2020

Comments on “How Errors Cumulate: Two Examples” by Roger Tourangeau

Kristen M. Olson

Comments on "How Errors Cumulate: Two Examples" by Roger Tourangeau

Kristen Olson

Department of Sociology, University of Nebraska-Lincoln, 703 Oldfather Hall, Lincoln, NE 68588-0324, USA. *email* kolson5@unl.edu.

Abstract

This paper provides a discussion of the Tourangeau (2019) Morris Hansen Lecture paper. I address issues related to compounding errors in web surveys and the relationship between nonresponse and measurement errors. I provide a potential model for understanding when error sources in nonprobability web surveys may compound or counteract one other. I also provide three conceptual models that help explicate the joint relationship between nonresponse and measurement errors.

Keywords: Measurement error, Mode effects, Morris Hansen Lecture, Nonresponse.

Published in *Journal of Survey Statistics and Methodology* 8 (2020), pp 442–452.

doi: 10.1093/jssam/smz025

Published 5 July 2019.

The author thanks the Morris Hansen Lecture Committee and Roger Tourangeau for the invitation to read and discuss this paper. Initial thinking about the joint roles of nonresponse and measurement errors was supported by the National Science Foundation (NSF SES-0620228). Work on mode predictors was supported in part by the University of Nebraska-Lincoln Office of Research and Economic Development and partially funded by Cooperative Agreement USDA-NASS 58-AEU-5-0023, supported by the National Science Foundation National Center for Science and Engineering Statistics.

1. Introduction

Tourangeau's paper provides two interesting case studies about the role of multiple error sources in survey data. The first case study is one in which errors occur at different stages of the representation process—errors first occur when creating a potential sample frame, then may be amplified when selecting sampled persons, possibly because of self-selection, and then are exacerbated with an individual's decision to participate. The second case study has to do with situations where different error sources may influence each other and, in particular, the relationship between nonresponse error and various measurement error outcomes.

Each of these case studies are important, especially as survey researchers experience falling response rates (e.g., Brick and Williams 2013; Dutwin and Lavrakas 2016; Williams and Brick 2018), changing coverage of frames (e.g., Peytchev and Neely 2013; Battaglia, Dillman, Frankel, Harter, Buskirk, et al. 2016; McGeeney and Kennedy 2017), and a wide variety of both probability and nonprobability methods of sampling individuals (e.g., Dutwin and Buskirk 2017; Mercer, Kreuter, Keeter, and Stuart 2017; MacInnis, Krosnick, Ho, and Cho 2018). It is equally important to consider when errors may not yield compounding or cumulating errors. We may be able to use one approach in our survey design to deliberately offset errors, potentially reducing those correlations between propensity to be measured and the survey variable of interest. I address each of these cases in turn.

2. Case one: Representation errors for nonprobability web surveys

It is clear that there are issues related to representation or nonobservation in web surveys, especially nonprobability web panels. Web surveys have errors related to coverage (Tourangeau, Conrad, and Couper 2013), sample selection (Yeager, Krosnick, Chang, Javitz, Levendusky, et al. 2011; MacInnis et al. 2018), and nonresponse (Shih and Fan 2007). Tourangeau argues that these errors compound in web surveys. In particular, nonobservation errors will compound when the propensity to be covered, selected, and participate are similarly correlated with the outcome variables of interest, thus resulting in biases that systematically move estimates in a cumulatively positive or cumulatively negative direction.

Tourangeau provides evidence that sample demographics on characteristics such as age, education, race, and income differ at each stage of the coverage, selection, and participation when examining deviations from population benchmarks. Although there are demographic differences at each stage of the representation process, it is not clear from Tourangeau's argument what kind of survey estimates are the most likely to be correlated with the propensity to be measured, be it on the frame or in the sample, or to participate in the survey at all. Further, the theoretical motivation behind a correlation to be covered, sampled, and participate is also not clear.

