

Self-Administered Field Surveys on Sensitive Topics

Matthew Nanes
Saint Louis University*

Dotan Haim
Florida State University†

June 23, 2020

Abstract

Research on sensitive topics uses a variety of methods to combat response bias in in-person surveys. Increasingly, researchers allow respondents to self-administer responses using electronic devices as an alternative to more complicated experimental approaches. Using an experiment embedded in a survey in the rural Philippines, we test the effects of several such methods on response rates and falsification. We asked respondents a sensitive question about reporting insurgents to the police alongside a non-sensitive question about school completion. We randomly assigned respondents to answer these questions either verbally, through a “forced choice” experiment, or through self-enumeration. We find that self-enumeration significantly reduced non-response compared to direct questioning, but find little evidence of differential rates of falsification. Forced choice yielded highly unlikely estimates, which we attribute to non-strategic falsification. These results suggest that self-administered surveys can be an effective for measuring sensitive topics on surveys when response rates are a priority.

Acknowledgments

Nico Ravanilla played a substantial role in designing and implementing the survey. The authors thank Konstantin Ash, Kolby Hanson, Connor Huff, and Steven Rogers for their comments. This work was supported by Evidence in Governance and Politics Metaketa IV. Both authors received significant financial support for their work through Evidence in Governance and Politics, and the University of California San Diego Policy Design and Evaluation Lab. Additionally, Nanes received significant financial support from The Asia Foundation unrelated to this research, and Haim received significant financial support from the United Nations Development Program unrelated to this research. The data and code to replicate all analyses in this article are available at the Journal of Experimental Political Science Dataverse within the Harvard Dataverse Network, at: <https://doi.org/10.7910/DVN/GL28QD>.

* Assistant Professor, Political Science. matthew.nanes@slu.edu. 3750 Lindell Blvd, St. Louis MO 63108. Twitter: @MatthewJNanes

† Assistant Professor, Political Science. dhaim@fsu.edu. Twitter: @HaimDotan

In-person surveys on sensitive topics like racial prejudice, exposure to violence, and attitudes about abortion suffer from social desirability. The resulting non-response and falsification causes noise and bias. Existing strategies for reducing response bias include using enumerators with characteristics that minimize subjects’ discomfort (Adida et al. 2016) and asking questions indirectly (Nanes and Lau 2018). A popular strategy uses randomization to hide subjects’ responses at the individual level while allowing researchers to calculate sample-level estimates. Common procedures include list experiments (Kuklinski et al. 1997; Blair and Imai 2012; Nanes 2019), endorsement experiments (Lyall et al. 2013), and forced choice¹ (Warner 1965; Blair et al. 2015).

While these techniques have allowed scholars to make great strides in answering previously under-explored research questions, concerns remain. Randomization devices place high cognitive demands on subjects and enumerators, create opportunities for measurement error, and decrease measurement precision (Gelman 2014; Kramon and Weghorst 2019; Blair et al. 2019). They produce only sample-level estimates, though techniques exist to retrieve some of the “lost” information (Corstange 2009; Imai 2011; Blair and Imai 2012; Ahlquist 2018). Finally, implementing these techniques is often prohibitively expensive because they increase survey length and require extensive piloting.

As an alternative, many scholars attempt to reduce response bias on sensitive questions using self-administered surveys. Extensive evidence suggests that allowing respondents to record information on their own reduces non-response and preference falsification (Tourangeau and Yan 2007; Kim et al. 2010). Recent increases in the use of electronic devices for in-person surveys in the developing world have made these methods more common (Bush and Prather 2019). However, nearly all the evidence in support of self-administration comes either from online surveys or pencil-and-paper surveys conducted by highly literate populations in the developed world (Gnambs and Kaspar 2014).

