methods, data, analyses | Vol. 13(1), 2019, pp. 3-6

Editorial: Social Desirability Bias in Surveys – Collecting and Analyzing Sensitive Data

Ben Jann[1], *Ivar Krumpal*[2] *& Felix Wolter*[3]

[1] University of Bern
[2] University of Leipzig
[3] Johannes Gutenberg University Mainz

Studying social phenomena and social problems often involves measuring and analyzing behaviors or attitudes that are sensitive in several ways. Topics such as delinquency, substance abuse, sexual issues, xenophobia or homophobia may oblige survey respondents to self-report information about very private issues or to report that they have acted against social or legal norms. Hence, survey participants could fear negative consequences of violating social desirability (SD) norms or of a disclosure of their private information to third parties (Tourangeau & Yan, 2007).

As cumulative empirical research has shown, this prompts respondents to engage in self-protective behavior when answering sensitive survey questions, namely by providing untruthful and biased answers (be it unconsciously or deliberately) or by refusing to answer at all (Krumpal, 2013; Lensvelt-Mulders, 2008; Wolter, 2012). This systematic misreporting or nonresponse leads to biased estimates and poor data quality. Statistical associations could be biased as well if the degree of misreporting varies systematically across subgroups or is related to other variables.

At the same time, research about sensitive topics and norm-violations is of particular interest for the social sciences and public discussions likewise: Public authorities, for instance, are interested in being informed about the prevalence of tax evasion, corruption, or illicit work. Media and political parties seek for accurate election forecasts. Researchers may want to study levels and determinants of deviant behaviors, political extremism, or health problems.

The demand for valid measurements of sensitive issues on the one hand and the well-confirmed difficulties due to SD bias on the other has occupied survey methodologists since the very beginning of modern survey research (Benson,

1941; Hyman, 1944). There are two main lines of research. The first one consists in theorizing about, identifying, and quantifying response biases and, if possible, in providing means for controlling such biases ex post, that is, after the data has been collected. One approach for instance concerns measuring and adjusting for socially desirable responding by using psychometric SD scales. The theoretical part of this research agenda seeks for explanations and clarifications of the mechanisms causing systematic misreporting or nonresponse. The second line of research aims at developing data collection techniques that alleviate or, at best, entirely avoid response biases. More conventional approaches in this regard encompass choosing a well-tailored (e.g., self-administered) survey mode or a mixed-mode design, using wording or filtering techniques, and reducing interviewer effects. Strategies that are more complex employ special questioning techniques that mostly pursue the goal of reducing misreporting by increasing the level of anonymity of the respondents' answers, for example via adding random statistical noise to the data. Randomized-response (RRT; Warner, 1965) and item count techniques (ICT; Droitcour et al., 1991) are probably the most prominent techniques in this regard.

Despite the long-standing research tradition in this field, one cannot allege that all problems have been solved. This holds for both theoretical and methodical questions on "best practices". For example, there is an ongoing theoretical discussion about the psychological mechanisms causing respondents to misreport on their true status (e.g., Holtgraves, 2004). Empirical findings regarding the performance of special questioning techniques such as RRT and ICT are mixed and often inconclusive (e.g., Holbrook & Krosnick, 2010). Hence, the objective of this special issue is to contribute to the ongoing debate about theoretical issues as well as about establishing best practices, survey designs, or measurement instruments for surveying sensitive topics.

The article by *Henrik Andersen* and *Jochen Mayerl* addresses the question whether socially desirable responding is more a deliberate, reflected editing of answers, or an automatic process occurring spontaneously. The authors find empirical evidence for both mechanisms depending on whether respondents report about positively connoted traits or about negatively connoted ones.

The paper by *Axel Franzen* and *Sebastian Mader* investigates whether "phantom questions", that is, questions on fictitious, non-existent issues, represent an opportunity to measure respondents' affinity for SD bias. The authors empirically compare classic SD scales (short versions of the Crowne-Marlowe SD scale) and phantom questions with respect to their internal and external consistency and validity.

