methods like online opt-in panels are important given both the increasing linguistic diversity of the U.S. as well as declining survey response rates (De Heer and De Leeuw 2002; Hanson and Espinosa 2016; Luiten et al. 2020; Rumbaut and Massey 2013). We conclude that the inclusion of open-ended probes and survey questions, while more time-consuming and expensive to analyze, is an important method of verifying data quality, particularly in multilingual research. Much research is still to be conducted on this timely topic.

8. Appendix

8.1. Wording of questions and probes

Survey question	Retrospective probe	Sample
Every ten years, the United States Census Bureau conducts a nationwide census, or a count of all individuals living in the United States. The next few questions are about the 2020 Census, which will occur on April 1st, 2020. How much have you seen or heard recently – within the last week or so – about the 2020 Census?	The first question you were asked was, "Every ten years, the United States Census Bureau conducts a nationwide census, or a count of all individuals living in the United States. The next few questions are about the 2020 Census, which will occur on April 1st, 2020. How much have you seen or heard recently – within the last week or so – about the 2020 Census?" This question mentions the "2020 Census." Based on your understanding, what is the "2020 Census"?	All

Continued

Survey question	Retrospective probe	Sample
Now we're going to show you an alternative version of the question. Version 1 (question you answered): "The next few questions are about the 2020 Census, which will occur on April 1st, 2020. How much have you seen or heard recently – within the last week or so – about the 2020 Census?" Version 2 (Alternative): "Every ten years, the United States Census Bureau conducts a nationwide census, or a count of all individuals living in the United States. The next few questions are about the 2020 Census, which will occur on April 1st, 2020. How much haveyou seen or heard recently – within the last week or so – about the 2020 Census?	Can you say more about why you prefer that version?	Half
Now we're going to show you an alternative version of the question. Version 1 (question you answered): "Every ten years, the United States Census Bureau conducts a nationwide census, or a count of all individuals living in the United States. The next few questions are about the 2020 Census, which will occur on April 1st, 2020. How much have you seen or heard recently – within the last week or so – about the 2020 Census?" Version 2 (Alternative): "The next few questions are about the 2020 Census, which will occur on April 1st, 2020. How much have you seen or heard recently – within the last week or so – about the 2020 Census?"	Can you say more about why you prefer that version?	All
How likely are you to participate in the 2020 Census? By participate we mean answer the questions on the census form. • Definitely will (1) • Probably will (2) • Might or might not (3) • Probably will not (4) • Definitely will not (5)	Another survey question was, "How likely are you to participate in the 2020 Census? By participate we mean answer the questions on the census form." What comes to mind when you see the phrase "questions on the census form" in this question?	All
Now we're going to show you an alternative version of the question. Version 1 (question you answered):	Can you say more about why you prefer that version?	All

Continued

Survey question	Retrospective probe	Sample
"How likely are you to participate in the 2020 Census? By participate we mean answer the questions on the census form." Version 2 (alternative): "How likely do you think it is that you or someone in your household will answer the questions on the census form?"		
Now we're going to show you an alternative version of the question. Version 1 (question you answered): "How likely are you to participate in the 2020 Census? By participate we mean answer the questions on the census form." Version 2 (alternative): "In 2020, the Census Bureau will mail nearly every household in the U.S. a questionnaire and an invitation to participate online. How likely do you think it is that you or someone in your household will mail back the questionnaire or submit it online?"	Can you say more about why you prefer that version?	Half
How concerned are you, if at all, that the Census Bureau will not keep answers to the 2020 Census confidential? • Extremely concerned (1) • Very concerned (2) • Somewhat concerned (3) • Not too concerned (4) • Not at all concerned (5)	What does the word "confidential" mean to you in this question?	Half
How concerned are you, if at all, that the Census Bureau will not keep answers to the 2020 Census confidential? • Extremely concerned (1) • Very concerned (2) • Somewhat concerned (3) • Not too concerned (4) • Not at all concerned (5)	In your own words, what does the phrase "the Census Bureau will not keep answers to the 2020 Census confidential" mean to you in this question?	All
How concerned are you, if at all, that the Census Bureau will share individuals' answers to the 2020 Census with other government agencies? • Extremely concerned (1) • Very concerned (2) • Somewhat concerned (3)	What government agencies were you thinking of when you were answering this question?	All

