Do electronic devices in face-to-face interviews change survey behavior? Evidence from a developing country

Sarah Sunn Bush¹ and Lauren Prather²

Abstract
A large literature shows that survey mode and survey technologies significantly affect item non-response and response distributions. Yet as researchers increasingly conduct surveys in the developing world, little attention has been devoted to understanding how new technologies—such as the use of electronic devices in face-to-face interviews—produce bias there. We hypothesize that using electronic devices instead of pen and paper can affect survey behavior via two pathways: a wealth effect and a surveillance effect. To test the hypotheses, we use data from a two-wave panel survey fielded in Tunisia. We investigate whether responses collected in Wave 1 with pen and paper changed when some individuals were interviewed in Wave 2 by interviewers using tablet computers. Consistent with the wealth effect hypothesis, more than half of the lowest income respondents reported a higher income in the second wave when interviewers used tablets. Conversely, we find little evidence that concerns about surveillance changed survey behavior.

Keywords
Public opinion, survey methods, developing countries

Introduction
Survey research in the developing world is typically conducted via face-to-face interviews. Traditionally, interviewers used paper-assisted personal interviewing (PAPI) to record responses, which would later be converted to a machine-readable format. Today, researchers increasingly employ computer-assisted personal interviewing (CAPI), using electronic devices such as tablets, smart phones, and laptops to record responses during face-to-face interviews.

CAPI is now used by many leading public opinion projects that rely on face-to-face interviews, such as the World Values Survey and Americas Barometer (Montalvo et al., 2019). In addition, independent researchers have produced high-quality studies using CAPI in such diverse places as Bangladesh (Winters et al., 2017), Lesotho (Clayton, 2018), and Peru (Hawkins et al., 2017). A March 2017 survey that we conducted of 76 political scientists who have fielded at least one survey in the developing world in the past five years confirmed the rising use of CAPI. Fully 84% of those surveyed reported having conducted their last survey using face-to-face methods. Of those, 53% reported that they used CAPI methods.¹ Researchers are switching to CAPI for myriad reasons including its cost, ease of use, better data quality, and ability to facilitate complex survey questionnaires.

Despite the increased use of CAPI in the developing world, to our knowledge, no previous study has examined its effects on survey behavior.² To explore these effects, our study exploits variation in the use of CAPI and PAPI methods in a nationally-representative panel survey of political attitudes in Tunisia. In Wave 1, interviewer teams conducted all interviews in person and recorded responses using pen and paper. In Wave 2, one member of each interviewer team recorded responses with hand-held tablets; the other interviewers used pen and paper again. Although the survey was not specifically designed to examine how CAPI

¹Department of Political Science, Yale University, USA
²School of Global Policy and Strategy, University of California San Diego, USA

Corresponding author:
Lauren Prather, School of Global Policy and Strategy, University of California San Diego, 9500 Gilman Drive, #0519, La Jolla, CA 92093, USA.
Email: lprather@ucsd.edu
changes survey behavior, we leverage the survey questions and design to test hypotheses about reported household income and government support. We find that CAPI generally did not affect item non-response, but we identify a significant effect on response distributions: half of the individuals with below-median household incomes as recorded in Wave 1 reported higher incomes when interviewed using tablets in Wave 2. This effect is consistent with our hypothesis that electronic devices increase the social distance between interviewer and respondent, introducing bias as lower income respondents change their responses about income to bridge that distance. In contrast, tablets did not cause people to report greater support for leaders or avoid answering questions about them.

The CAPI effect on reported income is noteworthy and encourages researchers to think carefully about survey design when seeking to accurately measure this core concept. Our survey of survey researchers suggests that some scholars already worry that tablets introduce bias; one individual wrote in an open-ended response that electronic devices “generate additional social desirability bias because they are so rare.” Although our null findings for the other hypotheses may imply that a shift from PAPI to CAPI will not alter survey data in the developing world, studies using other research designs or in different contexts could reveal more significant effects. Thus, we conclude this paper with a call for further research.

The effects of CAPI in developing countries

We build on the large survey methods literature to develop hypotheses about how the use of electronic devices in face-to-face interviews could affect item non-response and response distributions. First, we note that electronic devices such as smart phones and tablets are rare and high-status items in many countries. Although ownership is increasing, the median ownership rate in “emerging and developing nations” is 24% according to a 2014 survey of 32 countries (Pew Research Center, 2015: 11). Moreover, there is a significant digital divide, with better educated and wealthier individuals being more likely to own a smart phone.

