Reducing Item Non-Response to Vote Choice Questions: Evidence from a Survey Experiment in Mexico

Mollie J. Cohen*
University of Georgia

Kaitlen J. Cassell
YouGov America Inc.

Forthcoming at *Public Opinion Quarterly.*

September 18, 2022.

Abstract: Retrospective vote choice is a critical question asked in political science surveys. Yet, this question suffers from persistently high item non-response rates, which can bias estimates and limit scholars' ability to make sound inferences. In this paper, we develop a sensitive survey technique to decrease non-response to the vote choice question in a representative, face-to-face survey in Mexico City and Mexico State in 2018-2019. Respondents received different iterations of three treatments: an anonymity guarantee, a confidentiality reminder, and audio-assisted interviewing technology. The use of audio technology combined with a credible anonymity guarantee significantly improved item response. Both anonymity and confidentiality assurances improved the accuracy of response, which more closely resembled official results in the sensitive conditions. We then evaluate two non-rival mechanisms that might drive our findings: beliefs about response anonymity and re-engagement with the survey. We find that increased perceptions of response anonymity are associated with improved item response.

Acknowledgments: The authors thank the LAPOP Lab at Vanderbilt University for making the data available. Noam Lupu, Kristin Michelitch, Mariana Rodríguez, Keith Weghorst, Liz Zechmeister, and LAPOP lab participants provided invaluable feedback on the instrument. The authors also thank Oscar Castorena and Mariana Rodríguez for their programming assistance, Isaac Cisneros for lending his voice to the project, and Georgina Pizzolitto, Sebastián Larrea, Carole J. Wilson, and the survey team at Data OPM for support in the implementation of this study. The political behavior group at UGA, Shane Singh, and Amy Erica Smith gave helpful feedback on early drafts. The authors also thank the editor and three anonymous reviewers for their thoughtful comments.

* Corresponding author. Mollie J. Cohen, University of Georgia, Department of International Affairs, 202 Herty Drive, Athens, GA 30602, USA; email: mj.cohen@uga.edu.

Political scientists and pollsters rely on survey data to study a wide range of topics, from the abstract (e.g., support for democracy) to the sensitive (e.g., participation in vote-buying exchanges). Yet, polls regularly fail to accurately reflect outcomes of interest, affecting our ability to draw meaningful conclusions from them. Recent polling failures in elections in the United States (Kennedy et al. 2018; Panagopoulos et al. 2018), the UK (Sturgis et al. 2018), and elsewhere have resulted in widespread public discussion about the usefulness and trustworthiness of polls (Castillo-Manzano et al. 2018; Madson and Hillygus 2020), prompting survey researchers to ask: how can we improve the accuracy of survey responses, particularly to sensitive questions?

Scholars have developed tools to minimize known sources of bias that might influence response to sensitive questions (Lynn 2017; Chen and Haziza 2019). Yet, in even the best-designed survey, some respondents may refuse to answer sensitive questions (Berinsky 2008; Peytchev 2013) or provide socially desirable answers rather than honest ones (Krosnick 1999; Anduiza and Galais 2017). In the likely case that the refusal to answer sensitive questions is non-randomly distributed across political preferences, high item non-response rates limit our ability to address core political questions accurately and undermine the credibility of surveys in the eyes of the public.

In this paper, we present the results of a survey experiment conducted in Mexico in 2018-19 that aimed to improve response to a sensitive survey item at the center of political science research: retrospective vote choice. We embedded a pre-registered survey experiment in a face-to-face, representative study of Mexico City and Mexico State to evaluate the effectiveness of using audio-assisted interviewing strategies, in combination with confidentiality and anonymity assurances, to minimize item non-response to the vote choice question.[1] Interviews were

[1] The pre-analysis plan is available at https://osf.io/7rpt3. We have reordered and reworded hypotheses for clarity. Replication data and documentation are available at https://doi.org/10.7910/DVN/OVG2MD.

conducted using Computer-Assisted Personal Interviewing (CAPI) technology, which is widely used in face-to-face surveys (Montalvo et al. 2018; Bush and Prather 2019) and has been successfully used to self-administer questions on sensitive topics in developing contexts (Nanes and Haim 2020). Our study differs from past efforts by combining audio recordings – which render the content of questions and response options silent to bystanders – with CAPI technology. Guaranteeing response anonymity through this simple intervention improved both the rate of response to the vote choice item and the accuracy of responses, resulting in significantly less over-reporting of votes for the winner.

We then assess the evidence for two non-rival mechanisms that could drive improvements in item response: improved perceptions of response anonymity and re-engagement with the survey due to the novelty of the audio technology. We find that the treatment boosts item response in part by increasing respondents' beliefs that their answer is anonymous. While we find only limited evidence that novel question content re-engages participants in the survey, those who enjoyed the audio item were substantially more likely to respond to the vote choice question.

Our findings have important implications for a wide variety of sensitive questions in public opinion. From a practical standpoint, our results demonstrate that in face-to-face surveys using CAPI, including credible anonymity guarantees can significantly improve not only response rates to sensitive questions but also the accuracy of those responses at relatively low cost to survey researchers (see also Nanes and Haim 2020). The findings we present here are relevant for scholars of public opinion and survey researchers more broadly, particularly in the developing world, where most survey research is conducted via face-to-face interviews (Lupu and Michelitch 2018). While we apply this method to vote choice, we expect that this technique could be used to ask a broad range of sensitive items across varied political and social contexts.

Improving Item Response in Surveys

A vital challenge for survey researchers is ensuring that respondents provide truthful answers to survey questions rather than falsifying their preferences or not answering items at all. In the best-case scenario, if item non-response occurs at random, it can inflate standard errors, decreasing confidence in statistical results and making researchers less likely to reject null hypotheses when they should. In the worst-case scenario, when the decision to refuse to answer a survey question is non-random, high item non-response can bias study estimates and artificially deflate standard errors.

In the Americas, retrospective vote choice questions regularly have high levels of item non-response. In the nationally representative 2019 AmericasBarometer survey, for example, 14.2% of Mexicans who reported participating in the 2018 presidential election either could not recall or refused to reveal their vote choice. This item non-response rate is higher than refusals to answer a household income question (12.1%) and political ideology (9.2%) in the same study — two questions that regularly yield higher-than-average item non-response in cross-national public opinion surveys (Riphahn and Serfling 2005; Yan et al. 2010).

While attributes of survey participants (e.g., education) and question design issues (e.g., ambiguous question wording) can exacerbate item non-response, respondents also refuse to answer "sensitive" questions. There is substantial debate over what makes an item sensitive. Recently, Blair, Coppock, and Moor (2020) argue that four conditions must hold for an item to be sensitive: (1) there must be a relevant social referent (e.g., the interviewer, other bystanders), who (2) can learn the respondent's answer. The respondent must also believe that (3) the social referent would prefer she give a particular response, and (4) not providing the referent's preferred response could

entail costs. Under these conditions, respondents who hold socially undesirable views may be less likely to answer questions (Berinsky 1999) or might respond dishonestly (Tourangeau et al. 2000; Tourangeau and Yan 2007).

In the case of vote choice questions, each of these conditions is met. Relevant social referents could include the interviewer, respondents' family and friends, or politicians. Survey interviewers discover a respondent's behavior directly, by hearing their response; where no anonymity protections are in place, bystanders can learn this information indirectly, by overhearing the interview. The respondent must then believe that these social referents hold preferences over her reported behavior, and that providing an undesirable answer will entail some cost. In a democracy like Mexico, where citizens are generally able to freely vote for their preferred candidate, honestly reporting the "wrong" vote choice could result in self-presentation costs (i.e., embarrassment) or, in places where clientelism is pervasive, a lack of future vote buying offers.