Continued

Survey question	Retrospective probe	Sample
• Not too concerned (4) • Not at all concerned (5)		
How concerned are you, if at all, that the answers you provide to the 2020 Census will be used against you? • Extremely concerned (1) • Very concerned (2) • Somewhat concerned (3) • Not too concerned (4) • Not at all concerned (5)	Another survey question was, "How concerned are you, if at all, that the answers you provide to the 2020 Census will be used against you?" What concerns came to mind when you answered the question that appears above, if any	All
How concerned are you, if at all, that the answers you provide to the 2020 Census will be used against you? • Extremely concerned (1) • Very concerned (2) • Somewhat concerned (3) • Not too concerned (4) • Not at all concerned (5)	What does the phrase "used against you" mean to you in this question?	All
How concerned are you, if at all, that an unauthorized person from outside the Census Bureau will access your answers to the census? • Extremely concerned (1) • Very concerned (2) • Somewhat concerned (3) • Not too concerned (4) • Not at all concerned (5)	The next survey question was, "How concerned are you, if at all, that an unauthorized person from outside the Census Bureau will access your answers to the census?" What concerned came to mind when you answered the question that appears above, if any?	All

8.2. 2020 Census Attitudes Survey (Instrument Used for Pretesting Purposes)

Thank you for agreeing to participate in this research study. On the next screen, you will be asked to begin filling out a questionnaire. After you have filled out some of the questions, you will be asked to answer some follow-up questions about the previous question items. Your responses will help us make sure that everyone understand the questions the same way. The survey will take 10 minutes or less to complete. By law, the Census Bureau is required to keep your answers confidential. This survey is voluntary and you may refuse to answer any question. The results of this study will be used to inform future surveys and censuses. This collection has been approved by the Office of Management and Budget (OMB). This eight-digit OMB approval number, 0607-0725, confirms this approval. If this number was not displayed, we could not conduct this survey. We estimate that this survey will take about 10 minutes. By proceeding with this study, you give your consent to participate in this study. Your privacy is protected by the Privacy Act. There are a limited number of uses of your data that are permitted under the Privacy Act. You can find a list of these uses in the System of Records Notice (SORN). For this survey, the SORN is named "COMMERCE/CENSUS-5, Decennial Census Program." For more information, please

visit our website at census.gov and click on “Data Protection and Privacy Policy” at the bottom of the home page.

What is your age?

_____years

[PROGRAMMING NOTE: Terminate under 18 yrs]

Are you ....

1. Male
2. Female

What state do you live in?

1. Alabama
2. Alaska
3. Arizona
4. Arkansas
5. California
6. Colorado
7. Connecticut
8. Delaware
9. Florida
10. Georgia
11. Hawaii
12. Idaho
13. Illinois
14. Indiana
15. Iowa
16. Kansas
17. Kentucky
18. Louisiana
19. Maine
20. Maryland
21. Massachusetts
22. Michigan
23. Minnesota
24. Mississippi
25. Missouri
26. Montana
27. Nebraska
28. Nevada
29. New Hampshire
30. New Jersey
31. New Mexico
32. New York
33. North Carolina
34. North Dakota

35. Ohio
36. Oklahoma
37. Oregon
38. Pennsylvania
39. Rhode Island
40. South Carolina
41. South Dakota
42. Tennessee
43. Texas
44. Utah
45. Vermont
46. Virginia
47. Washington
48. West Virginia
49. Wisconsin
50. Wyoming
51. I do not reside in the United States

Were you born in the United States?
1. Yes
2. No

Are you of Hispanic, Latino, or Spanish origin?
1. Yes, Hispanic, Latino, Spanish origin
2. No, not of Hispanic, Latino, or Spanish origin