We hypothesize that the presence of a high-status electronic device during a CAPI interview in a developing country will activate social desirability effects related to income. Although survey interviewers are higher status than many respondents, electronic devices are a salient marker of wealth and feature prominently during interviews. The literature on social desirability bias demonstrates that survey respondents often try to express what they perceive as the socially desirable answer. In general, more social desirability bias tends to occur when responses are shared directly with an interviewer rather than personally completing a questionnaire. Interviewer characteristics (e.g., race, gender, and clothing) can further magnify this bias (e.g., Davis and Silver, 2003; Benstead, 2013; Blaydes and Gillum, 2013). We therefore expect that the presence of electronic devices in a CAPI survey will raise the sensitivity of questions about household income in developing countries, particularly among the poor. The mechanism is that an interviewer with an electronic device may appear to be of a higher status than an interviewer with pen and paper, since possessing such a device suggests something about the wealth of the interviewer or her employer. Thus, in the presence of an interviewer carrying a tablet, low-income respondents may feel embarrassed to report their low incomes to the interviewer.

These social desirability dynamics could have two effects. The first is that low-income respondents will be less comfortable telling their interviewer information about their economic status. In other words, these respondents will be less likely to respond to questions related to income when interviewed in the presence of an electronic device (H1). Additionally, the presence of an electronic device makes higher incomes seem more socially desirable because it suggests the interviewer has a relatively high socio-economic status. To avoid the feelings of embarrassment that might accompany the reporting of a low income to a high-status interviewer, lower income respondents will report a higher income than they otherwise would have (H2).

Turning to our second set of hypotheses, we note that governments often sponsor public opinion surveys, including in partially-democratic and authoritarian environments (Nalepa and Pop-Eleches, 2013, Corstange, 2014). In such contexts, governments may conduct surveys to gather information about their popular support. The use of surveys as a form of government surveillance could cause individuals to be suspicious about the consequences of participation.

Electronic devices have certain characteristics that make them worrisome from the perspective of surveillance. In the traditional PAPI method of data collection for face-to-face interviews, interviewers can only collect information about respondents that they are able to write down. If the interviewer does not ask the respondent identifying questions, then the respondent may be reassured that the process is anonymous. In contrast, most smart phones and tablets allow interviewers to record respondents using audio or video, often without any obvious sign or sound. For this reason, electronic devices make surveillance via surveys easier.

We hasten to note that recording respondents in this manner is not typical in academic surveys. However, in countries that are undemocratic—or that have recently transitioned to democracy—the public may be worried about surveillance and suspicious of people who are conducting surveys (Driscoll and Hidalgo, 2014). As such, the possibility of being recorded while expressing sensitive opinions may change responses. Such concerns should be especially relevant for people who oppose government leaders, since they would be more likely to suffer negatively from being watched.
Concerns about surveillance by electronic devices could have at least two effects. The first is that government opponents will feel less comfortable expressing their views about the government in the presence of electronic devices. The idea is that government opponents will be more worried than government supporters that their responses could be used against them. Therefore, they will be less likely to respond to questions about the government (H3). The second potential effect is that government opponents may choose to respond but falsify their responses, reporting greater regime support when interviewed in the presence of an electronic device. Since the presence of an electronic device during an interview makes it easier to record the respondent’s voice or image, government opponents may report greater support than they would have otherwise (H4).

Research design

We test our hypotheses in Tunisia. As noted in the introduction, we exploit variation in the use of CAPI and PAPI survey methods in an otherwise-unrelated political attitudes survey (Bush and Prather, 2017, 2018). Tunisia is a plausible setting to observe the effects we hypothesize. First, at the time of our study in 2014, economic issues and inequality were highly salient. Moreover, smart phones and tablets signalled high status as only 12% of Tunisians owned them (Pew Research Center, 2015: 11). Second, Tunisia had recently emerged from decades of a repressive dictatorship. Observers have noted the “near-comprehensive surveillance of political activities” in authoritarian Tunisia (Kallander, 2011). Although Tunisia had begun a democratic transition at the time of our study, many former regime members held political positions in the post-revolution government and were competing in the 2014 elections. Some non-governmental associations’ leaders expressed concerns about surveillance by the government during the transitional period (Bush, 2015: 204).

Figure 1 situates Tunisia within the context of developing countries in 2014 according to the proportion of their citizens who owned smart phones (Pew Research Center, 2015) and levels of freedom. On the one hand, it reveals that Tunisia was among the countries surveyed by Pew with the lowest levels of smart phone and tablet ownership. We expect the wealth-priming effects of CAPI to be stronger in such environments because electronic devices are more likely to be high-status items. In contrast, we expect wealth-priming effects to be attenuated when most people own smart phones and tablets. On the other hand, Figure 1 shows that Tunisia was among the countries surveyed by Pew with the most freedom in 2014, with an overall designation of “partly free.” As such, and despite the country’s authoritarian past, Tunisians may not have been as worried about surveillance as citizens in other countries. Stronger surveillance effects could obtain in less free settings such as China, Egypt, Vietnam, or Uganda.