We test whether allowing subjects to privately record their own responses reduces

¹Often called “randomized response”

social desirability bias in a context that matches the conditions faced by researchers studying sensitive topics relating to conflict and development. We embedded an experiment in an in-person survey conducted using electronic tablets in a rural, conflict-affected province in the Philippines. Experimentally-selected subsets of respondents answered the same question on their willingness to report insurgent activities using either 1) a direct verbal response, 2) self-enumeration, or 3) forced choice. We focus our analysis on forced choice because it allows us to keep the wording identical to the direct and self-enumerated question, sidestepping additional confounds that comparison with list or endorsement experiments would introduce. We view this technique as a good benchmark given studies find similar results for forced choice compared to these other approaches (Rosenfeld et al. 2016). While other studies found signs that subjects falsified about sensitive items despite using forced choice (Kraay and Murrell 2013), the promising application of this approach to violence-prone areas (Blair et al. 2015; Lyall et al. 2018) warrants further exploration.

Compared to the direct question approach, self-enumeration resulted in a significant increase in response rates but did not significantly affect the rate of reporting socially undesirable behaviors. Alarming, and consistent with Kramon and Weghorst (2019), we find that forced choice resulted in high rates of confusion among our respondents as well as highly inaccurate results. Thus, while self enumeration is not a panacea for all problems, it provides a low-cost method of increasing response rates without introducing the complications of randomized measurement devices.

Measuring Civilian Support for Insurgency in the Philippines

We seek to measure citizens' willingness to report insurgent activities to the authorities. Counterinsurgents rely on citizen-provided information to locate rebels, allowing government forces to use their power advantage in conventional combat (Berman et al. 2011). Citizens may be uncomfortable revealing their response to an enumerator. Supporting insurgents is illegal, and respondents may fear arrest or reprisals. They may also be embarrassed

to admit this socially-undesirable preference.

We survey 4,470 citizens in Sorsogon Province, Philippines, a hotbed of the communist insurgents New People’s Army (NPA). Sorsogon is divided into 541 barangays, the smallest administrative unit in the Philippines. Of these, 298 barangays were safe enough for us to conduct research. These barangays were not insurgent-free, but the enumerators used their affiliation with a well-known independent research organization and internationally-known university to gain permission to survey. We randomly selected 15 adults in each of the 298 barangays. Enumerators recorded answers in tablet computers.

One of Sorsogon’s strengths as a research site is the variety of citizens’ backgrounds in terms of education, exposure to globalization, and familiarity with technology. Our survey included many respondents who use a tablet or smartphone daily, and many who had never used such a device. In this sense, our sample provides a reasonably representative cross section typical of the wider developing world.

We conducted our experiment at the end of a survey containing questions for other research. For all previous questions, the enumerator read the question and answer choices out loud, the respondent said her answer choice, and the enumerator recorded the choice in the tablet (direct response). The experiment began by randomly assigning each survey respondent to answer both a placebo question and sensitive question in one of three ways: verbally, self-enumeration, or forced choice. For each subject, we asked both the placebo and sensitive questions using the same method.