To make sense of these multiple competing error sources, we can adapt existing models for how modes affect survey errors, including participation decisions, to both coverage and sample selection. Figure 1 shows a set of characteristics that my collaborators and I have examined as predictors of both mode and device decisions related to participation in probability-based surveys (Olson, Smyth, and Wood 2012; Smyth, Olson, and Millar 2014; Olson and Smyth 2018). They provide a useful categorization of the multiple influences that may jointly affect these three stages that drive errors of nonrepresentation, including having and using internet access, deciding to enroll in a web panel, and ultimately participating in a survey request. When these influences work in the same direction, we are likely to see the compounding error structure for which Tourangeau argues.

Figure 1 summarizes the various influences on coverage, selection, and participation in web surveys. Starting with the black oval at the top of figure 1, being familiar with and having access to the medium of interest—be it internet access at home or, more recently, a smartphone—affects one's ability to be part of the internet population (Couper 2000). This is the easiest error source to conceptually identify the items that are the most likely to be affected; is the estimate related to whether an individual has internet access or is the estimate related to knowing how to use the internet if they do have access to it? Characteristics that are strongly predictive of media familiarity and access are obvious indicators of having different types of media such as internet access or a smartphone for internet access,

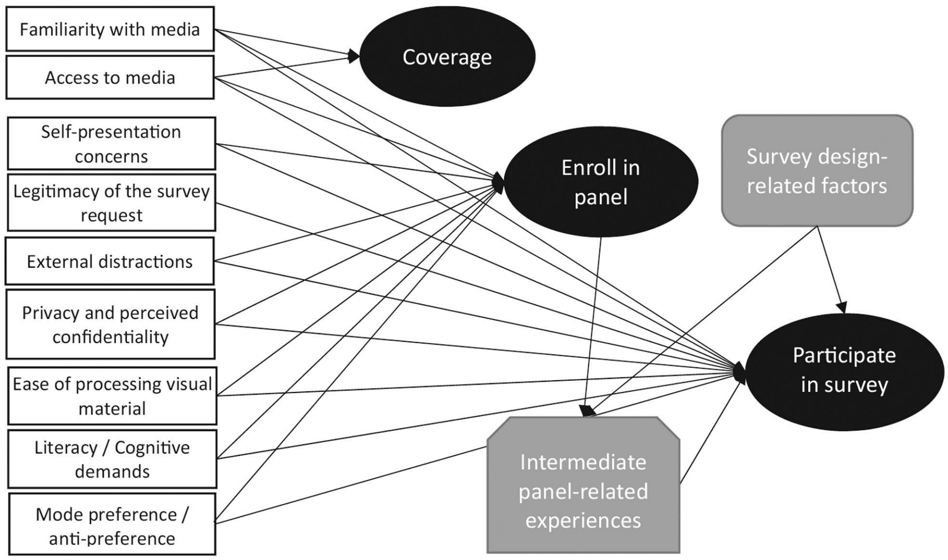

Figure 1. Potential Predictors of Coverage, Panel Enrollment, and Survey Participation in Web Surveys.

in addition to characteristics related to socioeconomic status such as income and potential usage characteristics such as age.

For example, Sterrett, Malato, Benz, Tompson, and English (2017) examine changes in internet access overall from 2006 to 2014 using the General Social Survey (GSS). According to the Sterrett et al.'s (2017) analysis of the GSS, internet coverage in the United States overall has increased over this time frame, from 69 percent to 86 percent of the US adult population; and simultaneously, differences on a variety of demographic characteristics between the covered and non-covered population has decreased, notably parity gains on education, income, and race distributions. Discrepancies still remain, however, as Tourangeau mentions. Recent data from the Pew Research Center (2018) show that 97 percent or higher of adults aged 18 to 49, college graduates, or individuals with a household income at $75,000 or above have internet access and a cell phone, and over 90 percent of these groups have a smartphone, compared with only 66 percent of adults age 65 and older. This also includes close correlates of access and familiarity, such as sex, in which men spend more time on a variety of internet activities, and race, in which black and Hispanic adults are much more likely to get online using their smartphone than having broadband internet access at home, according to Pew (see also Couper, Gremel, Axinn, Guyer, Wagner, et al. 2018). Web surveys based only on smartphone access, however, may have other decisions related to having a smartphone to access the internet compared with having internet access at home more generally (e.g., Antoun, Conrad, Couper, and West 2018).