Our study aims to improve response to the vote choice question by rendering the second condition irrelevant. If interviewers and observers do not know how a respondent has answered the vote choice question, then social desirability concerns should no longer apply, resulting in higher (and more honest) reporting. We reduce social referents' ability to learn participants' responses using three tools: a confidentiality reminder, a credible guarantee of anonymity, and a novel audio-assisted interview protocol. These interventions provide increasingly stringent protections. Whereas confidentiality signifies that interviewers will not share respondents' answers with others, anonymity indicates that participants' responses will be unknown by the interviewer and other bystanders.

Survey researchers have developed several techniques that credibly guarantee the confidentiality or anonymity of responses, thereby encouraging survey participants to answer

sensitive questions (Singer 1993; Cho and LaRose 1999; Singer et al. 2003; Chauchard 2013; Corbacho et al. 2016; Nanes and Haim 2020). Credible anonymity guarantees include changing the survey mode so that respondents self-administer paper questionnaires in part or in full (for the vote choice question, respondents are commonly asked to mark a paper ballot; e.g., Bishop and Fisher 1995),[2] online surveys, or audio computer-assisted self-interviewing (ACASI) modes, which we employ here. In studies that use ACASI modes, questions are asked via audio recording and individuals self-administer questionnaires using a computer. Because others are unable to learn respondents' answers (there is no interviewer present, and bystanders are unable to hear the questions being asked), interviews conducted using ACASI should be free of presentation concerns. This, in turn, should result in higher, and more honest, reporting. Indeed, experimental studies find that ACASI modes yield lower item non-response and higher reports of sensitive behaviors (Tourangeau and Smith 1996; Dykema et al. 2012; Robertson et al. 2018), which is interpreted as more honest reporting (Gnambs and Kaspar 2015). We expect that this finding will extend to vote choice. That is,

> **H_1:** The probability of non-response to the vote choice question will be lower in conditions that employ audio technology compared to the control condition.

In theory, ACASI modes guarantee response anonymity, removing presentation concerns and thereby improving response. However, in the context of a face-to-face interview, respondents might not realize the strength of these anonymity protections without additional information. We therefore assess whether asking the vote choice question using ACASI modes is more effective

[2] In one poll conducted by a reputable survey firm prior to Mexico's 2018 election, 14.3% of respondents did not answer a vote choice question asked using the secret ballot technique (authors' personal communication).

when combined with reminders that participants' responses are confidential or credible guarantees of response anonymity.

Guaranteeing respondent confidentiality at the beginning of a survey can improve study participation rates and response accuracy (Singer 1993), and is generally viewed as an effective tool for minimizing survey and item non-response (Singer et al. 1995). However, sensitive questions are rarely administered immediately following the consent process because building rapport with subjects over the course of an interview facilitates honest reporting (Krosnick and Presser 2010). Thus, there is a risk that participants will forget promises of confidentiality, which in turn might make them less likely to respond to sensitive items.

To overcome this concern, some recommend reminding participants that their responses will be confidential prior to administering sensitive items. However, evidence about the effectiveness of confidentiality reminders delivered prior to sensitive items is limited, and results from past experimental studies are mixed (Leon et al. 2021). Confidentiality reminders or more stringent guarantees of anonymity can improve some respondents' beliefs that their answers will be protected (Brough et al. 2022). This should make them more likely to respond to sensitive items, and to provide honest responses (e.g., Joinson et al. 2008; García-Yi and Grote 2012). Indeed, if respondents understand and believe reassurances of either confidentiality or anonymity (Whelan 2007; Brough et al. 2022), we should observe larger declines in item non-response as the relative strength of the protection increases (e.g., Ong and Weiss 2000; Ermakova et al. 2016). This discussion yields the following expectations:

> **H_2:** The probability of non-response to the vote choice question will be lower in conditions that employ either confidentiality reminders or an anonymity guarantee compared to the control condition.

H_3: Among the treatments administered via audio, the probability of non-response to the vote choice question will be *highest* in a condition using audio with no reminders and *lowest* in a condition with a combined confidentiality reminder and anonymity guarantee.

We do note that some experiments have shown that confidentiality reminders can backfire. Mid-interview reminders can lead respondents to view this information as especially important, thereby *increasing* confidentiality concerns and resulting in *lower* item response after such reminders (e.g., Frey 1986; Singer et al. 1992; Moore and Ames 2002).

Following from this scholarship, we also expect that credibly guaranteeing response anonymity will improve response *accuracy*. Because the audio treatments remove social referents' ability to identify a participant's response (and therefore to sanction a respondent for the "wrong" answer), individuals in the audio conditions should have no incentive to misrepresent their vote choice. As a result, responses should more closely match election outcomes in treatment conditions. This logic underlies a common assumption in studies of sensitive behaviors, that higher reported levels of sensitive behaviors indicate more truthful responses (Ong and Weiss 2000; Dykema et al. 2012; Robertson et al. 2018; but see Lelkes et al. 2012). However, the true prevalence of sensitive behaviors is often not verifiable. In this case, however, we are able to compare vote reports across treatment conditions to observed outcomes in respondents' place of residence. In brief, if the treatments work to minimize concerns about response confidentiality and anonymity, we expect that the *accuracy* of reporting will improve with these assurances.

H_4: Responses will more closely match official election outcomes in the most anonymous conditions (i.e., the audio-administered condition that contains both a confidentiality reminder and anonymity guarantee).[3]

[3] SM Section F examines expectations of greater improvements to response accuracy where AMLO won by larger margins.

How the Treatment Reduces Item Non-Response

We designed our audio treatments to increase response to the vote choice question by removing the ability for interviewers and bystanders ("social referents" in the language of Blair, Coppock, and Moor 2020) to learn participants' answers. That is, the audio treatment was designed to increase the perceived anonymity of responses. We therefore expect that participants will perceive improvements in the confidentiality or anonymity of their responses after receiving reminders, and that this perception will in turn increase item response. Indeed, experimental studies show that levels of anonymity in survey instruments affect study participants' *perceptions* of their anonymity (Bates and Cox 2008) and perceptions that responses are anonymous increase the provision of sensitive responses (Lavender and Anderson 2009). In short, we expect that treatments using audio technology will result in higher perceptions of response anonymity.

> **H_{1M}**: Beliefs about response anonymity for the vote question will be significantly higher in conditions employing audio technology, and highest in the most anonymous condition.

We evaluate the anonymity perceptions mechanism against a non-rival alternative hypothesis, that improved item non-response results from engagement with the audio technology, rather than perceived improvements to privacy. As survey length increases, respondents tend to pay less attention, which can increase item non-response (Backor et al. 2007; Krosnick and Presser 2010). The vote choice item was located past the midpoint of a 45-minute survey, when participants may have been fatigued and thus predisposed to provide less careful responses. Given the interview context, it is possible that changing the interview mode to one that past scholarship

shows respondents enjoy (e.g., NIMH 2008) could independently improve item response, with the audio question temporarily re-engaging respondents in the study.

We evaluate this possibility by analyzing a novel item that mirrored the structure of the vote choice question: respondents were asked to identify their favorite player on the Mexican national soccer team, which competed in the World Cup during the presidential campaign. This question was designed to be engaging compared to the rest of the survey; if the novelty of the item content could decrease item non-response, we wanted to capture that effect. We also expect that, if respondents find audio items engaging, they will say as much. If re-engagement due to the novelty of the survey *mode* is driving the effect, rather than the novelty of the item *content,* we expect the following hypothesis will hold:

> **H_{2M}:** On average, respondents will report that the audio-administered question (either vote choice or soccer) was more engaging compared to the survey as a whole.