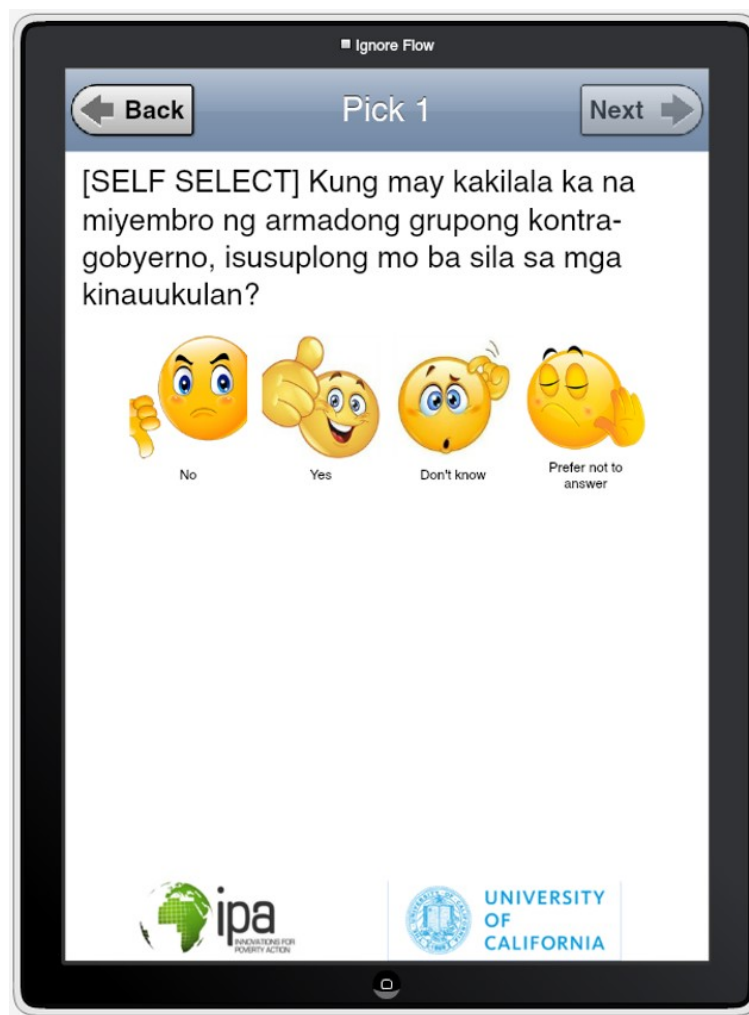
Placebo: “Did you complete high school?”

Sensitive: “If you knew about the activities of an anti-government group, would you report them to the authorities?”

Answer choices to both questions were “no,” “yes,” “don’t know,” and “refuse to answer.” For the approximately 1/3 assigned to self-enumerate, the enumerator stated: “You’ll enter your answer in the tablet yourself so that I cannot see, and then advance to the next

screen to keep your response private.” The enumerator read the question out loud, then turned the tablet around to face the subject. As the enumerator read each answer choice, he or she pointed to the corresponding options. Each option contained text and a symbol (Figure 1), allowing illiterate subjects to participate. The subject held the tablet so that the enumerator could not see the screen, tapped her response, and advanced the page. The enumerator then finalized the survey, making it impossible for anyone to access answers within the tablet. This procedure provided equal opportunity to operate the tablet even for respondents who declined to answer, mitigating concerns that the novelty of using the tablet might induce subjects to participate.

Figure 1: Self Enumeration Screenshot



Another 1/3 of subjects answered the questions using forced choice. Enumerators received extensive training implementing this technique, and answered subjects' questions until they understood what to do. The enumerator instructed:

“Next we’re going to try a technique to keep your responses private. I’ll give you a coin to flip and you will flip it and notice which side it lands on. If it lands on heads, please answer the next question honestly. If it lands on tails, please answer “yes,” regardless of your true answer. This way, I won’t know whether you answered “yes” because that’s your true answer or because of how the coin landed, so your answer stays private.”

The final one-third of subjects responded exactly as they did in the rest of the survey, by telling the enumerator their answer verbally and the enumerator recording it in the tablet.

