Public Opinion Quarterly, Vol. 83, Special Issue 2019, pp. 185–209

HOW TRANSPARENCY AFFECTS SURVEY RESPONSES

ELIZABETH C. CONNORS
YANNA KRUPNIKOV
JOHN BARRY RYAN*

Abstract Following a shift toward greater transparency, many academic journals across a variety of disciplines now require authors to post their data. At the same time, many university Institutional Review Boards (IRBs) have followed recent US federal guidelines and now require researchers to be more transparent with survey participants regarding what will happen to the collected data. In this paper, we take the first steps toward considering the interaction between these two survey research developments. Using a nationally representative panel, we show that informing survey participants that their de-identified data will be publicly shared by a researcher can affect how these participants answer certain questions. In some cases, public posting notifications can increase data quality (e.g., knowledge measures), but in other cases informing participants of the data's future use can exacerbate social desirability issues (e.g., turnout). Our results suggest conditional costs and benefits to the intersection between two critical ethical norms underlying survey research: data-sharing and informed consent.

King, Keohane, and Verba (1994, 8), quoting Merton (1949), argue that a project is not scientific unless the researchers' data and methods are public. Making data public increases the chances of re-analysis, which can detect the

Elizabeth C. Connors is a PhD candidate in the Department of Political Science at Stony Brook University, Stony Brook, NY, USA. Yanna Krupnikov and John Barry Ryan are associate professors of political science at Stony Brook University, Stony Brook, NY, USA. The authors would like to thank Arthur Lupia as well as seminar participants at the Massachusetts Institute of Technology, Stony Brook University, and Yale University. Previous versions of this paper were presented at the 2016 Annual Meeting of the Midwest Political Science Association and the 2017 NYU Center for Experimental Social Science Experimental Political Science Conference. This research was conducted with a grant from Time-Sharing Experiments for the Social Sciences (Y.K., PI). *Address correspondence to John Barry Ryan, Stony Brook University, Political Science Department, N711 Social and Behavioral Sciences, Stony Brook, NY 11794, USA; email: john.ryan@stonybrook.edu.

doi:10.1093/poq/nfz013 Advance Access publication June 18, 2019

robustness of results and even identify cases of outright data fraud. Across a variety of disciplines, academic journals are institutionalizing this norm with replication data policies.[1]

Yet, when "data points" are ordinary people choosing to participate in the research process, more than just publication and disciplinary standards are under consideration. Institutional Review Boards (IRBs) govern research involving human subjects and typically require the full, *informed* consent of participants—recent federal guidelines suggest that being informed means that participants must be aware of their data's future use.[2] In turn, researchers who use human subjects and want to publish their research may be increasingly required to inform their participants that their survey responses may be publicly posted in some form.

The norms of both data-sharing and informed consent are pivotal to the ethics of survey research. In this manuscript, we investigate the effects of these two norms on research participants and consider the possibility that the norms' intersection can influence people's survey responses. Through a series of tests, we take the first step toward considering what happens when participants are told their *de-identified* survey responses will be publicly shared with other scholars. We demonstrate that informing participants of this can, under certain conditions, lead people to change their answers to certain questions.

While our research design rests on a simple manipulation—whether participants are informed prior to participation that their de-identified responses will be part of a publicly shared data set—our approach uses the concept of self-monitoring to consider whether respondents who are more likely to engage in socially desirable responding are also more sensitive to data-sharing notifications. Across a variety of tests, our results suggest a conditional intersection between the norms of data transparency and fully informed participants.

This manuscript builds on prior work on ethics and privacy in the informed consent process (Singer 1993, 2003, 2011). At a time when scholars across disciplines have not only called for greater research transparency but have also begun advocating for systematic empirical investigations of the effects of these new transparency norms (Vasilevsky et al. 2013; Lee and Moher 2017), our study offers an early empirical consideration of the relationship between data-posting information in consent forms and respondent behavior. The results suggest that even practices that most researchers view as a net positive for the scientific process may have implications for measurement error and overall data quality.

1. For example: biomedical sciences (Vasilevsky et al. 2013), economics (Galiani, Gertler, and Romero 2017), political science (Lupia and Ellman 2014), psychology (Nelson, Simmons, and Simonsohn 2018), and sociology (Freese and Peterson 2017).

2. IRBs interpret federal guidelines on their own, causing variation in notification requirements. Currently, the Carnegie Foundation classifies 115 institutions as "very high research activity" (formerly known as R1). Of the 85 that had posted consent form guidelines on their IRB websites in 2016, 46 percent required researchers to include the future use of data in consent forms.

Public Posting and Survey Response: Two Perspectives

Research on survey response suggests two broad perspectives on the effect of informing survey participants of public posting. In what follows, we consider why information about data-sharing may be both unlikely and likely to affect survey response.

One perspective suggests that information embedded in a consent form is unlikely to have much influence on individual responding. For starters, individuals often pay little attention when participating in surveys (Krosnick 1991; Berinsky, Margolis, and Sances 2016). If respondents pay little attention to questions, they may be even *less* likely to pay attention to consent forms. This means data-sharing notices are unlikely to influence participants because participants are unlikely to have even read them.

A related argument suggests that even if participants *do* read data-sharing information, they are unlikely to be concerned about it. Surveys often attempt to maintain anonymity, and web surveys may create an even greater sense of anonymity by minimizing human interaction and emphasizing de-identification (Tourangeau 2004; though some research suggests participants often do not believe anonymity claims, e.g., Jensen, Potts, and Jensen 2005; Joinson, Woodley, and Reips 2007; Mueller et al. 2014). Even if people do closely consider all information provided by the survey, the fact that they do not typically give their names in surveys should diminish the effect of the data-sharing cue.