However, if the novelty of the *item* causes re-engagement, we expect to see higher engagement with the survey among those who receive the soccer question, not the vote question, by audio.

Case Selection

We conducted this study in Mexico, a democracy where election outcomes are consistently and correctly reported at low levels of aggregation and where likely costs to respondents for answering a vote choice question include social embarrassment or, in some cases, the removal of clientelistic goods in future elections. Mexico has been recognized internationally as an electoral democracy since 2000 (Andersson and Lindberg 2016). While insecurity due to crime and violence varies sub-nationally and national parties regularly engage in vote buying, Mexican citizens as a whole are

free to vote for their preferred candidate or party without fear of retribution. The 2018 presidential election was highly polarized and we believed that respondents might therefore withhold their vote choice from interviewers. Although the final election outcome was not surprising—Andres Manuel López Obrador ("AMLO") led the polls for months before winning 53.2% of the national vote, more than double the vote share of the second-place candidate—tensions ran higher than the results suggest.

For many, the 2018 election represented a referendum on the formerly hegemonic Institutional Revolutionary Party (PRI) and the National Action Party (PAN)—representatives of the political establishment that had held the presidency since the democratic transition (Crow 2010)—versus populist former Mexico City mayor and thrice presidential candidate AMLO. AMLO fueled anti-establishment rhetoric (Schettino 2017), while establishment candidates portrayed AMLO as a danger to Mexican democracy (Anderson 2018; Osorno 2018). Resulting high levels of polarization could have affected Mexicans' willingness to discuss politics in public (Gerber et al. 2012; Wells et al. 2018), potentially making them less willing to reveal their vote choice to interviewers. Depending on respondents' social circle and the local political context, we expected that both pro-establishment (PAN, PRI) and anti-establishment (AMLO) voters might self-censor.

Study Design

To assess the effectiveness of our treatment on response to the vote choice question, we partnered with the LAPOP Lab at Vanderbilt University to conduct a face-to-face survey of a representative sample of citizens in Mexico City and Mexico State (N=1,892). Fieldwork was conducted soon after the July 1, 2018 presidential election, from November 7, 2018 to February 21, 2019, using

CAPI technology. The unit response rate (RR1) for the survey was 7%.[4] LAPOP conducted extensive real-time quality control during data collection to ensure data quality (e.g., Cohen and Larrea 2018; Montalvo et al. 2018; Cohen and Warner 2021).

ACASI techniques are easiest to implement where individuals have access to technology and can conduct interviews independently and in a private location. These conditions do not hold in much of the developing world. Where respondents may not have access to personal computers, and neither privacy nor literacy is guaranteed among the target population (as in Mexico), the implementation of ACASI modes can be challenging. Our study pairs audio recordings with CAPI technology to approximate an ACASI environment within the context of a face-to-face study. Respondents who reported voting in the most recent election were randomly assigned by the survey software to one of six treatment conditions, described in Table 1. All 1,432 respondents who reported turning out received an identical vote choice question that read: "Who did you vote for in the last presidential election of 2018?"

Table 1. Distribution of Condition Attributes, Respondents, and Item Non-response

Condition	Interviewer	Audio
No reminder	N=247 Non-response: **8.1%**	N=261 Non-response: **5.8%**
Confidentiality reminder	N=253 Non-response**: 8.3%**	N=232 Non-response: **4.7%**
Anonymity guarantee		N=216 Non-response: **2.7%**
Confidentiality + anonymity		N=223 Non-response: **4.9%**

In the first two conditions (Control and Confidentiality Reminder), this item was read by the interviewer, who then provided numbered response options to study participants. In the

[4] RR1 is the most conservative of AAPOR's standard definitions of unit response. For additional detail about how LAPOP calculates unit response rates and tracks interview attempts, see Warner and Camargo-Toledo (2019).

remaining conditions, an identical question was administered via audio recording, which respondents listened to using headsets. Respondents in all conditions were instructed to answer using the *number*, not the candidate's name. To refuse to respond, participants volunteered that they "did not know" or "preferred not to say." This format differs from standard vote choice questions, which ask respondents to select a candidate from a list of alternatives or pose an open-ended question. By numbering the options, we made the response task identical for participants in the control and treatment conditions. We can thus be certain that differences in item response across conditions are due to the audio treatment, not to changes in the response task itself.

The audio question was read by a male native to Mexico City (see SM Section G). Interviewers placed clean, disposable headphone covers on headsets, and tested and adjusted the volume settings before administering the vote choice question (SM Table A1 provides relevant scripts). In all conditions, the order of candidates' names was randomized; in the audio conditions, no names were displayed on interviewers' screens, only radio buttons with associated numbers. As a result, neither interviewers nor bystanders could identify a respondent's vote choice in the audio conditions. "Other" was anchored as the final option; as two "other" candidates ran, this was not an identifying response. Adherence to the protocol was high: according to documentation from the quality control team, only 22 (1.5%) audio interviews were incorrectly applied. We do not drop these interviews.

Participants in the audio conditions received either a confidentiality reminder, an anonymity guarantee, neither of these, or both. The **confidentiality reminder** read: "Before we continue with the next question, I want to remind you that all of the responses you provide me will be kept confidential and anonymous." The **anonymity guarantee** read: "Before we continue with the next question, I want to inform you that I will not know what response you give to this question,

since, as you can see, my screen only displays the numbers of the options, and the order of the options changes with each interview." The combined treatment included both statements.

The confidentiality reminder includes language about anonymity, which directly mirrors informed consent language from the beginning of the study. This bundled treatment increases the strength of the confidentiality reminder and also assures that any differences across the reminder and guarantee conditions (H_3) are attributable to the *credible guarantee*, not the *reminder,* of anonymity.

Results

Table 1 presents the distribution of question attributes, the number of respondents, and item non-response rates (including both "don't know" responses and refusals) among respondents in each condition. Balance checks show minor differences in education levels, age, and state of residence across treatment conditions (SM Table B1). We present models estimated without controls in the paper body; however, our results are stronger when we control for imbalances (SM Table C1). All models are estimated without survey weights or adjusted standard errors. For results with standard errors adjusted to account for the complex sample design, see SM Table C1.

Turnout is over-reported in the sample, by 5.7 and 7.6 percentage points in Mexico State and City, respectively. Of the 932 respondents who were assigned to an audio condition, 77 (8.3%) refused the treatment and were randomly re-assigned to an interviewer-administered condition. Respondents who refused the audio treatment were slightly older and less likely to respond to the vote choice question compared to successfully treated individuals. Because we analyze non-response by participants' initial assignment ("intent-to-treat"), these differences bias our results away from statistical significance. A small number of interviewers had an unusually high number