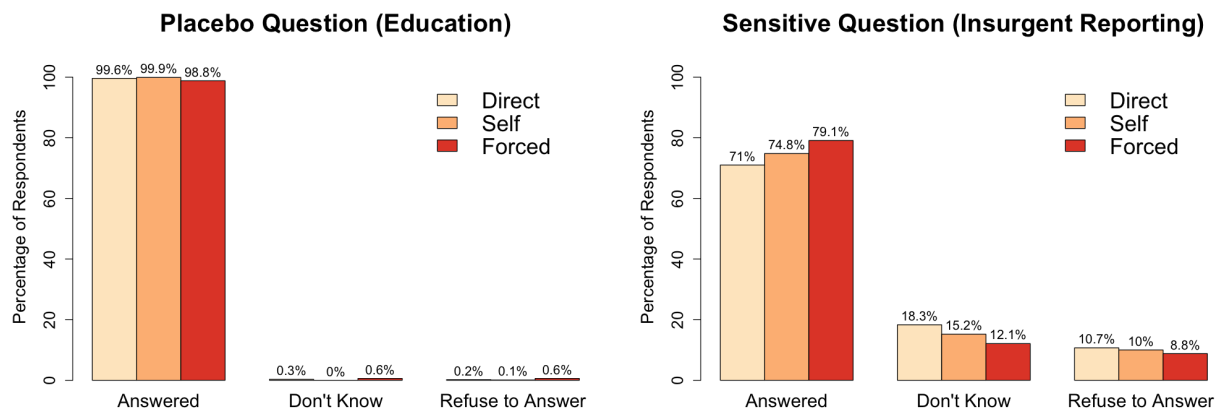
While our main interest is in enumerator-induced social desirability bias, the presence of bystanders presents another possible source of sensitivity. The respondent’s spouse or relative may wish to oversee the enumeration, or enumerators’ presence may attract neighbors’ attention. Bystanders can affect people’s willingness to answer, and the answer they select. The best solution is to conduct surveys in a private location. We instructed our enumerators to seek out a private room in the respondent’s house and avoid locations with onlookers. However, total privacy is often impossible, thus it is worth studying the effects of bystanders explicitly. Despite our enumerators’ efforts, 24.2% of our surveys experienced onlookers.² While not part of our randomized design, we explore the effects of onlookers in a multivariate regression framework in the supporting information and report some of the results below.

²We did not anticipate that so many interviews would occur in the presence of bystanders. In the appendix, we discuss the ethical implications of conducting sensitive interviews with bystanders present. Respondents decided whether to answer questions with full knowledge of bystanders’ presence, mitigating some ethical concerns. We also believe that the majority of bystanders were family members, reducing the concern that bystanders might reveal information to the authorities or insurgents. Still, a clear lesson is that we should have placed greater emphasis on privacy when training our enumerators.

Results

We begin by analyzing *response rates* across the three experimental groups in Figure 2. Clearly, reporting insurgents is much more sensitive than completing high school. 25.3% of respondents selected either “Don’t Know” or “Refuse to Answer” for reporting an armed group, while only 0.53% selected these options about high school graduation. Subjects were more likely to answer the question about reporting insurgents using self enumeration (74.8%) or forced choice (79.1%) than direct questioning (71.0%). Regression models (see Supporting Material) show that for the sensitive question, these differences in response rates are statistically significant.

Figure 2: Response Rates by Experimental Group



Although forced choice significantly decreases non-response rates, we cannot determine the distribution of non-responses relative to the results of subjects’ coin flips. While we know that the probability of the coin landing on heads is 50%, the probability of a respondent answering the question might be conditional on the result of the coin flip. If response rates increase only among those who are “forced” to answer yes, the method introduces bias by causing us to subtract out the wrong proportion of affirmative responses. Neither self- nor direct-enumeration suffer from this limitation.

The multivariate models in the Supporting Materials also help us disaggregate the mechanisms driving the higher response rate in the self-enumeration group. Self-enumeration

may reduce social desirability bias by shielding answers from the enumerator, or by reducing the likelihood of an onlooker overhearing. To distinguish between these possibilities, we include an interaction between the presence of a bystander and the use of self enumeration. While bystander presence is associated with increased non-response, the effects of self enumeration on response rates does not vary noticeably depending on bystander presence, leaving the enumerator’s presence as the most likely cause of non-response under direct enumeration.

While we propose that self enumeration improves response rates by enhancing privacy, an alternative possibility is that subjects exposed to the novelty of using the tablet were more willing to exert effort on the survey, while those responding verbally responded “don’t know” or “prefer not to answer” because they did not care to engage with the survey. If improved engagement increased response rates, it should have done so on both sensitive and non-sensitive questions. However, we observe a distinguishable difference in response rates between Self and Direct enumeration only on the question about insurgency, not on the placebo question. While we acknowledge that we cannot completely rule out this possible alternative explanation, it also leads to an improvement in response rates without raising concerns over response accuracy.