What is your race? Please select all that apply.
1. White
2. Black or African American
3. American Indian or Alaska Native
4. Asian
5. Native Hawaiian or Pacific Islander
6. Some other race

What is the highest degree or level of school you have completed?
1. No high school
2. Some high school
3. High school graduate or equivalent (for example GED)
4. Some college, but degree not received or is in progress
5. Associate degree (for example AA, AS)
6. Bachelor's degree (for example BA, BS, AB)
7. Graduate degree (for example master's, professional, doctorate)

Lang_Eng. Do you speak a language other than English at home?
1. Yes
2. No

ASK IF Lang_Eng = 1:

What is this language?

ASK IF Lang_Eng = 1:
How well do you speak English?

1. Very well
2. Well
3. Not well
4. Not at all

The next few questions are about the 2020 Census, which will occur on April 1st, 2020. How much have you seen or heard recently – within the last week or so – about the 2020 Census? Would you say a great deal, some, a little, or nothing at all?

1. A great deal
2. Some
3. A little
4. Nothing at all

Every ten years, the United States Census Bureau conducts a nationwide census, or a count of all individuals living in the United States. The next few questions are about the 2020 Census, which will occur on April 1st, 2020. How much have you seen or heard recently – within the last week or so – about the 2020 Census?

1. A great deal
2. Some
3. A little
4. Nothing at all

How likely are you to participate in the 2020 Census? By participate we mean answer the questions on the census form.

1. Definitely will
2. Probably will
3. Might or might not
4. Probably will not
5. Definitely will not

How concerned are you, if at all, that the Census Bureau will not keep answers to the 2020 Census confidential?

1. Extremely concerned
2. Very concerned
3. Somewhat concerned
4. Not too concerned
5. Not at all concerned

How concerned are you, if at all, that the Census Bureau will share individuals' answers to the 2020 Census with other government agencies?

1. Extremely concerned
2. Very concerned
3. Somewhat concerned

4. Not too concerned
5. Not at all concerned

How concerned are you, if at all, that the answers you provide to the 2020 Census will be used against you?

1. Extremely concerned
2. Very concerned
3. Somewhat concerned
4. Not too concerned
5. Not at all concerned

How concerned are you, if at all, that an unauthorized person from outside the Census Bureau will access your answers to the census?

1. Extremely concerned
2. Very concerned
3. Somewhat concerned
4. Not too concerned
5. Not at all concerned

/* RANDOM ROTATE SERIES */

Which one of the following is the most important concern, if any concern, about filling out the census form??Mark ONE box.

1. The Census Bureau will not keep answers to the 2020 Census confidential
2. The Census Bureau will share individuals' answers to the 2020 Census with other government agencies
3. The answers you provide to the 2020 Census will be used against you
4. An unauthorized person from outside the Census Bureau will access your answers to the census
5. No concerns
6. Other (Please specify)

_____

Now, we would like to ask you a few questions about some of the questions you just answered.

The first question you were asked was, "The next few questions are about the 2020 Census, which will occur on April 1st, 2020. How much have you seen or heard recently – within the last week or so – about the 2020 Census?"

_____

Now we're going to show you an alternative version of the question. Version 1 (question you answered): "The next few questions are about the 2020 Census, which will occur on April 1st, 2020. How much have you seen or heard recently – within the last week or so – about the 2020 Census?" Version 2 (Alternative): "Every ten years, the United States Census Bureau conducts a nationwide census, or a count of all individuals living in the United States. The next few questions are about the 2020 Census, which will occur on April 1st, 2020. How much have you seen or heard recently – within the last week or so – about the 2020 Census?

Are these questions asking the same thing or something different?
1. Same thing
2. Something different

Q67. Which version of the question do you prefer?
1. Version 1 (question you answered)
2. Version 2 (alternative)
3. No preference

ASK IF Q67 = 1 OR 2

Can you say more about why you prefer that version?