The second black oval represents the decision to enroll in a panel, either in a probability- or a nonprobability-based web panel. The diagram identifies multiple features that likely influence these decisions. Unfortunately, good empirical data on who joins the wide variety of nonprobability web panels that exist are sparse and generally considered proprietary (Baker, Blumberg, Brick, Couper, Courtright, et al. 2010), and differences across nonprobability web panel vendors are vast (Kennedy, Mercer, Keeter, Hatley, McGeeney, et al. 2016). As such, our insights into nonprobability web panel self-selection mechanisms (along with company panel maintenance and refreshment decisions) are limited. However, we can speculate on a variety of influences for panel participation. First, factors in an individual deciding to join a web panel include having any access to the internet, having familiarity with the media and possibly even familiarity with surveys themselves, and having *regular* access to it. We can speculate that one is unlikely to want to be part of a web panel if there is not easy access to the internet at home or at work. Tourangeau mentions individuals who have few external distractions or other responsibilities that keep an individual from being able to participate in the panel as potential causes, concurrent with findings by Kennedy et al. (2016), who found that many online samples contained higher than expected rates of individuals who may have more discretionary time. There may be other influences, including being comfortable disclosing information in a self-administered format, feeling that the panel has sufficient confidentiality protections, or in general, having limited privacy concerns. Furthermore, being comfortable processing visual material, as the web is a visual mode, being literate, and having sufficient cognitive ability to process the survey requests (e.g., more highly educated in nonprobability samples; Malhotra and Krosnick 2007; Yeager et al. 2011), and, in general, preferring to take surveys or engage in tasks online (Kennedy et al. [2016] found higher rates of volunteerism and civic engagement on web panels) may also be predictive of selecting into a web panel.

The third black oval identifies the influences that also affect participation decisions in a particular mode. These influences are likely also modified by perceived legitimacy of the survey request, intermediate experiences that the respondent had as part of the panel itself, and any other survey design-related factors that are at play. For instance, it is a common finding that members of both nonprobability and probability-based web panels are members of multiple web panels and have participated in many previous surveys, suggesting that participation decisions are a function of previous panel experiences (e.g., Baker et al. 2010).

How does the confluence of these various factors influence the covariance between the propensity to be measured (*P*) and the survey variable (*Y*)? The covariance between *P* and *Y* is likely to be strongest when there are strong associations between an auxiliary variable and *P* and *Y* (Little and Vartivarian 2005; Kreuter, Olson, Wagner, Yan, Ezzati-Rice et al. 2010). Thus, each of these sets of characteristics presented in figure 1 can be used to identify proxy measures for the larger constructs underlying them and anticipate associations with other related variables. For instance, employment status is a commonly used measure of external distractions, but it is also tightly correlated with income. Thus, as Tourangeau argues, one may be less likely to be part of a web panel if you are employed, but you may be more likely to have internet access because of higher income. As such, these various joint influences may counteract each other.

How multiple influences on *P* affect the correlation between *P* and *Y* depends on the direction and strength of the association between these characteristics, propensity, and the survey variable of interest. In simulation work that Kreuter and Olson (2011) did on survey nonresponse, they found that when there are two variables (Z_1 and Z_2) that influence both the propensity to respond (*P*) and the survey variable (*Y*), the direction and magnitude of the association between these variables, response propensity, and the survey variables matter for both nonresponse bias of the unadjusted estimate and the efficacy of weights that are created using either or both of these variables. In particular, "the implications for bias and mean square error of adjusted respondent means are substantially different when the predictors have relationships of the same directions compared to when they have opposite directions with either propensity or the survey variables" (Kreuter and Olson 2011, p. 326). Thus, for errors to cumulate over multiple error sources, the association among the causal influence and propensity and the survey variables of interest must be reinforcing (i.e., in the same direction). As examples, Antoun et al. (2018) saw this reinforcing effects in a smartphone-based survey; Bosnjak, Haas, Galesic, Kaczmirek, Bandilla, et al. (2013) also saw reinforcing effects across coverage and nonresponse in a probability-based web panel for some characteristics (e.g., education), but not in others (e.g., age, personality traits).