ELKA Consulting conducted the panel survey. Tunisian interviewers completed the interviews in person using Arabic. Although CAPI is increasing in developing countries, it was not commonplace in Tunisia in 2014. In fact, this study was the first by ELKA to collect data via CAPI. Some survey firms operated in Tunisia prior to the revolution, including our partner ELKA Consulting, although it focused on market research then.

Wave 1 took place after the parliamentary election in October 2014 and interviewed 1400 adults. The survey was nationally representative and sampled both male and female adult Tunisians. The response rate for Wave 1 was 71%. Wave 2 followed the presidential runoff election in December 2014. Multiple attempts were made to contact people from Wave 1 to minimize attrition. 1107 people were re-interviewed. Whenever possible, interviewers from Wave 1 conducted the Wave 2 interviews. Even with attrition, the sample was similar to the population of Tunisia on several key dimensions. Further details about the sample are in the Online Appendix.

In Wave 1, interviewers recorded responses using PAPI. In Wave 2, one member of each interviewer team was trained to record responses using Acer tablets. The other interviewers continued to use PAPI. This design created two groups of respondents: those in the CAPI condition (i.e., with interviewers that would be given tablets in Wave 2) and those in the PAPI condition (i.e., with interviewers that would not be given tablets in Wave 2). In all, 291 interviews in Wave 2 were completed using tablets. ELKA introduced CAPI to a subset of respondents in Wave 2 as a way of piloting CAPI while maintaining quality control.

We use two identification strategies. First, we use a within-subjects design to examine whether the shift from PAPI to CAPI caused respondents to respond less often or differently to a question about household income. Because we asked the same income question in both waves, we can estimate the effect of CAPI independent of the effect of the interviewers who were eventually given tablets. Though we cannot be sure that the CAPI and PAPI respondents’ incomes were trending in the same direction after Wave 1,
the short amount of time (two months) that elapsed between the two surveys made major changes in income unlikely. Second, we use a between-subjects design to examine whether CAPI respondents responded differently to PAPI respondents when asked a question about government support. We examine differences between subjects in Wave 2 since the same question about government support was not asked in both waves.

An important design consideration was how to assign tablets to interviewers. The primary concern was with maintaining data quality for the political attitudes survey. ELKA therefore considered interviewers’ ability to use, care for, and secure the devices. These considerations prompted the tablets to be assigned to the leader of each interviewer team. Leaders tended to be more experienced and slightly older than the other interviewers but similar in many other ways (e.g., education level, gender, attire, region of origin), as discussed in detail in the SI. Thus, the CAPI intervention is a bundled “treatment” with the effects identified below resulting from the “CAPI plus interviewer” combination. Estimating the effect of this bundled treatment is relevant since using CAPI requires hiring interviewers who can use electronic devices successfully and responsibly.

Although CAPI was not assigned to interviewers at random, CAPI was in expectation assigned to respondents at random since interviewers were assigned randomly to respondents in Wave 1. We check for balance across CAPI and PAPI interviewers in Wave 2 using measures recorded in Wave 1. In other words, we examine potential differences across respondents when all interviewers used PAPI. Doing so is particularly important for the between-subjects analysis of H3 and H4.

To assess balance, we check for differences between the CAPI and PAPI conditions on key individual characteristics of respondents measured in Wave 1 using comparison of means tests. Since we are unable to directly examine the parallel trends assumption that underpins a difference-in-differences analysis, balance tests help ensure that the individuals who were interviewed using CAPI in Wave 2 were similar to other individuals on observable characteristics in Wave 1. As Table 1 shows, the two groups did not exhibit statistically significant differences in terms of age, gender, employment status, or political interest in Wave 1. However, they reported higher incomes in Wave 1, were less likely to support the Nidaa Tounes political party, had higher levels of education, and had lower levels of political knowledge. In the between-subjects analysis below, we control for the variables that were not well balanced during Wave 1, although our findings are substantively similar without these controls (see SI). To shed further light on the likelihood of parallel trends, the SI discusses key events in Tunisia between Waves 1 and 2 and a placebo test.