The first question you were asked was, "Every ten years, the United States Census Bureau conducts a nationwide census, or a count of all individuals living in the United States. The next few questions are about the 2020 Census, which will occur on April 1st, 2020. How much have you seen or heard recently – within the last week or so – about the 2020 Census?" This question mentions the "2020 Census." Based on your understanding, what is the "2020 Census"?

Now we're going to show you an alternative version of the question. Version 1 (question you answered): "Every ten years, the United States Census Bureau conducts a nationwide census, or a count of all individuals living in the United States. The next few questions are about the 2020 Census, which will occur on April 1st, 2020. How much have you seen or heard recently – within the last week or so – about the 2020 Census?" Version 2 (Alternative): "The next few questions are about the 2020 Census, which will occur on April 1st, 2020. How much have you seen or heard recently – within the last week or so – about the 2020 Census?"

Are these questions asking the same thing or something different?
1. Same thing
2. Something different

Q48. Which version of the question do you prefer?
1. Version 1 (question you answered)
2. Version 2 (alternative)
3. No preference

ASK IF Q48 = 1 OR 2

Can you say more about why you prefer that version?

Another survey question was, "How likely are you to participate in the 2020 Census? By participate we mean answer the questions on the census form." What comes to mind when you see the phrase "questions on the census form" in this question?

Now we're going to show you an alternative version of the question.

Version 1 (question you answered):
"How likely are you to participate in the 2020 Census? By participate we mean answer the questions on the census form."

Version 2 (alternative):
"How likely do you think it is that you or someone in your household will answer the questions on the census form?"

Are these questions asking the same thing or something different?
1. Same
2. Something different

Q60. Which version of the question do you prefer?
1. Version 1 (question you answered)
2. Version 2 (alternative)
3. No preference

ASK IF Q60 = 1 OR 2

Can you say more about why you prefer that version?

Now we're going to show you an alternative version of the question.

Version 1 (question you answered):
"How likely are you to participate in the 2020 Census? By participate we mean answer the questions on the census form."

Version 2 (alternative):
"In 2020, the Census Bureau will mail nearly every household in the U.S. a questionnaire and an invitation to participate online. How likely do you think it is that you or someone in your household will mail back the questionnaire or submit it online?"

Are these questions asking the same thing or something different?
1. Same
2. Something different

Q86. Which version of the question do you prefer?
1. Version 1 (question you answered)
2. Version 2 (alternative)
3. No preference

ASK IF Q86 = 1 OR 2

Can you say more about why you prefer that version?

Another survey question was, "How concerned are you, if at all, that the Census Bureau will not keep answers to the 2020 Census confidential?"

What does the word "confidential" mean to you in this question?

In your own words, what does the phrase "the Census Bureau will not keep answers to the 2020 Census confidential" mean to you in this question?

_____

Another question was, "How concerned are you, if at all, that the Census Bureau will share individuals' answers to the 2020 Census with other government agencies?"

In your own words, what does the phrase "share individuals' answers with other government agencies" mean to you in the question above?

_____

What government agencies were you thinking of when you were answering this question?

_____

Another survey question was, "How concerned are you, if at all, that the answers you provide to the 2020 Census will be used against you?"

What concerns came to mind when you answered the question that appears above, if any

_____

What does the phrase "used against you" mean to you in this question?

_____

The next survey question was, "How concerned are you, if at all, that an unauthorized person from outside the Census Bureau will access your answers to the census?"

What concerned came to mind when you answered the question that appears above, if any?

_____

Thank you for your participation! If you have any additional thoughts on this survey, please provide them in the space below.

_____

8. References

Aranberri, N., G. Labaka, A. Díaz de Ilarraza, and K. Sarasola. 2017. "Ebaluatoia: Crowd Evaluation for English–Basque Machine Translation." *Language Resources and Evaluation* 51(4): 1053–1084. DOI: https://doi.org/10.1007/s10579-016-9335-x.

Barreto, M., L. Frasure-Yokley, E. Vargas, and J. Wong. 2018. "Best Practices in Collecting Online Data with Asian, Black, Latino, and White Respondents: Evidence from the 2016 Collaborative Multiracial Post-Election Survey." *Politics, Groups, and Identities* 6(1): 171–180. DOI: https://doi.org/10.1080/21565503.2017.1419433.