We may be able to use these offsetting influences to our advantage. For instance, Olson and Smyth (2017) examined whether adding an explicit question to the cover of a mail survey would improve the accuracy of within-household selection procedures. We anticipated that the explicit question may depress response rates to the extent that household reporters tend not to follow the instructions and hand the survey to the correct respondent. This is what we found: there was a marginally lower response rate among those cases that had the verification question. However, this yielded a better sample; the condition with the verification question was closer to national benchmark estimates on a variety of demographic and socioeconomic questions, and more accurate selection of respondents within the household, both overall and in households with more than two adults where the selection can go wrong. Tourangeau, Kreuter, and Eckman (2012) found a similar tradeoff between nonresponse and complete information for selection within a household in a telephone survey.

Thus, offsetting errors can work in our favor. It may be that there are estimates for which these offsetting influences are at play across error sources in nonprobability web surveys; thinking causally about the joint influences and what may be positively or negatively associated is very important for understanding exactly when the covariance between P and Y overall is greater than that at each stage.

3. Case two: Nonresponse and measurement errors

Let us now turn to case two. Tourangeau summarizes Olson (2013) on the relationship between nonresponse and measurement errors. There is a clear association between the propensity to be a unit

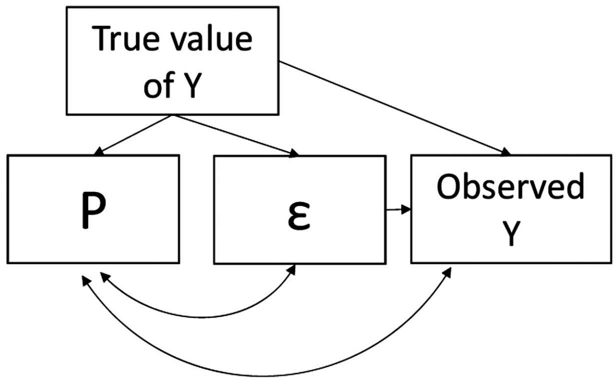

Figure 2. Conceptual Model for Relationship between Nonresponse and Measurement Error for Socially Desirable Characteristics.

nonrespondent and to fail to answer survey questions, but there is less of a clear relationship between unit nonresponse and other indicators of measurement error.

I appreciate the work that Tourangeau and his colleagues have done on socially desirable characteristics, nonresponse, and measurement error (e.g., Groves, Couper, Presser, Singer, Tourangeau, et al. 2006; Tourangeau, Groves, Kennedy, and Yan 2009; Fricker and Tourangeau 2010; Tourangeau, Groves, and Redline 2010). In some ways, this is the low hanging fruit, as are indicators of media familiarity and access for compounding errors in web surveys. Why might this be the case?

As shown in **Figure 2**, for some socially desirable characteristics, there is a clear link among the true value of the *Y* variable we are interested in, measurement errors in those reports (e), and survey participation (*P*). True nonvoters are less likely to participate in surveys, but when they do, they are less likely to accurately report that they did not vote (Tourangeau et al. 2010). Poor students are less likely to participate in surveys, but when they do, they are less likely to accurately report on their grades (Olson 2007; Sakshaug, Yan, and Tourangeau 2010). Similarly, Peytchev, Peytcheva, and Groves (2010) found that low response propensity respondents to the National Survey of Family Growth (NSFG) had the largest measurement errors in their reporting of ever having an abortion when comparing reports with interviewers with their reports in an audio computer-assisted self-interviewing (ACASI) system.