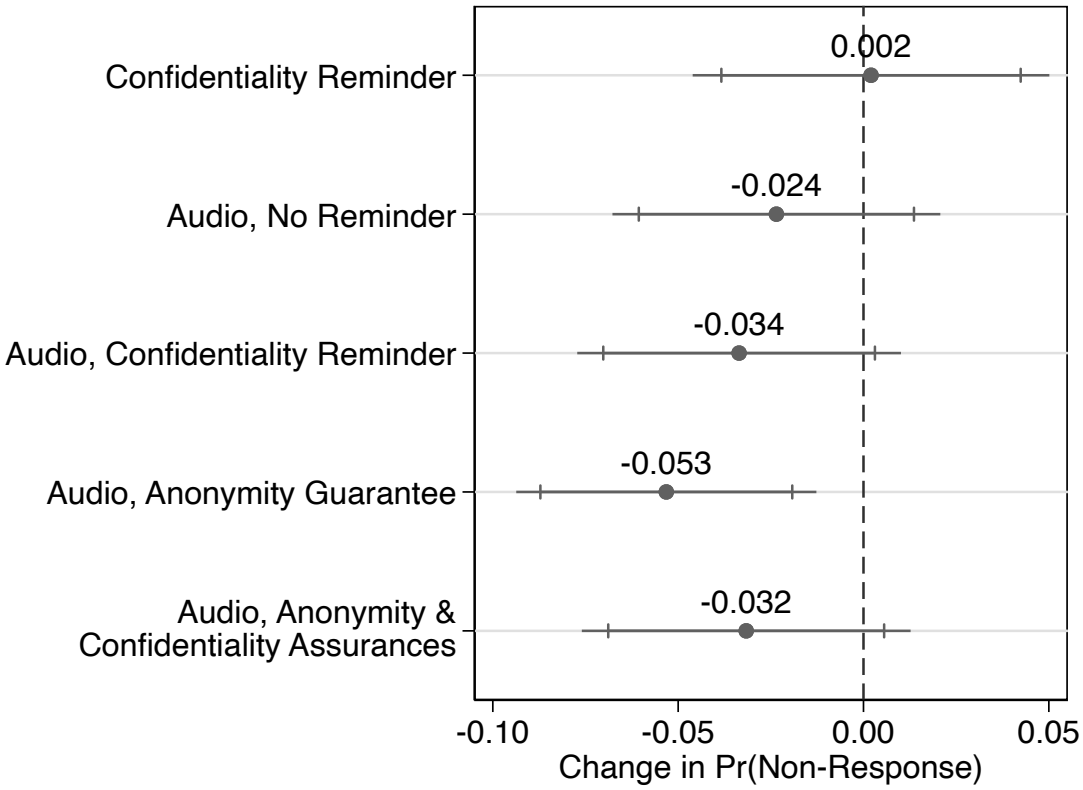
of respondents who refused the audio treatment. This could indicate that those interviewers shirked rather than administer the audio question. However, because interviewers' work is assigned by neighborhood, this pattern could also reflect differences in the population across work areas (SM Section D). To examine this possibility, we estimated design effects (DEFF) for coefficients in models clustering results by interviewer and by block. Design effects range from 0.64 to 1.15 and from 0.99 to 1.05 for interviewer and block, respectively (see SM Table C5), indicating that adjustments could *deflate* standard errors compared to simple random sampling. We present the most conservative results, estimated without adjusted standard errors, in the paper body.

In the control group, item non-response to the vote choice question is 8.1%. This value is substantially lower than we expected; for example, in the 2018 AmericasBarometer, which was conducted only a few months prior to our study, 14.2% of voters did not report their candidate choice. We suspect that this lower level of item non-response is due to differences in question format: we provided response options rather than ask an open-ended question, which likely facilitated candidate recall. We note this lower-than-expected rate of item non-response in the control condition because of its impact on the statistical power of our study. To reliably detect a four-percentage-point decrease in item non-response with 95% confidence (two-tailed) and power of 0.8, we would need 561 respondents in each group, rather than the approximately 220 respondents we recruited based on our expectations (an 8-10 percentage-point decrease), which we derived from past studies. We therefore present both 90% and 95% confidence intervals in figures below.

Figure 1 plots the predicted change in item non-response to the vote choice question for each condition compared to the control group (the interviewer-administered question with no confidentiality or anonymity assurance), estimated using a logistic regression analysis without

controls. For complete results and alternative specifications, see SM Table C1. Item non-response is significantly lower in only one of the four audio conditions (anonymity guarantee with no reminder) compared to the control group, with $p < 0.05$, two-tailed.

Figure 1. Average Item Non-Response by Treatment Condition



Note: Figure presents results from a logit model regressing item non-response on condition, estimated without controls, weights, or adjusted standard errors. Dots represent the change in predicted item non-response for each condition, compared to the control. Whiskers represent a 95% confidence interval, with vertical hash marks at 90% confidence, around estimates. Estimates with whiskers that cross the dashed vertical line are not different from the control group with $p < 0.05$, two-tailed.

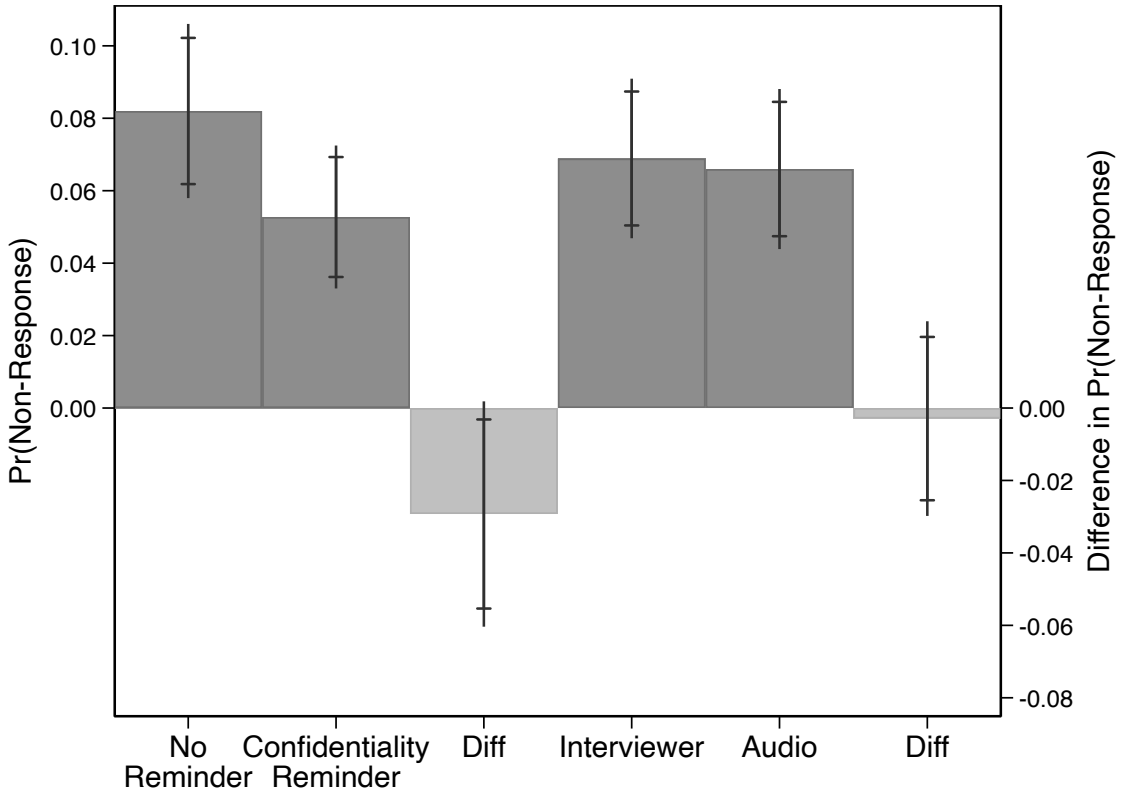
Compared to the control group, item non-response increases slightly (by 0.2 percentage-points, to 8.3%) among those who received a confidentiality reminder prior to the interviewer-administered vote choice question; this difference is small and is not statistically significant (Z = 0.08, p = 0.934, two-tailed). Predicted item non-response declines substantially in the audio conditions: by 2.4 percentage points (Z = -1.04, p = 0.298, two-tailed) among those who received the vote choice question via audio with no confidentiality reminder, 3.4 percentage points (Z = -

1.51, $p = 0.132$, two-tailed) among those who received the randomized audio and a confidentiality reminder, 5.3 percentage points ($Z = -2.58$, $p = 0.010$, two-tailed) among those who received the randomized audio with an anonymity guarantee, and 3.2 percentage points ($Z = -1.40$, $p = 0.162$, two-tailed) among those who received the anonymity guarantee and the confidentiality reminder with the audio item. The decrease in item non-response is only significant at standard thresholds ($p < 0.05$) for the audio treatment with an anonymity guarantee but no reminder; this strong result holds across all alternative model specifications (see SM Table C1). Using a more generous threshold ($p < 0.1$) of statistical significance, item non-response following the audio treatment with anonymity and confidentiality assurances declines significantly compared to control in three of six specifications. Neither the audio condition with no reminder, nor the audio condition with an anonymity guarantee, reach even this generous level of statistical significance. We note these results in consideration of issues of statistical power detailed above.