Behr, D., L. Kaczmirek, W. Bandilla, and M. Braun. 2012. "Asking Probing Questions in Web Surveys: Which Factors Have an Impact on the Quality of Responses?" *Social Science Computer Review* 30(4): 487–498. DOI: https://doi.org/10.1177/0894439311435305.

Behr, D., K. Meitinger, M. Braun, and L. Kaczmirek. 2017. "Web Probing-Implementing Probing Techniques from Cognitive Interviewing in Web Surveys with the Goal to Assess the Validity of Survey Question." *Social Science Open Access Repository*: 18. DOI: https://doi.org/10.15465/gesis-sg_en_023

Brown, A., G. López, and M. Hugo Lopez. 2016. *Digital Divide Narrows for Latinos as More Spanish Speakers and Immigrants Go Online*. Washington, DC: Pew Research Center. Available at: https://www.pewresearch.org/hispanic/2016/07/20/digital-divide-narrows-for-latinos-as-more-spanish-speakers-and-immigrants-go-online/ (accessed July 2021).

Buchanan, E. and J. Scofield. 2018. "Methods to Detect Low Quality Data and Its Implication for Psychological Research." *Behavior Research Methods* 50(6): 2586–2596. DOI: https://doi.org/10.3758/s13428-018-1035-6.

Callegaro, M., A. Villar, D.S. Yeager, and J.A. Krosnick. 2014. "A Critical Review of Studies Investigating the Quality of Data Obtained with Online Panels Based on Probability and Nonprobability Samples." In *Online Panel Research: A Data Quality Perspective*, edited by M. Callegaro, R.P. Baker, J. Bethlehem, A.S. Göritz, J.A. Krosnick, and P.J. Lavrakas: 23–53. Sussex, UK: Wiley.

Chandler, J. and G. Paolacci. 2017. "Lie for a Dime: When Most Prescreening Responses are Honest but Most Study Participants are Impostors." *Social Psychological and Personality Science* 8(5): 500–508. DOI: https://doi.org/10.1177/194855061769 8203.

Chandler, J., C. Rosenzweig, A. Moss, J. Robinson, and L. Litman. 2019. "Online Panels in Social Science Research: Expanding Sampling Methods Beyond Mechanical Turk." *Behavior Research Methods* 51(5): 2022–2038. DOI: https://doi.org/10.3758/s13428-019-01273-7.

Chmielewski, M., and S. Kucker. 2019. "An MTurk Crisis? Shifts in Data Quality and The Impact on Study Results." *Social Psychological and Personality Science* 11(4): 464–473. DOI: https://doi.org/10.1177/1948550619875149.

De Heer, W. and E. de Leeuw. 2002. "Trends in Household Survey Nonresponse: A Longitudinal and International Comparison." In *Survey Nonresponse*, edited by R.M. Groves, D.A. Dillman, J.L. Eltinge, and R.J.A. Little: 41–54. New York: Wiley.

Edgar, J. 2013. "Self-Administered Cognitive Interviewing." In 68th American Association for Public Opinion Research Conference, May 2013. Boston, MA. Available at: http://www.aapor.org/AAPOR_Main/media/AnnualMeetingProceedings/2013/Session_A-1-3-Edgar.pdf (accessed July 2021).

Edgar, J., J. Murphy, and M. Keating. 2016. "Comparing Traditional and Crowdsourcing Methods for Pretesting Survey Questions." *Sage Open* 6(4): 1–14. DOI: https://doi.org/10.1177/2158244016671770.

Goerman, P., L. Fernandez, and R. Quiroz. 2018. "Translation of Country-Specific Programs and Survey Error: Measuring the Education Level of Immigrants." *Translation & Interpreting* 10(2): 21–33. DOI: https://doi.org/10.1177/2158244016671770.