The study by *Manfred Antoni*, *Daniel Bela*, and *Basha Vicari* deals with SD bias in reported earnings. Linking survey data to administrative validation data on an individual level, the authors investigate the degree of over- and underreport-

ing depending on earnings levels, other individual characteristics, and interviewer effects.

Paula Fomby and *Narayan Sastry* discuss the use of interactive voice response technology (IVR) for collecting sensitive data among adolescents. The authors review questionnaire design, fieldwork protocols, data quality and completeness, and respondent burden of the IVR procedure employed in the Panel Study of Income Dynamics 2014 Child Development Supplement.

The paper by *Alessandra Gaia* and *Tarek Al Baghal* presents a new version of the ICT, namely the longitudinal ICT (L-ICT). While ICT is implemented in cross-sectional surveys with a random split into different sub-samples, L-ICT administers the long- and short-lists (one including the sensitive item, the other not) to the same respondents in different waves of a panel survey. The authors discuss general properties, pros, and cons of L-ICT and present empirical results from a first implementation in the *Understanding Society* Innovation Panel.

The article by *Anke Erdmann* presents empirical evidence on the performance of the triangular model (TM) for gathering sensitive survey data as compared to conventional direct questioning. The sensitive questions pertain to issues about mental stress among students. The author also addresses whether the TM has different effects for certain subgroups of respondents, such as for those scoring high on SD or depressiveness scales.

Finally, the study by *Felix Wolter* seizes a suggestion by Grant, Moon, and Gleason (2014) and introduces the person count technique (PCT), a new variant of ICT. PCT is empirically tested in an experimental survey against conventional direct questioning with respect to nonresponse and misreporting on attitude questions about asylum seekers.

Overall, we are confident that this special issue of mda provides various important contributions to both theoretical and practical challenges in the field of research on sensitive questions. We would like to thank all the authors for their valuable contributions and their patience during the review process. Our thanks also go to the editorial team of mda for their support, and the reviewers for their careful reading and commenting of the manuscripts.

References

Benson, L. E. (1941). Studies in Secret-Ballot Technique. *Public Opinion Quarterly, 5*(1), 79–82.

Droitcour, J., Caspar, R. A., Hubbard, M. L., Parsley, T. L., Visscher, W., & Ezzati, T. M. (1991). The Item Count Technique as a Method of Indirect Questioning: A Review of its Development and a Case Study Application. In P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz., & S. Sudman (Eds.), *Measurement Errors in Surveys* (pp. 185–210). New York: Wiley.

Grant, T., Moon, R., & Gleason, S. A. (2014). *Asking Many, Many Sensitive Questions: A Person-Count Method for Social Desirability Bias*: Unpublished Manuscript.

Holbrook, A. L., & Krosnick, J. A. (2010). Measuring Voter Turnout by Using the Randomized Response Technique. Evidence Calling into Question the Method's Validity. *Public Opinion Quarterly, 74*(2), 328–343.

Holtgraves, T. (2004). Social Desirability and Self-Reports: Testing Models of Socially Desirable Responding. *Personality and Social Psychology Bulletin, 30*(2), 161–172.

Hyman, H. (1944). Do They Tell the Truth? *Public Opinion Quarterly, 8*(4), 557–559.

Krumpal, I. (2013). Determinants of Social Desirability Bias in Sensitive Surveys: A Literature Review. *Quality & Quantity, 47*(4), 2025–2047.

Lensvelt-Mulders, G. J. L. M. (2008). Surveying Sensitive Topics. In E. D. de Leeuw, J. J. Hox, & D. A. Dillman (Eds.), *International Handbook of Survey Methodology* (pp. 461–478). New York: Lawrence Erlbaum.

Tourangeau, R., & Yan, T. (2007). Sensitive Questions in Surveys. *Psychological Bulletin, 133*(5), 859–883.

Warner, S. L. (1965). Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias. *Journal of the American Statistical Association, 60*(309), 63–69.

Wolter, F. (2012). *Heikle Fragen in Interviews. Eine Validierung der Randomized Response-Technik*. Wiesbaden: Springer VS.