Figure 2 examines the effect of the audio question and confidentiality reminders on item non-response in the four fully balanced factorial conditions (with and without audio, and with and without confidentiality reminders, excluding conditions with anonymity guarantees; see SM Table C2). The first panel predicts item non-response to the vote choice question in interviewer- versus audio-administered conditions (H_1). As expected, the audio treatments substantially decrease item non-response, although this effect does not reach standard significance thresholds in this model specification. The effect is substantial: item non-response declines by 35%, from 8.2% to 5.3% when pooling across conditions ($t = -1.79$, $p = 0.074$, two-tailed). The second panel shows the effect of confidentiality reminders on item non-response (H_2). Contrary to our expectations, the confidentiality reminder does not improve item non-response in these conditions. Item non-

response declines very slightly, from 6.9% to 6.6% (t = -0.18, p = 0.855) across conditions with and without a confidentiality reminder.

Figure 2. Levels and Change in Item Non-Response by Question Attributes



Note: Figure presents the predicted probability of item non-response across conditions, estimated using logit models predicting item non-response by treatment type. Models include only fully balanced conditions and are estimated without controls, weights, or adjusted standard errors. Dark gray bars represent predicted item non-response by treatment type, and light gray bars represent the difference between these probabilities. Whiskers represent 95% confidence interval, with horizontal hash marks at 90% confidence, around estimates (two-tailed). If whiskers cross the 0-line, an estimate is not distinguishable from 0 (for light gray bars, this means non-response is not distinguishable across groups).

H_3 anticipated that item non-response would be significantly lower in the condition including both an anonymity guarantee and a confidentiality reminder compared to treatments including these factors in isolation. Results presented in Figure 1 are not consistent with this expectation: differences in item non-response between these and the "most-confidential" condition (2.1 and 0.2 percentage points, compared to the audio-administered item with anonymity and

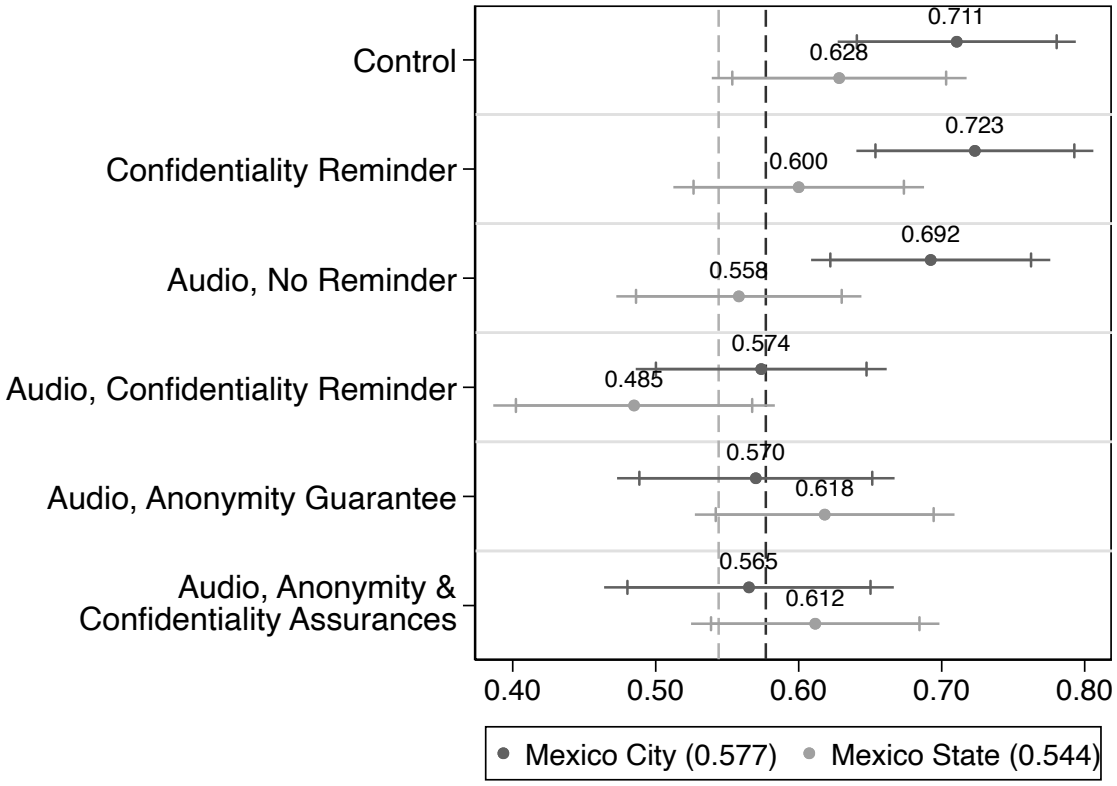
confidentiality assurances, respectively) are not statistically significant (*chi-squared* = 1.638, p = 0.239 and *chi-squared* = 0.01, p = 0.924, two-tailed, respectively). As designed, this study is underpowered to detect differences across audio treatments at standard levels of statistical significance; it is possible that we would detect differences in a larger sample.

Beyond decreased item non-response, we expected to observe improved response *accuracy* in conditions with more stringent confidentiality and anonymity assurances (H_4). To evaluate H_4, we computed AMLO's vote share in each state (the lowest level at which the data are representative) using data from Mexico's electoral management body, *Instituto Nacional Electoral* (INE). Figure 3 compares AMLO's estimated vote share in our data to official figures. AMLO's official vote share in Mexico City and Mexico State is denoted using color-coded dashed lines. Whiskers represent 95% confidence intervals (two-tailed) around point estimates, with vertical hash marks at 90% confidence; estimates that are *not* significant are not statistically different from AMLO's observed vote share. That is, insignificant results indicate more accurate reporting.

Over-reporting a vote for the winner varies considerably across conditions. In Mexico City, denoted in dark gray in the Figure, AMLO's vote share (57.7% according to INE) is significantly over-reported in both interviewer conditions (difference = 13.4 percentage points in the control condition, *chi-squared* = 9.88, p = 0.002; difference = 14.6 percentage points in the confidentiality reminder, *chi-squared* = 11.96, p = 0.001), and in the audio condition without any assurances (difference = 11.5 percentage points, *chi-squared* = 7.30, p = 0.007). However, in the remaining audio conditions, AMLO's predicted vote is not distinguishable from his official vote share. That is, we observe no over-reporting of votes for the winner in audio conditions including

confidentiality or anonymity assurances in Mexico City. This constitutes strong evidence in support of H_4.

Figure 3. AMLO's Predicted and Observed Vote Share in Mexico City, State



Note: Figure presents predicted probabilities of voting for AMLO, estimated following logit models regressing reported vote for AMLO on treatment condition in each state. Models are estimated without controls, weights, or adjusted standard errors. Dark gray dots are probabilities for Mexico City, light gray dots for Mexico State. Dashed lines indicate AMLO's official voteshare in Mexico City (dark gray) and Mexico State (light gray), per INE. Whiskers represent 95% confidence intervals, with vertical hash marks at 90% confidence, around the estimates (two-tailed). Whiskers that cross the associated dashed line are not distinguishable from AMLO's official vote share in that state with $p < 0.05$, two-tailed.

