Journal of Experimental Political Science (2023), **10**, 148–153
doi:10.1017/XPS.2022.8

SHORT REPORT

Fraud in Online Surveys: Evidence from a Nonprobability, Subpopulation Sample

Andrew M. Bell[1,2]* and Thomas Gift[3]

[1]Indiana University-Bloomington, Bloomington, IN, USA, [2]U.S. Army War College, Carlisle, PA, USA and [3]University College London, London, UK
*Corresponding author. Email: bellam@iu.edu

Abstract
We hired a well-known market research firm whose surveys have been published in leading political science journals, including *JEPS*. Based on a set of rigorous "screeners," we detected what appears to be exceedingly high rates of identity falsification: over 81 percent of respondents seemed to misrepresent their credentials to gain access to the survey and earn compensation. Similarly high rates of presumptive character falsification were present in panels from multiple sub-vendors procured by the firm. Moreover, we found additional, serious irregularities embedded in the data, including evidence of respondents using deliberate strategies to detect and circumvent one of our screeners, as well as pervasive, observable patterns reflecting that the survey had been taken repeatedly by a respondent or collection of respondents. This evidence offers reasons to be concerned about the quality of online nonprobability, subpopulation samples, and calls for further, systematic research.

Growing evidence points to problems with "character misrepresentation" in digital surveys (Ahler et al. 2021; Chandler and Paolacci 2017; Hydock 2018; Ryan 2020; Wessling et al. 2017). We present concerning results from a nonprobability online survey of a specific subpopulation fielded through a well-known, commercial firm. A set of rigorous "screeners" revealed extremely high rates of presumptive fraud: More than 81 percent of respondents appeared to misrepresent themselves as current or former US Army members – our subpopulation of interest – to complete the survey and earn compensation. Presumed falsification rates were similar across multiple established sub-vendors, indicating that the problems were not idiosyncratic to a particular panel. Data also indicate the use of deliberate tactics to circumvent one of our screeners and repeated participation from a respondent or group of respondents, further raising suspicions about the data.

This article has earned badges for transparent research practices: Open Data and Open Materials. For details see the Data Availability Statement.

These irregularities point to the potential for significant identity misrepresentation rates in online nonprobability, subpopulation surveys – rates that are orders of magnitude greater than those typically reported in standard online surveys (Callegaro et al. 2014; Cornesse et al. 2020; Kennedy et al. 2020a; Mullinix et al. 2016). Although online survey firms can vary markedly in their regulation of quality control – and all should be assessed for data problems (Kennedy et al. 2016) – risks for exceedingly high levels of fraud could be heightened under the conditions present in our study. Our findings call for further, systematic research into the validity of nonprobability online surveys, particularly those that sample specific subpopulations. They also underscore the imperative for researchers to develop clear tools and strategies (prior to statistical analyses and within preregistration plans) to ensure data integrity based on expert knowledge of research subjects.

Survey details

We contracted a nationally recognized market research firm,[1] whose samples have formed the basis of numerous, widely-cited political science studies, including articles published in *Journal of Experimental Political Science*, *American Political Science Review*, and *Journal of Politics*. The firm, which used multiple sub-vendors, fielded the survey over two separate rounds in April–May 2021.[2]

Screening process

We employed two screeners to confirm the authenticity of respondents with self-identified Army experience. First, we asked a "knowledge" question about the practice of saluting, one of the most essential elements of military protocol.[3] The question required knowing both the Army's rank hierarchy and that enlisted soldiers salute first. Multiple former Army officers consulted for this study validated the screen, with one stating: "Anyone who is answering that question incorrectly is either not reading the question or has not served in the military, let alone the Army."[4] Second, respondents reported specific information on their Army background including: highest rank achieved, source of officer commission, deployment years and locations, and unit type.[5] We coded "non-viable" responses as those that provided information contravening federal law or Army personnel policy. We coded a small number of responses "highly improbable" based on being contrary to Army personnel practices or historical evidence, but which could be theoretically plausible.[6]

[1]We withhold the survey firm's name and its associated sub-vendors due to liability reasons.

[2]Round 2 was launched after Round 1 was halted due to concerns about data quality. For detailed information on the survey, see Appendix (Supplementary Material).

[3]For text, see Appendix. Army rank structure and saluting protocols are so central to Army service that they are among the first subjects recruits learn in basic training (Army, 2019).

[4]Email from former US Army major, April 20, 2021.

[5]Information such as military rank is so defining of an Army member's career and identity – akin to a civilian's job – that respondents should almost never report these details incorrectly (Bell & Terry, 2021).

