

Conditional Pop-up Reminders Reduce Incidence of Rounding in Web Surveys

Field Methods
2022, Vol. 34(4) 334–345
© The Author(s) 2022



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1525822X221115829
journals.sagepub.com/home/fmx



Rainer Schnell¹, Sarah Redlich¹, and Anja S. Göritz²

Abstract

Frequency of behaviors or amounts of variables of interest are essential topics in many surveys. The use of heuristics might cause rounded answers, resulting in the increased occurrence of end-digits (called heaping or digit-preference). For web surveys (or CASI), we propose using a conditional prompt as input validation if digits indicating heaping are entered. We report an experiment, where respondents in an online access panel ($n = 2,590$) were randomly assigned to one of three groups: (1) no input validation; (2) conditional input validation if rounding was presumed; and (3) input validation every time a numerical value was entered. Conditional input validation reduces heaping for variables with high proportions of heaped values. Unconditional input validation seems to be less effective.

Introduction

Blair and Burton (1986, 1987) were the first to extensively review the problems associated with asking for the frequency of certain behaviors from respondents in an open-ended question. Respondents tend not to answer such questions in as much detail as intended by researchers. Instead, the

¹Research Methodology Group, University of Duisburg-Essen, Duisburg, Germany

²Department of Psychology, University of Freiburg, Freiburg, Germany

Corresponding Author:

Rainer Schnell, Research Methodology Group, University of Duisburg-Essen, 47057 Duisburg, Forsthausweg 2, Germany.

Email: rainer.schnell@uni-due.de

answers often are multiples of certain numbers, mostly five or 10. Such behavior on the part of the respondent is termed *heaping* (also called *digit-preference*).

Vaske and Beaman (2006) provided an overview of heaping in surveys and its implications. For instance, response heaps do not only occur at 0 and 5, but potentially at any number and gaps between response heaps increase with a growing range of responses (Vaske and Beaman 2006:286–91). An example of the occurrence of heaping has been recognized as a problem in demography (evident in census counts), where questions asking for a person's age often yield nonuniform distributed terminal digits (Hobbs 2004). Furthermore, Wang and Heitjan (2008) showed that heaping could lead to biased descriptive statistics by distorting the underlying distribution, and it can affect inferences.

Heaping might be seen as the consequence of not counting each and every occurrence of the behavior to be reported. Problems due to heaping can be approached in two ways: Either by preventing during data collection or correcting already available data. Most of the literature is dedicated to attempts for correcting available data (see, e.g., Allen et al. 2017; Bar and Lillard 2012; Crawford et al. 2015; Wolff and Augustin 2003; Zinn and Würbach 2015). For example, Zinn and Würbach (2015) used a zero-inflated log-normal distribution for self-reported income. In contrast, attempts at preventing heaping during data collection have been rare. One of the few examples is Becker and Diop-Sidibé (2003), who proposed using a calendar to reduce heaping when recording durations.

Tourangeau et al. (2004:235) list three causes for rounding: (1) imprecise encoding of the information in memory; (2) indeterminacy in the underlying quantities (e.g., the actual price of a house is unknown until it is being sold); and (3) the burden of retrieving numerous specific pieces of information.¹ From the perspective of a survey designer, only the last mechanism can be used to reduce heaping since the other two are properties of the memory encoding or correspond to the true state of nature. In some contexts, the subjectively perceived burden can be reduced by increasing the motivation of respondents, for example, by emphasizing the importance of the answer (Cannell et al. 1981:404). Therefore, to increase motivation, prompts have been used, for example, to reduce nonresponse (DeRouvray and Couper 2002) or improve the quality of the given answer (Conrad et al. 2005).²

This article reports on a randomized experiment, where two different input validation prompts are used to reduce heaping during data collection in a general population web survey ($n = 2,590$). Respondents are either (1) not prompted; (2) prompted if rounding was detected; or (3) prompted every time a numerical answer was given. The conditional feedback on rounding behavior could be seen as an external motivation to increase the cognitive effort. Therefore, an input validation prompt should reduce heaping.

Methods

The proportions of rounded answers are compared between groups using a test for differences in proportions. The test statistic is

$$z = \frac{\hat{\pi}_1 - \hat{\pi}_2}{\sqrt{(\hat{\pi}_p(1 - \hat{\pi}_p))(1/n_1 + 1/n_2)}} \quad (1)$$

with

$$\hat{\pi}_p = \frac{x_1 + x_2}{n_1 + n_2} \quad (2)$$

where π_1 and π_2 are the proportions of rounded answers within the two experimental groups, n_1 and n_2 the number of responses per group, and x_1 and x_2 are the number of persons with rounded answers (see, e.g., [Agresti et al. 2017:478f.](#)). One-sided tests were considered.