In Mexico State, AMLO's estimated vote share is marginally different from his official vote (54.4% according to INE) in the control condition (difference = 8.4 percentage points, *chi-squared* = 3.58, $p = 0.058$). In all other conditions, including the confidentiality reminder and all audio conditions, AMLO's reported vote share is not distinguishable from official election results.

As in Mexico City, these results are robust to controlling for demographic imbalances across treatments (SM Tables C3 and C4 provide complete results).

Results: Mechanisms

We expected that beliefs about response confidentiality would be significantly higher when respondents received an anonymity guarantee. To test this expectation, at the end of the survey we asked: "How much do you believe that your response to the question about your vote in the 2018 presidential election was confidential, that is, that I don't know what response you gave?" Response options were "a lot," "some," "a little," or "not at all"; we recoded this variable so that higher values indicate greater perceived confidentiality. Thirty-six participants (2.5% of voters) did not answer this item and are dropped from the analysis, yielding 1,396 total respondents. Participants in the "most anonymous" condition (audio with anonymity and confidentiality assurances) were more likely to respond to this item than those in the control group (coefficient = -1.152, t = -1.93, p = 0.053). Table 2 presents the association between confidentiality perceptions and treatment attributes (audio-administered, confidentiality assurances, and anonymity guarantees). While respondents in groups that received an audio treatment did not perceive greater anonymity than other respondents (coefficient = 0.011, t = 0.19, p = 0.870), those who received anonymity guarantees believed their response to the vote choice question was more confidential, on average, than respondents from all other groups (coefficient = 0.242, t = 3.10, p = 0.002), providing support for H_{1M}.

Table 2. Anonymity, Novelty Mechanisms by Treatment Condition

	Anonymity Perceptions (all voters)		Audio Enjoyment (all voters)	
Treatment Condition *Vote Choice*	Coefficient (SE)	t-statistic (P-value)	Coefficient (SE)	t-statistic (P-value)

Audio Treatment	0.011	0.19	0.382	5.89
	(0.067)	(0.870)	(0.065)	(0.000)
Confidentiality Assurance	-0.093	-1.38		
	(0.067)	(0.168)		
Anonymity Guarantee	0.242	3.10		
	(0.078)	(0.002)		
Constant	3.039	51.76	2.869	54.25
	(0.059)	(0.000)	(0.053)	(0.000)
Observations	1,396		1,271	

Note: Table presents OLS estimates, regressing each dependent variable on treatment attributes. Models estimated without covariates, weights, or adjusted standard errors.

Beliefs about response anonymity are also associated with response to the vote choice question. Pooling across treatments, 10.6% of those who believed that their responses were not anonymous at all refused to answer the vote choice question, compared to 4.0% of respondents who believed that their responses were very anonymous.

One alternative explanation is that item novelty drives decreasing item non-response. If respondents find the audio item more engaging than the rest of the survey, we expect them to say as much (H_{2M}). Re-engagement could occur because the audio technology itself is novel and engaging. Alternatively, the content of the audio question may matter, with only novel questions increasing engagement.

To differentiate between the effects of novel technology and novel content, we included a second audio treatment later in the questionnaire. All survey respondents were asked: "Which of the following players from the Mexican national soccer team do you think played the best in the [2018] World Cup in Russia?" This item mirrors the vote choice item by providing a list of five numbered options (prominent soccer players Giovanni Dos Santos, Javier (el "Chicharito") Hernández, Hirving Lozano, Héctor Herrera, and "other"), randomly varying the order of those options, and asking respondents to provide only the number, not the associated name. We asked about Mexico's national soccer team because the World Cup took place during the presidential

election (Mexico was eliminated on July 2nd, one day after the election) and received substantial media attention. We expected most citizens would remember this recent, high-salience event. If the novelty of the question's content drives down item non-response, we should see similar item response rates to the audio- and interviewer-administered soccer questions. We also expect that respondents will report greater enjoyment of the audio-administered soccer question, compared to the vote choice question.

The same speaker recorded audios for the soccer and vote choice items. The question included no confidentiality or anonymity reminders, as it was designed to isolate the effect of novelty. All voters who were asked the vote choice question by interviewer were assigned to receive the soccer question by audio, while half of abstainers were randomly assigned to receive the soccer audio treatment. Pooling across all participants, 38.9% of the sample received the soccer item via audio and 61.1% via interviewer. Of the 733 respondents who were assigned the soccer audio question, 87 were not treated. We find no imbalances in sociodemographic factors across treatment and control (see SM Table B1). As before, we analyze respondents according to their original group assignment.

We analyze the results separately for voters and abstainers because randomization was blocked by turnout. Item non-response to the soccer item was very high: 31.5% of all respondents did not answer this question. We expected and observed higher item non-response among women (39.0%, compared to 23.3% among men).[5] Among voters and abstainers, we do not observe significant differences in item non-response in the audio condition (SM Table E2).

The novelty hypothesis argues that respondents should favorably evaluate the survey item using novel technology. We assess this expectation by analyzing response to a question that read,

[5] In exploratory analyses we find that item non-response to the soccer question increased among men in the audio condition, but not among women (SM Table E4). We find no differences for the vote choice experiment by gender.

"And now, thinking about the question that you received by audio and comparing it to the other questions you answered, did you enjoy this question much more, more, the same, less, or much less than the other questions?" We recoded this variable so that higher values indicate more enjoyment of the audio item. The audio enjoyment question was only asked of successfully treated voters. Of 1,432 self-identified voters, 148 received no audio treatment; 13 (1.0%) successfully treated respondents did not answer the enjoyment question, yielding 1,271 voters in this analysis. Contrary to the novelty hypothesis, results presented in the second column of Table 2 show that voters who received the vote choice question by audio reported significantly *more* enjoyment of the audio item compared to those who received the audio-administered soccer question. While this difference is small (0.382 units on the five-point scale, or about one-third of a standard deviation), it is statistically significant ($t = 5.891$, $p = 0.000$, two-tailed). We cannot make this comparison among non-voters, as half of abstainers received no audio question.

In SM Table E3, we present the results for a second implication of this argument, that respondents who receive the "most novel" item will report greater enjoyment of the survey overall. We find no significant effect of the soccer treatment on survey enjoyment. However, enjoying the audio-administered vote choice question is a strong, negative predictor of item non-response to the vote choice question. In sum, enjoyment of the audio item appears to shape item non-response; the novelty of a question's *content* does not.

Discussion

Item non-response is a persistent problem for survey researchers, particularly when questions address sensitive topics such as vote choice. The tools scholars have developed to encourage response to sensitive items (e.g., list experiments, the randomized response technique) often

involve asking sensitive questions indirectly, which increases uncertainty around estimates and requires large samples. By asking a sensitive question directly while credibly concealing respondents' answers, we bypass these issues.

Implementing this technique substantially improved both the *rate* of item response to the vote question – regularly a sensitive item in public opinion surveys – and the *accuracy* of responses, yielding results very similar to official vote totals. Across all audio treatments, item non-response dropped by 43% compared to the control condition. When respondents received a credible anonymity guarantee but no additional confidentiality reminder, item non-response dropped by 66% (from 8.1% to 2.8%). Further, when participants were provided credible assurances of response confidentiality or anonymity, reported votes for the winner were statistically indistinguishable from official results. For researchers interested in electoral politics, this study suggests that even several months after an election, reporting bias can be remedied through the use of relatively inexpensive audio technology in face-to-face surveys.