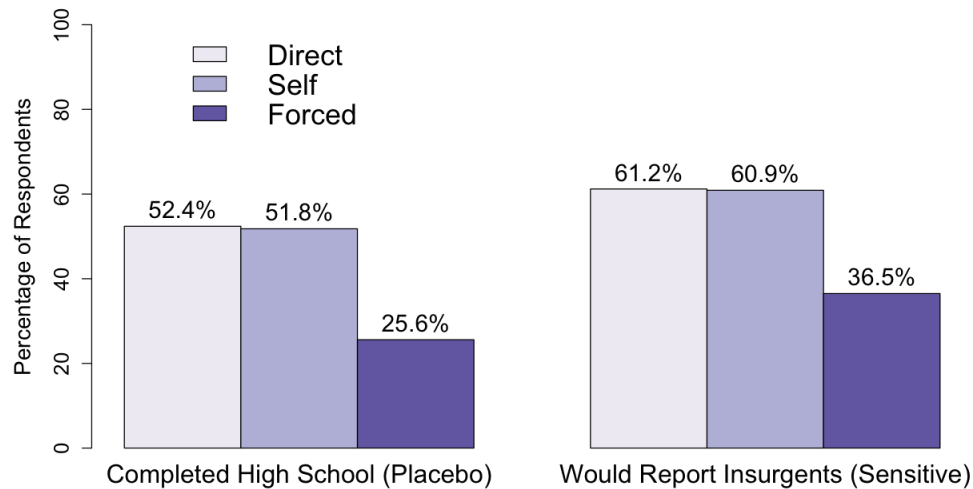
Finally, if survey method reduces non-response by addressing social desirability bias, then it should *not* affect non-response on a question that is not socially sensitive. Yet, we observe about a one percentage point *decrease* in response rates using forced choice compared to both direct and self enumeration, a small but statistically-significant difference, raising concerns that the forced choice method itself may suppress responses.

Falsification

Overall, 51.8% of respondents say they would report rebels to the authorities. We expect social desirability bias to make respondents more likely to say that they would report insurgents to the authorities, as supporting insurgents is illegal and socially undesirable.

Figure 3 shows that direct- and self-enumeration yield similar responses on both the placebo question and the sensitive question about reporting rebels. On the other hand, the forced choice method decreased affirmative responses by about 50% for the placebo and 41% for the sensitive question.

Figure 3: Affirmative Answers by Experimental Group



Direct and Self are observed frequencies, while Forced depicts an estimated value. Statistically-significant differences are noted in-text, and we present regression models in the Supporting Information.

We present multivariate models in the Supporting Materials. Of note, we find a negative relationship between the presence of bystanders and willingness to report insurgents, hinting at a potential problem with our design. We expected that the enumerator’s presence should *inflate* affirmative responses, as respondents likely view enumerators as authority figures. On the other hand, social desirability bias from onlookers *deflates* affirmative responses – likely because respondents worry that an onlooker might report back to the NPA. To the extent that self enumeration reduces social desirability bias from both sources, the effects might cancel each other out.

The dramatic decrease in affirmative responses to the sensitive question when using

forced choice is in line with our expectations about the direction of bias. Were this the only question on the survey asked using forced choice, we might reasonably interpret the reduction as an improvement in accuracy. Indeed, a major problem with randomization devices is that they may lead researchers to misinterpret whether results are due to improvements in measurement or the introduction of a new source of bias. In our case, however, the corresponding change in responses to the placebo suggests that these responses may be biased, as answers to the non-sensitive question should not have changed when we improved privacy. According to the Philippine Statistics Authority, nation-wide 38.7% of the population age 6 and up has completed high school.³ Our survey is limited to adults age 18 and up, so our estimate should be higher. In two prior surveys in Sorsogon Province, we find that 53.9% and 51.95% of respondents claimed to have graduated high school. Consistent with these figures, 52.4% and 51.8% of respondents who answered the question using direct- and self-enumeration, respectively, claim to have completed high school. On the other hand, forced choice yields an unlikely estimate of only 25.6%.