The statistical effect size for the difference between two proportions (p_1 and p_2) is usually evaluated using Cohen's h which is defined as ([Gleser and Olkin 2009:364](#))

$$h = 2 \arcsin\sqrt{p_1} - 2 \arcsin\sqrt{p_2} \quad (3)$$

Data

Results reported here are based on an experiment using all panel members of an academically managed non-probability online panel (WiSo Panel, [Göritz 2014](#)). As with many other non-probability online surveys ([Callegaro et al. 2015](#)), this panel has used different online recruitment methods such as banner ads, search engines, online networks, and newsletters. In addition, panel members are encouraged to recruit further panel members ([Crutzen and Göritz 2012:196](#)). Of course, non-probability samples should not be used for point estimates for a population ([Baker et al. 2010](#)), but estimating experimental effects is widely considered as legitimate within bounds given by the heterogeneity of the sampled population ([Kohler et al. 2019:166](#)).

The resulting sample ($n = 2,590$) consisted of 60.8% women and 39.2% men; ages ranged from 15 to 89 years (with a mean of 44). A comparison of the distribution of educational level in the sample and the German census is shown in [Table 1](#). As in many other access panels, the educational level of the respondents is biased in favor of more educated respondents ([McCutcheon et al. 2014:115](#)).³

The survey contained questions on:

- (1) the number of alcoholic beverages per week,
- (2) the number of cigarettes smoked per day,

Table 1. Educational Level in the Sample (Unweighted, in Percent). Census Data for People above 14 Years of Age. The Sample Contains a Larger Proportion of Highly Educated People than the General Public.

Educational level	Sample	Census
No degree	1.2	7.1
Nine years of school	11.4	35.8
Vocational qualification	28.1	28.7
Higher education entry qualification	31.2	12.7
University degree	26.1	13.6
Doctorate	2.0	1.2

- (3) kilometers driven by car per week,
- (4) expenses in euros for clothing in the previous month,
- (5) the number of orders of electronic items in the previous 12 months.⁴

The questions were presented in random order. The respondents were (disproportionately) randomly assigned to one of three groups:

- (1) no input validation ($n = 653$, $\approx 25\%$),
- (2) conditional input validation, that is, a prompt whenever an answer given to a question was probably rounded ($n = 1,293$, $\approx 50\%$),
- (3) input validation with a prompt for every question (unconditional input validation) if a number larger than zero was entered ($n = 644$, $\approx 25\%$).

For the conditional input validation, an answer was seen as rounded if it was a multiple of five for alcohol consumption, cigarettes, and order of electronic items. Additionally, for alcohol consumption, a weekly reoccurrence (multiple of seven) and for order of electronic items a monthly reoccurrence (multiple of 12) was considered a rounded answer. For kilometers driven and clothing expenses, an answer was seen as rounded if it was a multiple of 50.

The input validation prompt read: “If you guessed this answer, please can you give a precise answer” (translated from German).⁵

After editing for incomplete, zero, and implausible responses, 2,590 cases with 1,594 responses for alcohol consumption, 775 responses for cigarette counts, 2,190 responses for kilometers driven by car, 1,987 responses for clothing expenses, and 1,411 responses for orders of electronic items remained.⁶

Results

Table 2 and Figure 1 show the proportion of rounded responses per variable. Between 14 and 71% of all responses were rounded (considering a multiplier of five).

Table 2. Proportions of Responses of Multiples for Each Input Validation Group: Conditional (C), Unconditional (UC), No Input Validation (No). Conditional Input Validation Shows Lower Proportions of Heaped Responses than No, as Well as Unconditional Input Validation.