Adding a verbal confidentiality reminder prior to asking a standard vote question did not improve item response, and was associated with substantial over-reporting of votes for the winner in Mexico City. Unlike some past studies (e.g., Frey 1986), we do not observe backfiring following the reminder. Rather, consistent with recent work (Leon et al. 2021), we find that such reminders do not improve item response. Finally, we find that perceptions of response anonymity and enjoyment of the audio item – independent of the novelty of the audio item's *content* – are associated with improvements in item non-response.

We use audio technology to improve item response to one sensitive question, vote choice, in a developing context. However, we think that the utility of this technique likely extends to other sensitive questions and contexts, as well. Our treatment inhibits the ability of "social referents" to

uncover respondents' answers; we therefore expect the treatment would improve item response in any circumstance where a question's sensitivity hinges on interviewers' and bystanders' ability to discover a respondent's answer. We expect that this method would improve response to sensitive items with unordered, categorical response options, including items using a binary, yes/ no logic. Our method is therefore well suited to answer questions about individuals' experiences with corruption, vote-buying exchanges, or election violence, among others. Randomizing the order of numbered response options across interviews and providing respondents with numbered response options would allow researchers to ask these items directly, while also protecting response anonymity from interested third parties.

References

Anderson, Jon Lee. 2018. “A New Revolution in Mexico.” *The New Yorker*. https://www.newyorker.com/magazine/2018/06/25/a-new-revolution-in-mexico#.

Andersson, Frida, and Steffan I. Lindberg. 2016. “2016 V-Dem Institute *Country Brief Mexico*.” https://www.v-dem.net/media/filer_public/7a/0c/7a0cd24b-b994-4fbb-bde6-9943de44b176/country_brief_mexico.pdf.

Anduiza, Eva, and Carol Galais. 2017. "Answering without reading: IMCs and strong satisficing in online surveys." *International Journal of Public Opinion Research* 29(3): 497-519.

Atkeson, Lonna R. (1999). “Sure, I voted for the winner!” Overreport of the primary vote for the party nominee in the national election studies. *Political Behavior* 21(3): 197-215.

Backor, Kristen, Saar Golde, and Norman Nie. 2007. "Estimating survey fatigue in time use study." *International Association for Time Use Research Conference. Washington, DC*.

Bates, Scott C., and Jared M. Cox. 2008. "The impact of computer versus paper–pencil survey, and individual versus group administration, on self-reports of sensitive behaviors." *Computers in human behavior* 24(3): 903-16.

Berinsky, Adam J. 1999. "The two faces of public opinion." *American Journal of Political Science*: 1209-30.

Berinsky, Adam J. 2008. Survey non-response. *The SAGE handbook of public opinion research*: 309-22.

Blair, Graeme, Alexander Coppock, and Margaret Moor. 2020. "When to worry about sensitivity bias: A social reference theory and evidence from 30 years of list experiments." *American Political Science Review* 114(4): 1297-1315.

Brough, Aaron R., David A. Norton, Shannon L. Sciarappa, and Leslie K. John. 2022. "The Bulletproof Glass Effect: Unintended Consequences of Privacy Notices." *Journal of Marketing Research* 59(4): 739-54.

Burchell, Brendan, and Catherine Marsh. 1992. “The effect of questionnaire length on survey response." *Quality and Quantity* 26(3): 233-44.

Bush, Sarah Sunn, and Lauren Prather. 2019. "Do electronic devices in face-to-face interviews change survey behavior? Evidence from a developing country." *Research & Politics* 6(2): 2053168019844645.

Bishop, George F., and Bonnie S. Fisher. 1995. "“Secret ballots” and self-reports in an exit-poll experiment." *Public Opinion Quarterly* 59(4): 568-588.

Carreras, Miguel. 2016. “Institutions, Governmental Performance and the Rise of Political Newcomers.” *European Journal of Political Research*: 1–17.

Castillo-Manzano, José I., Lourdes López Valpuesta, and Rafael de Pozo Barajas. 2018. “At a time of insurgent parties, can societies believe in election polls? The Spanish experience.” *Revista de Economía Aplicada* 26(78): 81-97.

Chauchard, Simon. 2013. "Using MP3 players in surveys: The impact of a low-tech self-administration mode on reporting of sensitive attitudes." *Public Opinion Quarterly* 77(1): 220-31.

Chen, Sixia, and David Haziza. 2019. “Recent developments in dealing with item non-response in surveys: a critical review.” *International Statistical Review* 87: S192-S218.

Cho, Hyunyi, and Robert LaRose. 1999. “Privacy issues in Internet surveys.” *Social science computer review* 17(4): 421-34.

Cohen, Mollie J., and Sebastian Larrea. 2018. “Assessing and Improving Interview Quality in the 2016/17 AmericasBarometer.” LAPOP Lab Methodological Note No. 2.

Cohen, Mollie J. and Zach Warner. 2020. “How to get better survey data more efficiently.” *Political Analysis* 29(2): 121-38. DOI: http://dx.doi.org/10.1017/pan.2020.20.

Connors, Elizabeth C., Yanna Krupnikov, and John Barry Ryan. 2019. "How Transparency Affects Survey Responses." *Public Opinion Quarterly* 83(1): 185-209.

Coppock, Alexander. 2017. “Did shy Trump supporters bias the 2016 polls? Evidence from a nationally-representative list experiment.” *Statistics, Politics and Policy* 8(1): 29-40.

Corbacho, Ana, Daniel W. Gingerich, Virginia Oliveros, and Mauricio Ruiz-Vega. 2016. “Corruption as a self- fulfilling prophecy: evidence from a survey experiment in Costa Rica.” *American Journal of Political Science* 60(4): 1077-92.

Couper, Mick P., Eleanor Singer, Frederick G. Conrad, and Robert M. Groves. 2008. "Risk of disclosure, perceptions of risk, and concerns about privacy and confidentiality as factors in survey participation." *Journal of Official Statistics* 24(2): 255-75.

Crow, D. 2010. “The party's over: citizen conceptions of democracy and political dissatisfaction in Mexico.” *Comparative Politics* 43(1): 41-61.

Dykema, Jennifer, Kerryann Diloreto, Jessica L. Price, Eric White, and Nora Cate Schaeffer. 2012. "ACASI gender-of-interviewer voice effects on reports to questions about sensitive behaviors among young adults." *Public Opinion Quarterly* 76(2): 311-25.

Ermakova, Tatiana, Benjamin Fabian, and Rüdiger Zarnekow. 2016. "Improving individual acceptance of health clouds through confidentiality assurance." *Applied Clinical Informatics* 7(04): 983-93.

Frey, James H. 1986. "An experiment with a confidentiality reminder in a telephone survey." *Public Opinion Quarterly* 50(2): 267-69.

Garcia-Yi, Jaqueline, and Ulrike Grote. 2012. "Data collection: Experiences and lessons learned by asking sensitive questions in a remote coca growing region in Peru." *Survey Methodolog,* 38: 131-42.

Gerber, Alan S., Gregory A. Huber, David Doherty, and Conor M. Dowling. 2012. "Disagreement and the avoidance of political discussion: Aggregate relationships and differences across personality traits." *American Journal of Political Science* 56(4): 849-74.

Gnambs, Timo, and K. Kaspar. 2015. "Disclosure of sensitive behaviors across self-administered survey modes: a meta-analysis." *Behavior Research Methods* 47(4): 1237-59.