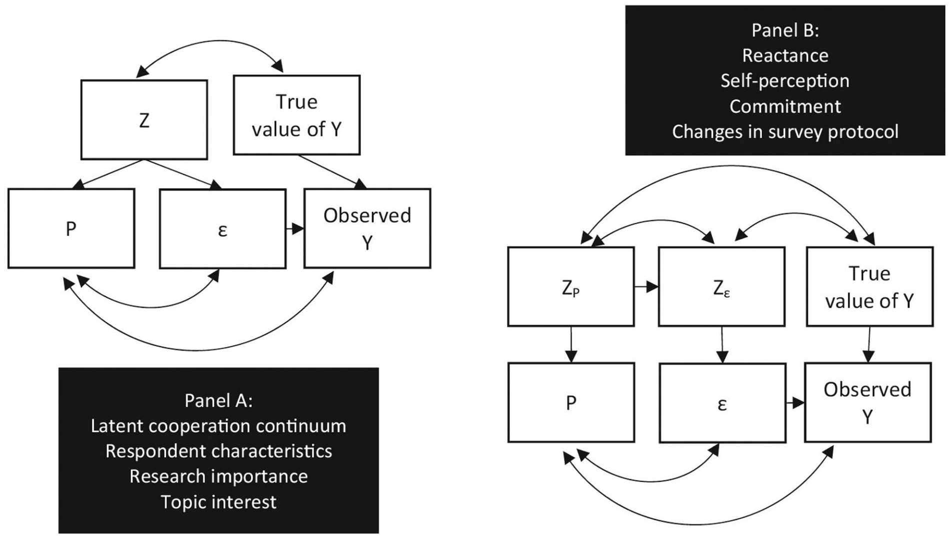

Figure 3. Conceptual Model for Relationship between Nonresponse and Measurement Error for Other Influences.

It would be interesting to look at other characteristics considered to be socially desirable or undesirable and the ways in which those questions are asked. For instance, Abraham, Helms, and Presser (2009) have shown that volunteers are more likely to participate than nonvolunteers in the American Time Use Survey (ATUS). But they argue that the measurement of volunteering in the ATUS is relatively free from social desirability bias, given the time diary in which the ATUS data are collected. Thus, it appears that the extent to which nonresponse and measurement errors are related or compounding for socially desirable characteristics is specific to question wording or the measurement stimulus—something that is generally overlooked.

Tourangeau describes a variety of other theoretical models for the relationship between nonresponse and measurement error, as outlined in Olson (2013). We can organize these other theoretical models for the relationship between nonresponse and measurement error into these diagrams, as shown in **Figure 3.**

For instance, as shown in panel A of figure 3, it is plausible that there is a measure *Z* that predicts both the propensity to respondent and measurement error that is separate from the true value of the survey variable itself. This may be the latent trait of cooperation, as so commonly posited. It is possible that this common cause between *P* and e is what is driving the clear and consistent relationship between unit and item nonresponse. For other measurement error outcomes, this set of common cause characteristics would lead to an association between propensity and measurement error, although the degree to which they affect the observed Y depends on the strength of the arrows between P and e, e and y, and P and Y. The respondent characteristics, research importance, and topic interest models fall into this diagram. For instance, one may be interested in the topic of the survey, and it may cause participation in a survey; there is plenty of evidence that shows that this occurs (e.g., Groves, Presser, and Dipko 2004; Groves et al. 2006;). Seeing the topic of the survey may also prime the respondent to overreport interest in this topic. It is plausible that this common cause of topic interest is what leads to the higher reports of birding even among birders themselves in the birders versus mall design experiment (Groves et al. 2006).