### Results

Both waves asked, “What is your monthly family income in Tunisian dinars?” Respondents were given six answer options in keeping with how this question is typically asked in Tunisia. Figure 2 shows the distribution of responses in Wave 1. The median respondent had a household income of 200–500 dinars. Thus, we consider “relatively poor”

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean CAPI</th>
<th>Mean PAPI</th>
<th>P-value for test that: CAPI=PAPI</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income</td>
<td>2.46</td>
<td>2.30</td>
<td>.04</td>
<td>1033</td>
</tr>
<tr>
<td>(1.085)</td>
<td>(1.022)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nidaa Supporter</td>
<td>0.69</td>
<td>0.78</td>
<td>.01</td>
<td>1063</td>
</tr>
<tr>
<td>(0.462)</td>
<td>(0.417)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>39.93</td>
<td>40.22</td>
<td>.80</td>
<td>1084</td>
</tr>
<tr>
<td>(16.196)</td>
<td>(16.316)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>0.54</td>
<td>0.53</td>
<td>.80</td>
<td>1084</td>
</tr>
<tr>
<td>(0.500)</td>
<td>(0.500)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>3.25</td>
<td>3.01</td>
<td>.01</td>
<td>1084</td>
</tr>
<tr>
<td>(1.376)</td>
<td>(1.341)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>0.36</td>
<td>0.37</td>
<td>.89</td>
<td>1084</td>
</tr>
<tr>
<td>(0.482)</td>
<td>(0.482)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Political Interest</td>
<td>2.25</td>
<td>2.28</td>
<td>.62</td>
<td>1084</td>
</tr>
<tr>
<td>(0.942)</td>
<td>(0.947)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Political Knowledge</td>
<td>1.94</td>
<td>2.04</td>
<td>.07</td>
<td>1107</td>
</tr>
<tr>
<td>(0.851)</td>
<td>(0.822)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: All means calculated from Wave 1 measures when all interviewers used PAPI. The CAPI column reports Wave 1 means of respondents interviewed by interviewers who used CAPI in Wave 2. The PAPI column reports Wave 1 means of respondents interviewed by interviewers who used PAPI in Wave 2. Standard deviations in parentheses.
respondents to be those who reported a household income of under 200 dinars. Since respondents placed themselves into income “bins,” it is unlikely that any CAPI effects resulted from rounding errors or slight increases.

We did not find support for H1, that relatively poor respondents would respond less to the income question. All but three of the poorer respondents as measured in Wave 1 reported their income in Wave 2. To test H2, we use a difference-in-differences approach. Our outcome variable is reported income, measured at two points in time: Wave 1 and Wave 2. We regress income on indicators for Wave 2, interviewers given tablets in Wave 2, and their interaction. The coefficient estimate for the interaction term is the average treatment effect (presented in Figure 3); a table with the full results is in the SI. As it shows, low-income respondents reported significantly higher incomes in Wave 2 in the presence of tablets. The increase is about 0.31 on the 6-point scale. To put this finding in context: more than half of the low-income respondents interviewed with tablets (26 out of 46) reported a higher income in Wave 2. For comparison, Figure 3 includes respondents who reported other income levels, among whom tablets did not have a clear effect. As shown in the SI, our findings are similar if we use ordered logistic regressions or regress Wave 2 income on variables indicating Wave 1 income, assignment to CAPI, and their interaction. Differences are substantial and unlikely to be due to random response instability or increases in true income, both of which have been similar for CAPI and PAPI respondents.

Next, we consider the surveillance hypotheses about the effects of CAPI on support for government leaders. Because identical questions about regime support were not asked in both waves, we analyze variation between subjects in Wave 2. Wave 1 asked respondents, “How satisfied are you with Nidaa Tounes [the ruling party]?“ Approximately one-quarter of the respondents responded that they were “dissatisfied” or “very dissatisfied” with Nidaa Tounes. We code them as government opponents. Wave 2 asked about satisfaction with the newly-elected president, Beji Caid Essebsi, also of Nidaa Tounes. We use responses to this question as our outcome variable.

Figure 2. Reported monthly household income in Tunisian dinars (dt), Wave 1.

Figure 3. Effect of tablets on reported income. The coefficient estimates and 95% confidence intervals are from OLS regressions of income on an indicator for assignment to the tablets in Wave 2, an indicator for Wave 2, and their interaction. The treatment effect is given by the interaction term. Standard errors were clustered on respondent. “Don’t know” responses were dropped.

Figure 4 presents our findings; again the SI contains the tables. We find no evidence that CAPI caused government opponents to respond less to a question about satisfaction with Essebsi (H3). Similarly, we do not find a clear effect of tablets on reported satisfaction with Essebsi among government opponents (H4). In the SI, we show that there is no evidence of a tablet effect that is conditional on age or education, two variables that could relate to individuals’ familiarity with tablets’ recording functions and thus concerns about surveillance.