Hanson, M. and L. Espinosa. 2016. "Culture, Ethnicity, and Linguistic Diversity: Implications for Early Childhood Special Education." In *Handbook of Early Childhood Special Education*, edited by B. Reichow, Boyd, E. Barton, and S. Odom: 455–471. Springer, Cham.

Hillygus, S., N. Jackson, and M. Young. 2014. "Professional respondents in nonprobability online panels." In *Online Panel Research: A Data Quality Perspective*, edited by M. Callegaro, R.P. Baker, J. Bethlehem, A.S. Göritz, J.A. Krosnick, and P.J. Lavrakas: 219–237. Sussex, UK: Wiley.

Kennedy, C., N. Hatley, A. Lau, A. Mercer, S. Keeter, J. Ferno, and D. Asare-Marfo. 2020. *Assessing the Risks to Online Polls from Bogus Respondents*. Washington, DC: Pew Research Center. Available at: https://www.pewresearch.org/methods/wp-content/uploads/sites/10/2020/02/PM_02.18.20_dataquality_FULL.REPORT.pdf (accessed July 2021).

Kennedy, C., N. Hatley, A. Lau, A. Mercer, S. Keeter, J. Ferno, and D. Asare-Marfo. 2021, "Strategies for Detecting Insincere Respondents in Online Polling." *Public Opinion Quarterly* 85(4): 1050–1075. DOI: https://doi.org/10.1093/poq/nfab057.

Kennedy, R., S. Clifford, T. Burleigh, P. Waggoner, R. Jewell, and N. Winter. 2020. "The Shape of and Solutions to the MTurk Quality Crisis." *Political Science Research and Methods* 8(4): 614–629. DOI: https://doi.org/10.1093/poq/nfab057

Lenzner, T. and C. Neuert. 2017. "Pretesting Survey Questions Via Web Probing–Does it Produce Similar Results to Face-to-Face Cognitive Interviewing?" *Survey Practice* 10(4): 2768. DOI: https://doi.org/10.29115/SP-2017-0020.

Luiten, A., J. Hox, and E. de Leeuw. 2020. "Survey Nonresponse Trends and Fieldwork Effort in the 21st Century: Results of an International Study Across Countries and Surveys." *Journal of Official Statistics* 36(3): 469–487. DOI: https://doi.org/10.2478/jos-2020-0025.

Matthijsse, S., E. de Leeuw, and J. Hox. 2015. "Internet Panels, Professional Respondents, and Data Quality." *Methodology* 11(3). DOI: https://doi.org/10.1027/1614-2241/a000094.

Meitinger, K. and D. Behr. 2016. "Comparing Cognitive Interviewing and Online Probing: Do They Find Similar Results?" *Field Methods* 28(4): 363–380. DOI: https://doi.org/10.1177/1525822X15625866.

Meitinger, K., D. Behr, and M. Braun. 2019. "Using Apples and Oranges to Judge Quality?: Selection of Appropriate Cross-National Indicators of Response Quality in Open-Ended Questions." Social Science Computer Review 39(3): 1–22. DOI: https://doi.org/10.1177/0894439319859848.

Murphy, J., Edgar, J., and Keating, M. 2014. "Crowdsourcing in the Cognitive Interviewing Process." In the Annual Meeting of the American Association for Public Opinion Research, May, 2014. Anaheim, CA. Available at: http://dc-aapor.org/2014%20conference%20slides/EdgarMurphyKeating.pdf (accessed January 2021).

Neuert, C., and T. Lenzner. 2019. "Effects of the Number of Open-Ended Probing Questions on Response Quality in Cognitive Online Pretests." *Social Science Computer Review* 39(3): 456–468. DOI: https://doi.org/10.1177/0894439319866397.

Ortega-Santos, Iván. 2019. "Crowdsourcing for Hispanic Linguistics: Amazon's Mechanical Turk as a source of Spanish data." Borealis – An International Journal of Hispanic Linguistics 8(1): 187–215. DOI: https://doi.org/10.7557/1.8.1.4670.