The other theoretical models linking nonresponse and measurement error displayed in panel B of figure 3 provide an even more distant connection between the causes of propensity and the causes of measurement error. In these models, there is something (Z_p) that influences the act of participation, and this Z_p also affects something that affects how people answer questions (Z_e). For instance, getting a lot of survey follow-up attempts may be a Z_p that creates a feeling of hostility in respondents; the Z_e causes them to provide less accurate data. This more distant collection of influences then will have more diffuse effects on the joint relationship between nonresponse and measurement errors, depending on the set of Z_ps and Z_es. With all of the conceptual models that exist, summarizing these models more succinctly will help us think more globally about nonresponse and measurement error.

There is still a lot that we do not know about why it is so hard to consistently find a link between nonresponse and measurement error. Does the link depend on the mode of data collection? Does the way in which error sources are linked vary across subgroups? Kreuter, Müller, and Trappmann (2010) provided evidence of variation in nonresponse and measurement error relationships across subgroups when looking at administrative data on welfare benefit receipt in Germany. How dependent is the link between these error sources on how questions themselves are constructed? Does an association depend on what exactly the respondent is told during recruitment? There are many other ways that we can think about these error sources together.

One other instance that Tourangeau did not mention, where we see joint relationships between errors, is something that I call *masquerading errors*. Here, we have errors that we can observe as one form of error, but they are actually hiding a second form of error. The example that is most salient here is that of interviewer-related variance. It is well established that interviewers have a clustering effect on survey estimates, reducing the precision of estimates and increasing their standard errors. In work by West, Kreuter, and Jaenichen (2013) and West and Olson (2010), it is clear that these interviewer variance effects are sometimes due to nonresponse, sometimes due to measurement error, and sometimes due to both. This is especially true for a respondent's age and age-related variables; different interviewers recruit respondents of different ages, but it manifests as what would look like a measurement error on age in our traditional models. Thus, nonresponse error is masquerading as measurement error.

4. Conclusion

Tourangeau's article provides an excellent overview of what we know about these multiple error sources. In sum, survey errors are not independent. They may compound each other, they may offset each other, they may be associated with one another, and they may hide one error as looking like a separate one. There is clearly more work to be done to understand how the different components of the Total Survey Error framework push and pull one other and ultimately affect our survey estimates.

References

Abraham, K. G., S. Helms, and S. Presser (2009), "How Social Processes Distort Measurement: The Impact of Survey Nonresponse on Estimates of Volunteer Work in the United States," *American Journal of Sociology*, 114, 1129–1165.

Antoun, C., F. G. Conrad, M. P. Couper, and B. T. West (2019), "Simultaneous estimation of multiple sources of error in a smartphone-based survey," *Journal of Survey Statistics and Methodology*, 7, 93–117.

Baker, R., S. J. Blumberg, J. M. Brick, M. P. Couper, M. Courtright, J. M. Dennis, D. Dillman, M.R. Frankel, P. Garland, R.M. Groves, C. Kennedy, J. Krosnick, P.J. Lavrakas, S. Lee, M. Link, L. Piekarski, K. Rao, R.K. Thomas, and D. Zahs. (2010), "Research Synthesis: AAPOR Report on Online Panels," *Public Opinion Quarterly*, 74, 711–781.

Battaglia, M. P., D. A. Dillman, M. R. Frankel, R. Harter, T. D. Buskirk, C. B. McPhee, J. M. DeMatteis, and T. Yancey (2016), "Sampling, Data Collection, and Weighting Procedures for Address-Based Sample Surveys," *Journal of Survey Statistics and Methodology*, 4, 476–500.

Bosnjak, M., I. Haas, M. Galesic, L. Kaczmirek, W. Bandilla, and M. P. Couper (2013), "Sample Composition Discrepancies in Different Stages of a Probability-Based Online Panel," *Field Methods*, 25, 339–360.

Brick, J. M., and D. Williams (2013), "Explaining Rising Nonresponse Rates in Cross-Sectional Surveys," *The ANNALS of the American Academy of Political and Social Science*, 645, 36–59.