Heaping on	Orders				Alcohol				Cigarettes			
	No	C	UC	Total	No	C	UC	Total	No	C	UC	Total
Multiplier of												
5	0.15	0.12	0.17	0.14	0.23	0.19	0.23	0.21	0.75	0.67	0.73	0.71
7					0.05	0.06	0.08	0.06				
12	0.01	0.01	0.01	0.01								

Heaping on	Km				Expenses			
	No	C	UC	Total	No	C	UC	Total
Multiplier of								
50	0.59	0.51	0.52	0.53	0.55	0.40	0.47	0.45

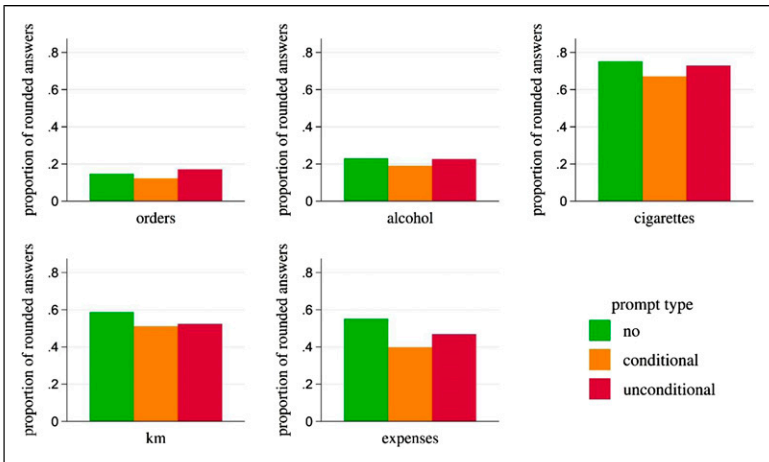


Figure 1. Proportions of rounded answers (five for electronic orders, alcoholic beverages, and cigarettes; 50 for driven kilometers, and expenses). Conditional input validation always yields lower proportions than no or unconditional input validation.

As expected, for clothing expenses (in euros) and kilometers driven by car pronounced heaping was found for multiples of 50. Furthermore, for cigarettes, large proportions of rounded responses are seen for multiples of five. For the other two variables of interest (beverages and electronic orders), heaping could neither be observed for multiples of five nor for weekly or

monthly reoccurrences. Therefore, given that 75% of all respondents reported numbers less than five (alcohol) or four (electronic orders), it is likely, that for these rare occurrences, the respondents actually recalled the individual episodes.

The experimental results (see Table 2) correspond to the theoretically expected order of the amount of heaping. For four variables of interest, the largest proportion of rounded answers is found in the group with no input validation (the exception is electronic orders, where the largest proportion is found in the unconditional prompting group).

For all five variables, a conditional input validation resulted in lower proportions of rounded answers than in the other two groups. In contrast, the unconditional reminder resulted for three variables in a decrease. Furthermore, on all five variables, the decrease is smaller for unconditional than the decrease for the conditional prompt.

These effects of the experimental conditions on differences in the proportions of rounded answers were tested by pairwise comparisons (no prompt versus conditional prompting, no prompt versus unconditional prompt; see Table 3).

For three variables (cigarettes, kilometers, and expenses), the decrease in heaping was significant for the conditional prompt. For the unconditional prompt, the decrease was significant for two variables (kilometers and expenses). As can be seen in Figure 1, the three variables that showed a significant reduction in heaping for conditional prompts are those variables with the largest amount of heaping in all experimental groups.⁷

The literature repeatedly reported more heaping with increasing magnitude of the response (Vaske and Beaman 2006:291). Due to the high skewness of

Table 3. Z Statistics: Pairwise Comparison of Proportions of Rounded Answers of Experimental Groups (Unweighted; One-sided). Variables Are Ordered According to Median Values for all Respondents. All Observed Differences Correspond to Theoretical Expectations: Prompting Reduces Heaping, Larger Reductions Are Observed for Conditional Prompting. for Both Experimental Conditions, the Size of the Effect Increases from Top to Bottom (i.e., with Increasing Median Value of the Variable).

	Conditional versus no prompt	Unconditional versus no prompt
Orders	-1.12	0.81
Alcohol	-1.64	-0.16
Cigarettes	-2.02*	-0.52
Km	-2.96**	-2.13*
Expenses	-5.65**	-2.62**

* $p < 0.05$, ** $p < 0.01$.

the answers to many behavioral frequency questions, we considered the median response given as a better indicator for the magnitude than the range or the maximum. Therefore, we ordered the variables according to the median value of the variable of interest (disregarding experimental groups in the computation of the median). For both experimental conditions, the size of the experimental effect increases with an increasing median value of the variable (note the decreasing value of the test statistic in [Table 3](#) from top to bottom in both columns). Thus, prompting seems to have a larger effect if higher values of response variables are to be expected.

The effect sizes for the variables of interest are shown in [Table 4](#). The proposed conditional prompt yields an average effect of 0.17. [Cohen \(1988: 184\)](#) considered values of h about 0.2 as small effects. However, conditional prompts resulted in an absolute decrease of 3% (orders) to 15% (expenses).