[6]For examples of "non-viable" and "highly improbable" responses, see Appendix.

https://doi.org/10.1017/XPS.2022.8 Published online by Cambridge University Press

Key findings

- **Total invalid: 81.8 percent**
  - 43.3 percent of total respondents failed the Army knowledge question.
  - 35.5 percent of respondents passed the knowledge screen but gave answers about Army service that were non-viable under federal law or military administrative rules.
  - 3.0 percent of respondents reported information about an Army background and career that was highly improbable.
- **Total valid: 18.2 percent**

Notably, the number of invalid responses remained consistent across multiple vendors, each operating in a different survey round (83.7 in Round 1 and 78.9 in Round 2), suggesting that the problem was not isolated to a specific panel (Table 1).

Additional irregularities

1. Respondents appeared to employ deliberate tactics to circumvent the knowledge screener.

Evidence. Beginning on April 17 in Round 1, the proportion of respondents who passed the knowledge question with non-viable demographic and career answers jumped sharply from 15.4 to 65.9 percent (Table 2). Possible explanations for this unexpected "break" in the data might include collusion (via the sharing of information on an email listserv or crowdsourced website), the retaking of surveys on devices with unique IP addresses, or the use of survey "bots."

2. Respondents appeared to use deliberate tactics to detect the knowledge question as a screener.

Evidence. In Round 1 – where failure to pass the knowledge question terminated the survey immediately – 49.8 percent passed the knowledge question but provided non-viable career responses. Conversely, in Round 2 – where we modified the setup so that failure resulted in termination later in the survey, thereby "hiding" the screener – the number of non-viable fell to 13.7 percent (Table 1). Conversely, the percentage of respondents who failed the screen rose from 30.6 percent (Round 1) to 62.7 percent (Round 2). This suggests that, absent a "hidden" screener, later respondents in Round 1 may have gleaned information from earlier respondents or survey attempts to identify the screener and defeat it.[7]

3. A respondent, or collection of respondents, appeared to take the survey numerous times.

Evidence. Categorizing conservatively, we identified – at a minimum – 73 suspicious instances of repeated (and unusual) responses regarding Army background and deployment experience.[8] The sequential clustering of these responses – six

[7]Duplicate IP addresses accounted for 11 of the total responses. These responses are included in the analysis.

[8]For examples, see Appendix.

https://doi.org/10.1017/XPS.2022.8 Published online by Cambridge University Press

Table 1
Summary of response categories

Category	Round 1 Count	Round 1 %	Round 2 Count	Round 2 %	Total Count	Total %
Valid	40	16.3%	34	21.1%	74	18.2%
Failed screen	75	30.6%	101	62.7%	176	43.3%
Passed screen, non-viable	122	49.8%	22	13.7%	144	35.5%
Passed screen, improbable	8	3.3%	4	2.5%	12	3.0%
Subtotal	245	100.0%	161	100.0%	406	100.0%

Table 2
Irregularity 1: Increase in "non-viable" responses (Round 1)

	Pre-break		Post-break	
Category	Count	%	Count	%
Valid	28	35.9%	12	7.2%
Failed screen	37	47.4%	38	22.8%
Passed screen, non-viable	12	15.4%	110	65.9%
Passed screen, improbable	1	1.3%	7	4.2%
Total	78	100.0%	167	100.0%

distinct waves of repeating answers across the 3 days of responses – suggests that such repetition was not coincidental. We also observed obvious repetitive patterns of survey takers seeming to misrepresent personal demographic information.

Conclusion

We see this analysis as an opportunity for learning. Despite taking precautions to screen out invalid respondents, we found high rates of presumptive fraud. This reinforces that researchers should be especially cautious when employing online surveys using nonprobability samples of specific subpopulations. Given that only about 7 percent of the US population is military or ex-military (Vespa 2020), our results are consistent with incentives for fraud increasing as the size of the subpopulation qualifying to participate in surveys decreases (Chandler and Paolacci 2017). Combined with other techniques, employing a diversity of screeners predicated on expert understanding of research subjects – including factors like demographics and content knowledge – can improve the odds of detecting falsified responses. Future research should systematically assess the quality of nonprobability surveys (Hauser and Schwarz 2016; Lopez and Hillygus 2018; Kennedy et al. 2020b; Thomas and Clifford 2017). By implementing rigorous screeners on diverse

https://doi.org/10.1017/XPS.2022.8 Published online by Cambridge University Press

populations, replicated across many firms and sub-vendors, this could illuminate whether our results are endemic to nonprobability surveys that sample specific sub-populations and what the broader implications are for internal and external validity.

Supplementary Material. To view supplementary material for this article, please visit https://doi.org/10.1017/XPS.2022.8

Data Availability. The data, code, and additional materials required 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/Y1FEOX"

Acknowledgments. The authors thank Christopher DeSante, Timothy Ryan and Steven Webster, two anonymous U.S. Army officers, as well as anonymous reviewers for their helpful comments and suggestions.