Couper, M. P. (2000), "Web Surveys: A Review of Issues and Approaches," *Public Opinion Quarterly*, 64, 464–494.

Couper, M. P., G. Gremel, W. Axinn, H. Guyer, J. Wagner, and B. T. West (2018), "New Options for National Population Surveys: The Implications of Internet and Smartphone Coverage," *Social Science Research*, 73, 221–235.

Dutwin, D., and T. D. Buskirk (2017), "Apples to Oranges or Gala versus Golden Delicious? Comparing Data Quality of Nonprobability Internet Samples to Low Response Rate Probability Samples," *Public Opinion Quarterly*, 81, 213–239.

Dutwin, D., and P. J. Lavrakas (2016), "Trends in Telephone Outcomes, 2008-2015," *Survey Practice*, 9, 1.

Fricker, S., and R. Tourangeau (2010), "Examining the Relationship between Nonresponse Propensity and Data Quality in Two National Household Surveys," *Public Opinion Quarterly*, 74, 934–955.

Groves, R. M., M. P. Couper, S. Presser, E. Singer, R. Tourangeau, G. Piani Acosta, and L. Nelson (2006), "Experiments in Producing Nonresponse Bias," *Public Opinion Quarterly*, 70, 720–736.

Groves, R. M., S. Presser, and S. Dipko (2004), "The Role of Topic Interest in Survey Participation Decisions," *Public Opinion Quarterly*, 68, 2–31.

Kennedy, C., A. Mercer, S. Keeter, N. Hatley, K. McGeeney, and A. Gimenez (2016), "Evaluating Online Nonprobability Surveys." Available at http://www.pewresearch.org/methods/2016/05/02/evaluating-online-nonprobability-surveys/ (May 2, 2016). Last accessed December 7, 2018.

Kreuter, F., G. Müller, and M. Trappmann (2010), "Nonresponse and Measurement Error in Employment Research: Making Use of Administrative Data," *Public Opinion Quarterly*, 74, 880–906.

Kreuter, F., and K. Olson (2011), "Multiple Auxiliary Variables in Nonresponse Adjustment," *Sociological Methods and Research*, 40, 311–332.

Kreuter, F., K. Olson, J. Wagner, T. Yan, T. M. Ezzati-Rice, C. Casas-Cordero, M. Lemay, A. Peytchev, R. M. Groves, and T. E. Raghunathan (2010), "Using Proxy

Measures and Other Correlates of Survey Outcomes to Adjust for Nonresponse: Examples from Multiple Surveys," *Journal of the Royal Statistical Society Series A-Statistics in Society*, 173, 389–407.

Little, R. J. A., and S. Vartivarian (2005), "Does Weighting for Nonresponse Increase the Variance of Survey Means?," *Survey Methodology*, 31, 161–168.

MacInnis, B., J. A. Krosnick, A. S. Ho, and M.-J. Cho (2018), "The Accuracy of Measurements with Probability and Nonprobability Survey Samples: Replication and Extension," *Public Opinion Quarterly*, 82, 707–744.

Malhotra, N., and J. A. Krosnick (2007), "The Effect of Survey Mode and Sampling on Inferences about Political Attitudes and Behavior: Comparing the 2000 and 2004 ANES to Internet Surveys with Nonprobability Samples," *Political Analysis*, 15, 286–323.

McGeeney, K., and C. Kennedy (2017), "A Next Step in the Evolution of Landline Sampling: Evaluating the Assignment-Based Frame," *Journal of Survey Statistics and Methodology*, 5, 37–47.

Mercer, A. W., F. Kreuter, S. Keeter, and E. A. Stuart (2017), "Theory and Practice in Nonprobability Surveys: Parallels between Causal Inference and Survey Inference," *Public Opinion Quarterly*, 81, 250–271.

Olson, K. (2007), "An Investigation of the Nonresponse/Measurement Error Nexus," unpublished Ph.D. dissertation, University of Michigan, Ann Arbor, MI.