In summary, for all variables, the conditional prompt resulted in a higher amount of decrease than the unconditional prompt, and the difference between both prompts increases with the magnitude of the response. Therefore, given the small additional costs due to the conditional prompt, we would recommend using conditional prompts in behavioral frequency questions in web surveys.

Discussion

Behavioral frequency questions and questions asking for numerical estimates tend to produce multi-modal frequency distributions due to heaping. Therefore, statistical modeling might be more difficult and might even result in biased estimates ([Wang and Heitjan 2008](#)).

Although most of the literature focuses on correction methods after the data collection, here we proposed a method to reduce heaping during the data collection. The conditional prompt after presumed rounded answers significantly reduced the amount of heaping in all three variables with severe rounding. Conditional input validation reduces heaping for variables with

Table 4. Effect Size (Cohen's h) for the Difference between Experimental Groups and Control Group.

	Conditional versus no prompt	Unconditional versus no prompt
Orders	-0.09	0.05
Alcohol	-0.10	0.00
Cigarettes	-0.18	-0.05
Km	-0.16	-0.14
Expenses	-0.30	-0.16

high proportions of heaped values. Unconditional input validation seems to be less effective.

This diminishing result of repetitive and unreliable warnings and alarms might be similar to the long-known “cry wolf” effect in engineering alarm systems (Bliss et al. 1995; Breznitz 1984). If the goal of the alarm is perceived as important and the alarm has a high positive predictive value, the cry wolf effect is not to be expected. However, if neither is present, the cry wolf effect will be seen (Johnson et al. 2017). Therefore, the success of the input validation will depend on whether the respondent task was perceived as important and if the prompt is perceived as legitimate. Later research should examine if this mechanism is actually the cause of the observed differences in the effectiveness of the prompts.

However, prompting seems to reduce heaping and larger reductions have been observed for conditional prompting. Therefore, a prompt motivated by a presumed rounded answer might be the best design option for reducing heaping in numerical answers to behavioral frequency questions.

The results reported here are based on behavioral frequency questions. If this effect also applies to other questions that might lead to rounding (Holbrook et al. 2014) such as personal characteristic questions, questions about age at the time of an event, questions about percentages, or feeling-thermometers remains to be studied.

A limitation of this between-subjects experiment is due to the software used for implementation. The program does not allow the storage of the initial answers after a changed response following a prompt. In further research, the initial answer should be recorded and compared to later corrections. In addition, at least a subsample of respondents correcting their previous answer should receive a follow-up question if the updated answer is considered an accurate response or a measure to prevent further prompts. A further limitation of this study is common in research on heaping: The lack of external validation data, particularly for behavioral frequency questions, makes the evaluation of measures to reduce heaping difficult. The mentioned follow-up question might partially compensate for the absence of external validation data to evaluate the correctness of the given response. Furthermore, in a sequence of numerical questions, the optimal time spacing of conditional prompts needs to be determined.

Another limitation is due to the non-probability sampling. However, most web surveys (Baker et al. 2010) are based on non-probability sampling methods. The panel used here holds participants from all walks of life who were recruited from a variety of sources. Although the treatment effect in a general population survey might be different (Kohler et al. 2019), it is unlikely that the effect will not be detected in a general population survey. Whether a given non-probability sample is fit for the purpose of the study depends on the purpose and the sampling methods used (Baker et al. 2013). The statistical

conditions necessary for the generalization of experimental results are described by Kohler et al. (2019). An overestimated treatment effect would be due to a high correlation between the selection probability for the web survey and the size of the motivational effect of the prompt. Since we found no effects of education, gender, and age on the size of the effect, we see no evidence for an overestimation due to non-probability sampling. Therefore, we consider the sample as fit for the purpose of this study.

Finally, replications of the experiment in other general population surveys seem to be necessary.

Acknowledgments

We thank the four anonymous reviewers for the detailed reviews and helpful comments to improve this article.