Conclusion

This study is the first investigation of how electronic devices affect survey behavior in face-to-face interviews in the developing world. We show that low-income respondents reported higher incomes when interviewed using CAPI than when interviewed using PAPI. However, tablets did not cause people to respond less to an income question or alter responses to a question about regime support. These results are mostly encouraging for researchers using data collected via CAPI. At the same time, researchers should recognize that low-income respondents may report higher incomes with CAPI. One implication is that income inequality may appear lower with CAPI. This finding is relevant for social scientists, governments, and development organizations that rely on longitudinal data collected via face-to-face household surveys. Biases in income measures also matter for researchers interested in questions ranging from the sources of economic policy preferences to the role of economic deprivation in individuals’ decisions to participate in protest and violence.

Researchers can build on these findings in several ways. To begin, the effects of CAPI can be studied in other contexts, including less democratic environments.
In particular, the surveillance-priming effects of CAPI could emerge more clearly in contexts where surveillance is ongoing and salient. Such contexts are important sites for research on this topic and should include countries in the Middle East and North Africa, where the repressive environment has drawn researchers’ attention to the sensitivity of many political questions (Corstange, 2009; Truex and Tavana, forthcoming). In addition, scholars could develop and test hypotheses about how CAPI affects other survey responses. A final avenue would be to assign CAPI devices randomly to interviewers to disentangle the bundled treatment of CAPI and experienced interviewers that we present in this study, though we recommend extensive training for less-experienced interviewers.

We conclude by building on other researchers’ experiences with CAPI to offer several practical recommendations (Benstead et al., 2017). First, alongside more commonly-reported details such as sampling methods and response rates, researchers should report whether their survey was conducted in person and, if so, whether CAPI was used. Second, researchers conducting longitudinal surveys might switch gradually from PAPI to CAPI, using random assignment of CAPI when possible to assess any biases that electronic devices introduce. Finally, researchers should attempt to reduce the potential biases introduced by electronic devices. For example, a wealth-priming effect associated with tablets might be mitigated by using electronic devices that do not appear to be too expensive or by using objective measures of income. Although there is no magic bullet for dealing with social desirability biases, an extensive literature suggests several promising strategies.

**Acknowledgements**

We thank participants at seminars at Temple University, UCSD, and the Visions in Methodology workshop for their helpful feedback. We also thank John Ahlquist, Chantal Berman, Mohamed Ikbal Elloumi and Elka Consulting, Chad Hazlett, Elizabeth Nugent, Molly Roberts, Bryn Rosenfeld, Caroline Tolbert, and two anonymous reviewers for their assistance and feedback.

**Declaration of conflicting interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Science Foundation under Grant No. 1456505. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. This research received clearance from the Temple University Institutional Review Board via Protocol 22385 and the Stanford University Institutional Review Board via Protocol 31932 (“International Election Observation and Perceptions of Election Credibility: A Case Study in the Arab World”).

**Supplemental material**


**Notes**

1. For more details of this survey, see the Online Appendix.
2. In the 1990s, researchers began to examine the effects of CAPI on survey behavior in the USA, comparing the use of laptops during interviews to other methods. Many of these studies focused on the effects of self-administered surveys rather than the presence of computers (Baker, 1992; Baker and Johnson, 1995; Tourangeau and Smith, 1996). A more recent literature on CAPI focuses on implementation issues, such as survey length and measurement error (Escandon et al., 2008; Caeyers et al., 2012; Leisher, 2014) and recruitment (Li et al., 2018).
3. See the SI for a review of this literature.
4. Although some interviewees may realize that the device is not owned by the interviewer, the device conveys information about her employer’s status.
5. We averaged countries’ civil liberties and political rights scores, which range from 1 to 7. We rescaled the scores so that 7 represents the most free countries. Data from Freedom House are available at https://freedomhouse.org/report-types/freedom-world (accessed March 16, 2017).

6. Wave 1 interviewers made the first re-contact attempt. Subsequent attempts were sometimes made by other interviewers from the same team. As the SI shows, the income results are robust to dropping respondents who were not interviewed by the same interviewer in both waves.

7. As a robustness check, we also include the respondents in the 200–500 dinars category. As the SI shows, the wealth-priming effect is attenuated but in the same direction.

8. For this analysis, non-responses are dropped.

9. In the SI, we report results using alternative Wave 1 measures of regime support based on vote choice in the parliamentary election and intended vote choice in the presidential election.

References