Karpowitz, Christopher F., J. Quin Monson, Lindsay Nielson, Kelly D. Patterson, Steven A. Snell. 2011. "Political Norms and the Private Act of Voting." *Public Opinion Quarterly* 75(4): 659–85. https://doi.org/10.1093/poq/nfr024

Kennedy, Courtney, Mark Blumenthal, Scott Clement, Joshua D Clinton, Claire Durand, Charles Franklin, Kyley McGeeney, Lee Miringoff, Kristen Olson, Douglas Rivers, Lydia Saad, G Evans Witt, Christopher Wlezien. "An Evaluation of the 2016 Election Polls in the United States." *Public Opinion Quarterly* 82(1): 1–33. https://doi.org/10.1093/poq/nfx047.

Krosnick, Jon A. 1999. "Survey research." *Annual Review of Psychology* 50(1): 537-67.

Krosnick, Jon A. and Stanley Presser. 2010. "Question and Questionnaire Design." *Handbook of Survey Research*: 263-314.

Kuran, Timur. 1991. "Now out of never: The element of surprise in the East European revolution of 1989." *World Politics: A Quarterly Journal of International Relations* 7-48.

Lavender, Jason M., and Drew A. Anderson. 2009. "Effect of perceived anonymity in assessments of eating disordered behaviors and attitudes." *International Journal of Eating Disorders* 42(6): 546-51.

Lelkes, Yphtach, Jon A. Krosnick, David M. Marx, Charles M. Judd, and Bernadette Park. 2012. "Complete anonymity compromises the accuracy of self-reports." *Journal of Experimental Social Psychology* 48(6): 1291-99.

Leon, Carmen M, Eva Aizpurua, Sophie van der Valk. 2021. "The Impact of Confidentiality Assurances on Participants' Responses to Sensitive Questions." *International Journal of Public Opinion Research* 33(4): 1024–38. https://doi.org/10.1093/ijpor/edaa039

Lupu, Noam, and Kristin Michelitch. 2018. "Advances in Survey Methods for the Developing World." *Annual Review of Political Science* 21: 195-214.

Lynn, Peter. 2017. "From standardised to targeted survey procedures for tackling non-response and attrition." *Survey Research Methods* 11(1): 93-103. https://doi.org/10.18148/srm/2017.v11i1.6734

Joinson, Adam N., Carina Paine, Tom Buchanan, and Ulf-Dietrich Reips. 2008. "Measuring self-disclosure online: Blurring and non-response to sensitive items in web-based surveys." *Computers in Human Behavior* 24(5): 2158-71.

Madson, Gabriel J., and Sunshine D. Hillygus. 2020. "All the best polls agree with me: Bias in evaluations of political polling." *Political Behavior* 42(4): 1055-72.

Montalvo, J. Daniel, Mitchell A. Seligson, and Elizabeth J. Zechmeister. 2018. "Improving Adherence to Area Probability Sample Designs: Using LAPOP's Remote Interview Geo-

Locating of Households in Real-Time (RIGHT) System." LAPOP Lab Methodological Note No. 4.

Moore, Roland S., and Genevieve M. Ames. 2002. "Survey confidentiality vs. anonymity: young men's self-reported substance use." *Journal of Alcohol and Drug Education* 47(2): 32.

Morin-Chassé, Alexandre, Damien Bol, Laura B. Stephenson, and Simon Labbé St-Vincent. 2017. "How to survey about electoral turnout? The efficacy of the face-saving response items in 19 different contexts." *Political Science Research and Methods* 5(3): 575-84.

Nanes, Matthew, and Dotan Haim. 2020. "Self-Administered Field Surveys on Sensitive Topics." *Journal of Experimental Political Science*: 1-8.

NIMH Multisite HIV/STD Prevention Trial for African American Couples Group. 2008. "Designing an audio computer-assisted self-interview system in a multisite trial: A brief report." *Journal of acquired immune deficiency syndromes* 49(Suppl 1): S52-S58.

Ong, Anthony D., and David J. Weiss. 2000. "The impact of anonymity on responses to sensitive questions 1." *Journal of Applied Social Psychology* 30(8): 1691-1708.

Osorno, Guillermo. 2018. "The Riddle of López Obrador." *The New York Times*: 4–7. https://www.nytimes.com/2018/06/17/opinion/mexico-andres-manuel-lopez-obrador-amlo.html.

Panagopoulos, Costas, Kyle Endres, and Aaron C. Weinschenk. 2018. "Preelection poll accuracy and bias in the 2016 US general elections." *Journal of Elections, Public Opinion and Parties* 28(2): 157-72.

Peytchev, Andy. 2013. "Consequences of survey nonresponse." *The ANNALS of the American Academy of Political and Social Science* 645(1): 88-111.

Riphahn, Regina T., and Oliver Serfling. 2005. “Item non-response on income and wealth questions.” *Empirical Economics* 30(2): 521-38.

Robertson, Ronald E., Felix W. Tran, Lauren N. Lewark, and Robert Epstein. 2018. "Estimates of non-heterosexual prevalence: the roles of anonymity and privacy in survey methodology." *Archives of Sexual Behavior* 47(4): 1069-84.

Schettino, Marcario. 2017. “Mexico 2018: It’s Not the Economy, Stupid.” *Americas Quarterly*: 1–4. http://americasquarterly.org/content/mexico-2018-its-not-economy-stupid.

Singer, Eleanor. 1993. "Informed consent and survey response: a summary of the empirical literature." *Journal of Official Statistics* 9(2): 361.

Singer, Eleanor, Hans-Juergen Hippler, and Norbert Schwarz. 1992. "Confidentiality assurances in surveys: Reassurance or threat?." *International Journal of Public Opinion Research* 4(3): 256-68.

Singer, Eleanor, Nancy A. Mathiowetz, and Mick P. Couper. 1993. “The impact of privacy and confidentiality concerns on survey participation the case of the 1990 US Census.” *Public Opinion Quarterly* 57(4): 465-82.

Singer, Eleanor, John Van Hoewyk, and Randall J. Neugebauer. 2003. "Attitudes and behavior: The impact of privacy and confidentiality concerns on participation in the 2000 census." *Public Opinion Quarterly* 67(3): 368-84.

Singer, Eleanor, Dawn R. Von Thurn, and Esther R. Miller. 1995. "Confidentiality assurances and response: A quantitative review of the experimental literature." *Public Opinion Quarterly* 59(1): 66-77.

Sturgis, Patrick, Jouni Kuha, Nick Baker, Mario Callegaro, Stephen Fisher, Jane Green, Will Jennings, Benjamin E. Lauderdale, and Patten Smith. 2018. “An assessment of the causes

of the errors in the 2015 UK general election opinion polls." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 181(3): 757-81.

Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. 2000. *The psychology of survey response*. Cambridge University Press.

Tourangeau, Roger, and Ting Yan. 2007. "Sensitive questions in surveys." *Psychological Bulletin* 133(5): 859.

Tourangeau, Roger, and Tom W. Smith. 1996. "Asking sensitive questions: The impact of data collection mode, question format, and question context." *Public Opinion Quarterly* 60(2): 275-304.

Warner, Zach, and Gabriel N. Camargo-Toledo. 2019. "How Does LAPOP Calculate Response Rates?" LAPOP Lab Methodological Note No. 5.

Wells, Chris, Katherine J. Cramer, Michael W. Wagner, German Alvarez, Lewis A. Friedland, Dhavan N. Shah, Leticia Bode, Stephanie Edgerly, Itay Gabay, and Charles Franklin. 2017. "When we stop talking politics: The maintenance and closing of conversation in contentious times." *Journal of Communication* 67(1): 131-57.

Whelan, Thomas J. 2007. "Anonymity and confidentiality: Do survey respondents know the difference." *Poster presented at the 30th annual meeting of the Society of Southeastern Social Psychologists, Durham, NC*.

Yan, Ting, Richard Curtin, and Matthew Jans. 2010. "Trends in income nonresponse over two decades." *Journal of Official Statistics* 26(1): 145.