Declaration of Conflicting Interests

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

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Therefore, some forms of heaping might be due to perceived cognitive burden and thus might be prone to satisficing (Krosnick 1991; Roberts et al. 2019). However, systematic studies of rounding and satisficing are rare (e.g., Gideon et al. 2017; Holbrook et al. 2014; Turner et al. 2015) and currently seem to suggest that heaping might not be a form of satisficing. We consider the study by Holbrook et al. (2014) as not conclusive: The fact that they used paid respondents might have impacted on the motivation to answer behavioral frequency questions and their interpretation is based on logistic regression models whose model fits are not reported.
2. The use of input validation prompts could be seen as an example of the sixth type (complex calculation validations) of the typology of real-time validations in web-based surveys by Peytchev and Crawford (2005).
3. Since weighting only reduces bias if the weighting variables are strongly related to the variables of interest (Bethlehem and Biffignandi 2012), we considered weighting as of no use. A reanalysis with a weight including education led to the same conclusions.
4. The questions read (translated from German): Alcohol: "How many alcoholic beverages do you usually drink in a week? By one alcoholic beverage, we mean one bottle of beer 0.33 l, one glass of wine 0.125 l, one glass of sparkling wine, a liquor

or a bottle of alcopops.” Cigarettes: “Approximately how many cigarettes do you smoke per day?” Kilometers: “How many kilometers per week do you drive or ride in a private car?” Expenses: “What amount of money in EUR did you spend on your personal clothing last month?” Orders: “In the last 12 months, how often have you ordered electronic items (including cameras and computer hardware) for personal use over the Internet?”

5. In German: “Falls Sie bei dieser Antwort geschätzt haben sollten, bitten wir Sie um eine genauere Antwort.”
6. Of the initial 2,921 sample members, 300 failed to answer any question of interest. Two respondents did not complete the questionnaire. Respondents reporting zero occurrences of the behavior being asked for were excluded since it indicates the absence of the behavior in question ($n = 31$). Finally, responses more than three standard deviations from the mean were considered as outliers and excluded. For kilometers driven, the upper limit of plausible values was considered as the maximum legal working hours for a week multiplied by 100 (upper limit = 6,000 kilometers).
7. Furthermore, separate logistic regressions for each variable for the prediction of rounding versus non-rounding were used to control for kind of input validation, order of the questions, and demographics of respondents. All models yielded pseudo- R^2 values close to zero. However, the coefficient for conditional reminders always indicates less rounding than the coefficient for unconditional reminders. Furthermore, we did not observe any impact of respondents’ fatigue as measured by time of day or order of questions on the amount of rounding.

References

- Agresti, A., C. Franklin, and B. Klingenberg. 2017. *Statistics: The art and science of learning from data*. 4th ed. Harlow, London: Pearson.
- Allen, C. M., S. D. Griffith, S. Shiffman, and D. F. Heitjan. 2017. Proximity and gravity: Modeling heaped self-reports. *Statistics in Medicine* 36:3200–15.
- 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. AAPOR report on online panels. *Public Opinion Quarterly* 74:711–81.
- Baker, R., J. M. Brick, N. A. Bates, M. Battaglia, M. P. Couper, J. A. Dever, K. J. Gile, and R. Tourangeau. 2013. Summary report of the AAPOR task force on non-probability sampling. *Journal of Survey Statistics and Methodology* 1:90–105.
- Bar, H. Y., and D. R. Lillard. 2012. Accounting for heaping in retrospectively reported event data—A mixture-model approach. *Statistics in Medicine* 31:3347–65.
- Becker, S., and N. Diop-Sidibé. 2003. Does use of the calendar in surveys reduce heaping? *Studies in Family Planning* 34:127–32.
- Bethlehem, J., and S. Biffignandi. 2012. *Handbook of web surveys*. Hoboken, NJ: Wiley.