We speculate that the substantially lower estimates using forced choice for both sensitive and non-sensitive questions result from non-compliance. We suspect that a large number of respondents whose coin toss assigned them to answer “yes” actually answered truthfully out of confusion. Thus, by subtracting out the 50 percent of respondents who were “forced” to answer “yes,” we subtract out an unknown quantity of true “yeses.” An alternative explanation might be that the use of a complicated method to ask a non-sensitive question confused respondents, leading to non-strategic falsification, whereas those answering the sensitive question using random-response understood the need for anonymity and therefore answered accurately. While we cannot dismiss this possibility entirely, we note that the *same respondents* answered both the sensitive and non-sensitive questions with forced choice, making confusion on one question but not the other unlikely.

The negligible difference in affirmative responses between direct and self enumeration

³<https://psa.gov.ph/content/education-women-and-men>

means that non-responders, of whom there were more under direct enumeration, do not differ systematically from responders and therefore did not introduce measurable bias. Yet, the lower non-response rates self enumeration elicits are still desirable for researchers, as they yield a larger usable sample and therefore allow for more precisely-estimated tests using the data. Furthermore, our failure to find bias from non-response in this scenario does not rule out its possibility on other questions or in other samples.

Discussion

Our results suggest that while randomized survey devices may address non-response from social desirability, they do so at the cost of inaccuracy. In our case, forced choice produced estimates inconsistent with existing evidence even on a non-sensitive topic. Without the benefit of comparison against a placebo with a known quantity, we might have mistaken the lower rate of affirmative answers to the sensitive question as an improvement in accuracy. Our findings serve as a stark reminder to researchers to employ randomized survey devices with caution. On the other hand, we find evidence that a much simpler method, allowing respondents to record their answers privately using a tablet computer, provides noticeable improvements in response rates without a loss of accuracy. Where it is not plausible to use tablets, researchers might ask subjects to record their answers on paper and then deposit them in a locked box for a similar effect.

In the Supporting Materials, we show that forced choice is more susceptible to enumerator effects than direct- or self-enumeration, and that it takes longer to enumerate than the other methods. While we fail to find conclusive evidence that self enumeration reduces enumerator-induced social desirability bias, we show that it yields responses that are at least as accurate as verbal response and more accurate than forced choice, while also eliciting fewer non-responses to a sensitive question than verbal response, making it an appropriate method for asking sensitive questions where an enumerator’s presence is required.

Two additional results stand out. First, bystanders were associated with increased

non-response and decreased likelihood of a respondent saying she would report insurgents to the authorities. These findings are a stark reminder that even the best research designs and enumeration procedures must take basic precautions like conducting surveys in private. Second, a large portion of respondents say they would *not* report the activities of an anti-government group to the authorities. Given the importance of citizen-provided information in combating insurgents, these figures are troubling for governments and international organizations combating insurgency. Future research might explore whether this lack of willingness to report is due to low government legitimacy, difficulty of reporting information, concerns about reprisals, or some other cause. The answers to this question determine the steps that the government, as well as outside agencies involved in countering violent extremism, should take to improve security and service delivery to citizens.

In closing, we note a disconnect in political science between developing measurement tools and developing estimation techniques for those tools. Scholars have made substantial improvements in analyzing data from experiment-based survey methods. However, these tools do not correct for inaccurate underlying data. We do not suggest that researchers should avoid randomized measurement devices altogether. However, we believe it is important to acknowledge that they introduce a new set of problems, for example trading social desirability bias for non-strategic response error. Where researchers are concerned about non-response due to social desirability bias, self-enumeration can improve response rates while providing responses that are as accurate and efficient as direct questioning.