Conflicts of Interest. The authors declare no conflicts of interest.

Ethics Statement. This survey was approved by the Indiana University-Bloomington IRB (Protocol #: 1910663858). The research adheres to APSA's Principles and Guidance for Human Subjects Research. See Supplemental Appendix for more information.

References

Ahler, D. J., C. E. Roush and G. Sood. 2021. The micro-task market for lemons: Data quality on amazon's mechanical turk. *Political Science Research and Methods*, 1–20.

Army. 2014. *Department of the Army Pamphlet 600–3, Commissioned Offcer Professional Development and Career Management*. Washington, DC: Headquarters, Department of the Army.

Army. 2019. *U.S. Army Training and Doctrine Command Pamphlet 600-4, The Soldier's Blue Book, The Guide for Initial Entry Training Soldiers*. Fort Eustis, VA: Department of the Army.

Bell, A. M. and F. Terry. 2021. Combatant Rank and Socialization to Norms of Re- straint: Examining the Australian and Philippine Armies. *International Interactions 47*(5): 825–854.

Callegaro, M., A. Villar, D. Yeager and J. A. Krosnick. 2014. A Critical Review of Studies Investigating the Quality of Data Obtained with Online Panels based on Probability and Nonprobability Samples. In *Online Panel Research: A Data Quality Perspective*, eds. Baker, R., Bethlehem, J., Goritz, A. S., Kros-nick, J. A., Callegaro, M. and Lavrakas, P. J. West Sussex, UK: John Wiley & Sons.

Chandler, J. J. and G. Paolacci. 2017. Lie for a Dime: When Most Prescreening Responses are Honest but most Study Participants are Impostors. *Social Psychological and Personality Science 8*(5): 500–508.

Cornesse, C., A. G. Blom, D. Dutwin, J. A. Krosnick, E. D. De Leeuw, S. Legleye, J. Pasek and D. Pennay. 2020. A Review of Conceptual Approaches and Empirical Evidence on Probability and Nonprobability Sample Survey Research. *Journal of Survey Statistics and Methodology 8*(1): 4–36.

Hauser, D. J. and N. Schwarz. 2016. Attentive Turkers: MTurk Participants Perform Better on Online Attention Checks than do Subject Pool Participants. *Behavior Research Methods 48*(1): 400–407.

Hydock, C. 2018. Assessing and Overcoming Participant Dishonesty in Online Data Collection. *Behavior Research Methods 50*: 1563–1567.

Kennedy, C., A. Mercer, S. Keeter , N. Hatley, K. McGeeney and A.-j. Gimenez. 2016. Evaluating Online Nonprobability Surveys. *Pew Research* May 2, 2016.

Kennedy, C., N. Hatley, A. Lau, A. Mercer, S. Keeter, J. Ferno and D. Asare-Marf. 2020a. Assessing the Risks to Online Polls From Bogus Respondents. *Pew* Feb. 18, 2020.

Kennedy, R., S. Clifford, T. Burleigh, P. D. Waggoner, R. Jewell and N. J. G. Winter. 2020b. The Shape of and Solutions to the mturk Quality Crisis. *Political Science Research and Methods 8*. 614–629.

Lopez, J. and D. S. Hillygus. 2018. Why So Serious?: Survey Trolls and Misinformation. *SSRN Electronic Journal*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3131087

Mullinix, K. J., T. J. Leeper, J. N. Druckman and J. Freese. 2016. The Generalizability of Survey Experiments. *Journal of Experimental Political Science 2*(2): 109–138.

Ryan, T. J. 2020. Fraudulent Responses on Amazon Mechanical Turk: A Fresh Cautionary Tale. Retrieved from https://timryan.web.unc.edu/2020/12/22/fraudulent-responses-on-amazon-mechanicalturk-a-fresh-cautionary-tale/

Thomas, K. A. and S. Clifford. 2017. Validity and Mechanical Turk: An Assessment of Exclusion Methods and Interactive Experiments. *Computers in Human Behavior 77*(1): 184–197.

Vespa, J. E. 2020. Those Who Served: America's Veterans From World War II to the War on Terror. *American Community Survey Report, U.S. Census Bureau* June 2020.

Wessling, K. S., J. Huber and O. Netzer. 2017. Mturk Character Misrepresentation: Assessment and Solutions. *Journal of Consumer Research 44*(1): 211–230.

Cite this article: Bell AM and Gift T (2023). Fraud in Online Surveys: Evidence from a Nonprobability, Subpopulation Sample. *Journal of Experimental Political Science* **10**, 148–153. https://doi.org/10.1017/XPS.2022.8