- Blair, E., and S. Burton. 1986. Processes used in the formulation of behavioral frequency reports in surveys. In *Proceedings of the Section on Survey Research*. American Statistical Association, 481–87. http://www.asasrms.org/Proceedings/papers/1986_090.pdf (accessed May 16, 2022).
- Blair, E., and S. Burton. 1987. Cognitive processes used by survey respondents to answer behavioral frequency questions. *Journal of Consumer Research* 14: 280–88.
- Bliss, J. P., R. D. Gilson, and J. E. Deaton. 1995. Human probability matching behaviour in response to alarms of varying reliability. *Ergonomics* 38:2300–12.
- Breznitz, S. 1984. *Cry wolf: The psychology of false alarms*. Hillsdale, NJ: Lawrence Erlbaum.
- Callegaro, M., K. L. Manfreda, and V. Vehovar. 2015. *Web survey methodology*. Los Angeles: Sage.
- Cannell, C. F., P. V. Miller, and L. Oksenberg. 1981. Research on interviewing techniques. *Sociological Methodology* 12:389–437.
- Cohen, J. 1988. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Conrad, F. G., M. P. Couper, R. Tourangeau, and M. Galesic. 2005. Interactive feedback can improve the quality of responses in web surveys. In *Proceedings of the survey research methods section*. ASA, 3835–40. <http://www.asasrms.org/Proceedings/y2005/files/JSM2005-000938.pdf> (accessed May 16, 2022).
- Crawford, F. W., R. E. Weiss, and M. A. Suchard. 2015. Sex, lies, and self-reported counts: Bayesian mixture models for heaping in longitudinal count data via birth–death processes. *Annals of Applied Statistics* 9:572–96.
- Crutzen, R., and A. S. Göritz. 2012. Public awareness and practical knowledge regarding hepatitis A, B, and C: A two-country survey. *Journal of Infection and Public Health* 5:195–98.
- DeRouvray, C., and M. P. Couper. 2002. Designing a strategy for reducing “No Opinion” responses in web-based surveys. *Social Science Computer Review* 20: 3–9.
- Gideon, M., B. Helleppie-McFall, and J. W. Hsu. 2017. Heaping at round numbers on financial questions: The role of satisficing. *Survey Research Methods* 11:189–214.
- Gleser, L. J., and I. Olkin. 2009. Stochastically dependent effect sizes. In *The handbook of research synthesis and meta-analysis*. 2nd ed., eds. H. Cooper, L. V. Hedges, and J. C. Valentine, 357–76. New York: Russell Sage.
- Göritz, A. S. 2014. Determinants of the starting rate and the completion rate in online panel studies. In *Online panel research: A data quality perspective*, eds. M. Callegaro, R. Baker, J. Bethlehem, A. S. Göritz, J. A. Krosnick, and P. J. Lavrakas, 154–70. Chichester, UK: Wiley.
- Hobbs, F. 2004. Age and sex composition. In *The methods and materials of demography*. 2nd ed., eds. J. S. Siegel and D. A. Swanson, 125–73. London: Elsevier.

- Holbrook, A. L., S. Anand, T. P. Johnson, Y. I. Cho, S. Shavitt, N. Chávez, and S. Weiner. 2014. Response heaping in interviewer-administered surveys: Is it really a form of satisficing? *Public Opinion Quarterly* 78:591–633.
- Johnson, K. R., J. I. Hagadorn, and D. W. Sink. 2017. Alarm safety and alarm fatigue. *Clinics in Perinatology* 44:713–28.
- Kohler, U., F. Kreuter, and E. A. Stuart. 2019. Nonprobability sampling and causal analysis. *Annual Review of Statistics and Its Application* 6:149–72.
- Krosnick, J. A. 1991. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology* 5:213–36.
- McCutcheon, A. L., K. Rao, and O. Kaminska. 2014. The untold story of multi-mode (online and mail) consumer panels: From optimal recruitment to retention and attrition. In *Online panel research: A data quality perspective*, eds. M. Callegaro, R. Baker, J. Bethlehem, A. S. Göritz, J. A. Krosnick, and P. J. Lavrakas, 104–26. Chichester, UK: Wiley.
- Peytchev, A., and S. Crawford. 2005. A typology of real-time validations in web-based surveys. *Social Science Computer Review* 23:235–49.
- Roberts, C., E. Gilbert, N. Allum, and L. Eisner. 2019. Satisficing in surveys: A systematic review of the literature. *Public Opinion Quarterly* 83:598–626.
- Tourangeau, R., L. J. Rips, and K. Rasinski. 2004. *The psychology of survey response*. Cambridge: Cambridge University Press.
- Turner, G., P. Sturgis, and D. Martin. 2015. Can response latencies be used to detect survey satisficing on cognitively demanding questions? *Journal of Survey Statistics and Methodology* 3:89–108.
- Vaske, J. J., and J. Beaman. 2006. Lessons learned in detecting and correcting response heaping: Conceptual, methodological, and empirical observations. *Human Dimensions of Wildlife* 11:285–96.
- Wang, H., and D. F. Heitjan. 2008. Modeling heaping in self-reported cigarette counts. *Statistics in Medicine* 27:3789–804.
- Wolff, J., and T. Augustin. 2003. Heaping and its consequences for duration analysis: A simulation study. *Advances in Statistical Analysis* 87:59–86.
- Zinn, S., and A. Würbach. 2015. A statistical approach to address the problem of heaping in self-reported income data. *Journal of Applied Statistics* 43:682–703.