Porter, C.O., R. Outlaw, J.P. Gale, and T.S. Cho. 2019. "The Use of Online Panel Data in Management Research: A Review and Recommendations". *Journal of Management* 45(1): 319–344. DOI: https://doi.org/10.1177/0149206318811569.

Prince, K.R., A.R. Litovsky, and D.G. Friedman-Wheeler. 2012. "Internet-Mediated Research: Beware of Bots." *The Behavior Therapist*, 35(5): 85–88. Available at: https://services.abct.org/i4a/doclibrary/index.cfm?category_id = 15

Rumbaut, R.G., and D.S. Massey. 2013. "Immigration & Language Diversity in the United States." *Daedalus* 142(3): 141–154. DOI: https://doi.org/10.1162/DAED_a_ 00224.

Shepperd, J.A., G. Pogge, J.M. Hunleth, S. Ruiz, and E.A. Waters. 2021. "Guidelines for Conducting Virtual Cognitive Interviews During a Pandemic." *Journal of Medical Internet Research* 23(3): 25173e. DOI: https://doi.org/10.2196/25173.

Simone, M. 2019a. "How to Battle the Bots Wrecking Your Online Study," *Behavioral Scientist*. Available at: https://behavioralscientist.org/how-to-battle-the-bots-wrecking-your-online-study/. (accessed January 2021).

Simone, M. 2019b. "Bots started sabotaging my online research. I fought back," STAT. Available at: https://www.statnews.com/2019/11/21/bots-started-sabotaging-my-online-research-i-fought-back/. (accessed January 2021).

Storozuk, A., M. Ashley, V. Delage, and E. Maloney. 2020. "Got Bots? Practical Recommendations to Protect Online Survey Data from Bot Attacks." *Quantitative Methods for* Psychology 16(5): 472–481. DOI: https://doi.org/10.20982/tqmp. 16.5.p472.

Team Y&R. 2020. *Census Tracking Survey*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research. Available at: https://doi.org/10.3886/E125761V3.

Trejo, Y.A.G., and A. Schoua-Glusberg. 2017. "Device and Internet Use among Spanish-Dominant Hispanics: Implications for Web Survey Design and Testing." *Survey Practice* 10(3). DOI: https://doi.org/10.29115/SP-2017-0016.

University of Maryland. 2014. "Institutional Review Board Investigator Handbook." College Park, MD: University of Maryland. Available at: https://research.umd.edu/-sites/default/files/documents/irb-forms/Investigator%20Handbook%20FINAL%20112114.pdf (accessed July 2021).

Upadhyay, U.D., and H. Lipkovich. 2020. "Using Online Technologies to Improve Diversity and Inclusion in Cognitive Interviews with Young People." *BMC Medical Research Methodology* 20(1): 1–10. DOI: https://doi.org/10.1186/s12874-020-01024-9.

Willis, G.B. 2005. *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Thousand Oaks, CA: Sage Publications. DOI: https://dx.doi.org/10.4135/9781412983655.

Yarrish, C., L. Groshon, J. Mitchell, A. Appelbaum, S. Klock, T. Winternitz, and D. Friedman-Wheeler. 2019. "Finding the Signal in the Noise: Minimizing Responses from Bots and Inattentive Humans in Online Research." *The Behavior Therapist* 42(7): 235–242. Available at: https://services.abct.org/i4a/doclibrary/index.cfm?category_id = 22

Yu, E., A. Fobia, J. Graber, J. Holzberg, R. Kaplan, B. Kopp, K. Kubzdela, B. Mockovak, R. Morrison, P. Scanlon. 2019. *Experiences Using Online Testing to Support Survey-Methods Research and Pre-Testing in the Federal Government.* Research and Methodology Directorate, Center for Behavioral Science Methods Research Report Series (Survey Methodology 2019-06). Washington, DC: U.S. Census Bureau. Available at: http://www.census.gov/content/dam/Census/library/working-papers/2019/adrm/rsm2019-06.pdf (accessed January 2021).

Received January 2021
Revised July 2021
Accepted March 2022