References

- Adida, C. L., Ferree, K. E., Posner, D. N. and Robinson, A. L. (2016), ‘Whos asking? interviewer coethnicity effects in african survey data’, *Comparative Political Studies* **49**(12), 1630–1660.
- Ahlquist, J. S. (2018), ‘List experiment design, non-strategic respondent error, and item count technique estimators’, *Political Analysis* **26**(1), 34–53.
- Berman, E., Shapiro, J. N. and Felter, J. H. (2011), ‘Can hearts and minds be bought? the economics of counterinsurgency in iraq’, *Journal of Political Economy* **119**(4), 766–819.
- Blair, G., Chou, W. and Imai, K. (2019), ‘List experiments with measurement error’, *Political Analysis* **27**(4), 455–480.
- Blair, G. and Imai, K. (2012), ‘Statistical analysis of list experiments’, *Political Analysis* **20**(1), 47–77.
- Blair, G., Imai, K. and Zhou, Y.-Y. (2015), ‘Design and analysis of the randomized response technique’, *Journal of the American Statistical Association* **110**(511), 1304–1319.
- Bush, S. S. and Prather, L. (2019), ‘Do electronic devices in face-to-face interviews change survey behavior? evidence from a developing country’, *Research & Politics* **6**(2), 1–7.
- Corstange, D. (2009), ‘Sensitive questions, truthful answers? modeling the list experiment with listit’, *Political Analysis* **17**(1), 45–63.
- Gelman, A. (2014), Thinking of doing a list experiment? heres a list of reasons why you should think again, Technical report. <https://statmodeling.stat.columbia.edu/2014/04/23/thinking-list-experiment-heres-list-reasons-think/>.
- Gnambs, T. and Kaspar, K. (2014), ‘Disclosure of sensitive behaviors across self-administered survey modes: a meta-analysis’, *Behavior Research Methods* **47**(4), 1237–1259.

- Imai, K. (2011), ‘Multivariate regression analysis for the item count technique’, *Journal of the American Statistical Association* **106**(494), 407–416.
- Kim, J., Kang, J.-h., Kim, S., Smith, T. W., Son, J. and Berkthold, J. (2010), ‘Comparison between self-administered questionnaire and computer-assisted self-interview for supplemental survey nonresponse’, *Field Methods* **22**(1), 57–69.
- Kraay, A. and Murrell, P. (2013), *Misunderestimating corruption*, The World Bank.
- Kramon, E. and Weghorst, K. (2019), ‘(mis) measuring sensitive attitudes with the list experimentsolutions to list experiment breakdown in kenya’, *Public Opinion Quarterly* .
- Kuklinski, J. H., Cobb, M. D. and Gilens, M. (1997), ‘Racial attitudes and the” new south”’, *The Journal of Politics* **59**(2), 323–349.
- Lyall, J., Blair, G. and Imai, K. (2013), ‘Explaining support for combatants during wartime: A survey experiment in afghanistan’, *American Political Science Review* **107**(04), 679–705.
- Lyall, J., Zhou, Y. and Imai, K. (2018), ‘Reducing insurgent support among at-risk populations: Experimental evidence from cash transfers and livelihood training in afghanistan’, *SSRN Working Paper* **3026531**.
- Nanes, M. and Haim, D. (2020), ‘Replication data for: Self-administered field surveys on sensitive topics’, *Harvard Dataverse* . <https://doi.org/10.7910/DVN/GL28QD>.
- Nanes, M. J. (2019), ‘Police integration and support for anti-government violence in divided societies: Evidence from iraq’, *Journal of Peace Research* .
- Nanes, M. and Lau, B. (2018), ‘Surveys and countering violent extremism: A practitioner guide’.
- Rosenfeld, B., Imai, K. and Shapiro, J. N. (2016), ‘An empirical validation study of popular survey methodologies for sensitive questions’, *American Journal of Political Science* **60**(3), 783–802.

Tourangeau, R. and Yan, T. (2007), ‘Sensitive questions in surveys.’, *Psychological bulletin* **133**(5), 859.

Warner, S. L. (1965), ‘Randomized response: A survey technique for eliminating evasive answer bias’, *Journal of the American Statistical Association* **60**(309), 63–69.