———. (2013), "Do Non-Response Follow-Ups Improve or Reduce Data Quality?: A Review of the Existing Literature," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176, 129–145.

Olson, K., and J. D. Smyth (2017), "Within-Household Selection in Mail Surveys: Explicit Questions Are Better than Cover Letter Instructions," *Public Opinion Quarterly*, 81, 688–713.

———. (2018), "Are You Busy? Please Do This Survey on Your Phone: The Effects of Cover Letter Language on Mobile Web Participation in a Mixed-Mode General Population Survey," paper presented at the Annual Conference of the American Association for Public Opinion Research, Denver, CO.

Olson, K., J. D. Smyth, and H. Wood (2012), "Does Giving People Their Preferred Survey Mode Actually Increase Survey Participation Rates?," *Public Opinion Quarterly*, 76, 611–635.

Pew Research Center (2018), "Internet/Broadband Fact Sheet." Available at http://www.pewinternet. org/fact-sheet/internet-broadband/ (February 8, 2018). Last accessed December 6, 2018.

Peytchev, A., and B. Neely (2013), "RDD Telephone Surveys: Toward a Single-Frame Cell-Phone Design," *Public Opinion Quarterly*, 77, 283–304.

Peytchev, A., E. Peytcheva, and R. M. Groves (2010), "Measurement Error, Unit Nonresponse and Self-Reports of Abortion Experiences," *Public Opinion Quarterly*, 74, 319–327.

Sakshaug, J. W., T. Yan, and R. Tourangeau (2010), "Nonresponse Error, Measurement Error, and Mode of Data Collection: Tradeoffs in a Multi-Mode Survey of Sensitive and Non-Sensitive Items," *Public Opinion Quarterly*, 74, 907–933.

Shih, T.-H., and X. Fan (2007), "Response Rates and Mode Preferences in Web-Mail Mixed- Mode Surveys: A Meta-Analysis," *International Journal of Internet Science*, 2, 59–82.

Smyth, J. D., K. Olson, and M. M. Millar (2014), "Identifying Predictors of Survey Mode Preference," *Social Science Research*, 48, 135–144.

Sterrett, D., D. Malato, J. Benz, T. Tompson, and N. English (2017), "Assessing Changes in Coverage Bias ofWeb Surveys in the United States," *Public Opinion Quarterly*, 81, 338–356.

Tourangeau, R., F. G. Conrad, and M. P. Couper (2013), *The Science of Web Surveys*, New York, NY: Oxford University Press.

Tourangeau, R., R. M. Groves, C. Kennedy, and T. Yan (2009), "The Presentation of a Web Survey, Nonresponse and Measurement Error among Members of Web Panel," *Journal of Official Statistics*, 25, 299–321.

Tourangeau, R., R. M. Groves, and C. D. Redline (2010), "Sensitive Topics and Reluctant Respondents: Demonstrating a Link between Nonresponse Bias and Measurement Error," *Public Opinion Quarterly*, 74, 413–432.

Tourangeau, R., F. Kreuter, and S. Eckman (2012), "Motivated Underreporting in Screening Interviews," *Public Opinion Quarterly*, 76, 453–469.

West, B. T., F. Kreuter, and U. Jaenichen (2013), "Interviewer" Effects in Face-to-Face Surveys: A Function of Sampling, Measurement Error, or Nonresponse?," *Journal of Official Statistics*, 29, 277–297.

West, B. T., and K. Olson (2010), "How Much of Interviewer Variance Is Really Nonresponse Error Variance?," *Public Opinion Quarterly*, 74, 1004–1026.

Williams, D., and J. M. Brick (2018), "Trends in U.S. Face-to-Face Household Survey Nonresponse and Level of Effort," *Journal of Survey Statistics and Methodology*, 6, 186–211.

Yeager, D. S., J. A. Krosnick, L. Chang, H. S. Javitz, M. S. Levendusky, A. Simpser, and R.Wang (2011), "Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples," *Public Opinion Quarterly*, 75, 709–747.