Another approach suggests that even small changes in survey context can affect response. Research on survey response has recognized that respondents can and often do misrepresent their beliefs (Zaller and Feldman 1992). These misrepresentations are often a function of what Berinsky (2004) calls "self-presentation concerns"—people give answers they hope will improve their appearance, rather than answers that speak to their attitudes and intentions. Broadly defined, these tendencies lead people to select answers that are "socially desirable," although beliefs about what is socially desirable can vary across individuals (Klar and Krupnikov 2016).

Certain survey contexts exacerbate these self-presentation tendencies (Schuman, Presser, and Ludwig 1981; Sudman, Bradburn, and Schwarz 1996), meaning that people not only misrepresent their positions due to social desirability perceptions, but that survey-specific factors can further influence participants' truthfulness. In particular, people are especially likely to respond in a socially desirable manner when reminded that surveys are, essentially, social interactions similar to conversations (Berinsky 2004). Reminders that someone will be interpreting responses can lead people to become particularly cautious with closed-ended questions, as choosing a single response is their only chance to send a signal about their personal qualities (Krupnikov, Piston, and Bauer 2016).

A number of factors can remind people there is a person on the "other side" of the survey. Certainly, explicit cues reminding people that a researcher is "watching" encourage socially desirable responses (Haley and Fessler 2005), but even less direct interventions can have the same effect. Clifford and Jerit (2015), for example, find that telling respondents they will receive feedback at the end of a survey can inadvertently increase socially desirable responding. Moreover, these types of survey cues can influence responses even in "conditions of complete anonymity" (Rigdon et al. 2009, 359). In sum, people misrepresent their opinions when reminded that the ultimate purpose of a survey is to evaluate public opinion.

Following these arguments, notifying people that their de-identified data will be publicly posted may serve as another cue reminding participants of the "other side" of the survey: there is not only the individual researcher, but *also* an entire research community. Hence, the public nature of survey responses is highlighted, encouraging respondents to answer in a "public" manner, rather than to simply reveal a preference.

SELF-MONITORING

If information about the public posting of de-identified data exacerbates respondents' awareness of the social components of the survey, some people are more likely to be sensitive to these social pressures than others (Paulhus 1991; Berinsky 2004). In turn, these differences in sensitivities will lead to conditional behavioral responses to the public posting notification (Paulhus et al. 2003; Berinsky and Lavine 2012). In other words, if information about the posting of de-identified data affects the extent to which people are reminded that "someone is watching," people who are more sensitive to these types of social cues should be more likely to engage in socially desirable responding than those who are less sensitive (Gangestead and Snyder 2000; Berinsky 2004; Berinsky and Lavine 2012).

A factor that captures the extent to which people are likely to change their behaviors in response to social cues is *self-monitoring* (Snyder 1974, 1979; Gangestead and Snyder 2000; Berinsky and Lavine 2007, 2012). Rooted in theories of impression management, self-monitoring follows from the idea that most people have some capacity to adjust their behaviors to fit specific social contexts (Goffman 1955), but proposes that some people are more likely than others to do so (Snyder 1974, 1979).

While high self-monitors are often more likely to engage in socially desirable responding in surveys (Weber et al. 2014), what differentiates self-monitoring—both as a concept and, importantly, as a measurement scale—is its focus on impression management. While other approaches capture socially desirable behavior as a function of both self-deception *and* impression management (Paulhus 1991), self-monitoring focuses on the extent to which people aim to engage in "self-presentations designed to impress an audience" (Paulhus et al. 2003, 899). Given our theoretic argument that information about data-posting could affect the social context of the survey, this approach is beneficial. Moreover, self-monitoring has been tested and validated for the types of survey measures we include in our empirical analysis (e.g., Berinsky 2004 directly tests self-monitoring against the Paulhus 1991 impression management scale).

If information about the posting of de-identified data affects the extent to which people perceive their responses to be "public," then we may expect differential responses to this information from high and low self-monitors. Since those higher in self-monitoring are more willing to change their survey answers (and even real-life behaviors) if they perceive that doing so will make a better impression on others (Snyder 1974; Gangestead and Snyder 2000), we would anticipate that high self-monitors will be more likely to respond to information about data-posting by shifting to responding in a manner they perceive to be socially desirable, or, at the very least, in ways that aim to avoid antipathy from others.

In contrast, low self-monitors are guided by the goal of "self-verification" (Banaji and Prentice 1994): the need to ensure that others see their "authentic selves" (Premeaux and Bedeian 2003). As Snyder describes, the key question for low self-monitors is "Who am I and how can I be me in this situation?" (1979, 103). While previous research has interpreted this to mean that low self-monitors will be more consistent in their responses to survey questions regardless of context changes (Weber et al. 2014), this is not necessarily the case (DeBono 1987; Day and Schleicher 2006). Driven by self-verification goals (Banaji and Prentice 1994), low self-monitors adjust their behaviors to make their positions clearer to others, even if these will not be impressive (DeBono 1987; Day and Schleicher 2006). As a result, low self-monitors may appear *less likely* to engage in socially desirable responding if they are told the data will be public.

In sum, if information about the posting of de-identified data is altering the perceived social context of the survey—and the extent to which people perceive it to be "public"—we should observe behavioral differences between high and low self-monitors. While high self-monitors may be more likely to turn to socially desirable responding, low self-monitors may become less likely to engage in this type of behavior.

Methods

RESPONDENTS

We begin with the idea that research on human subjects rests on two types of transparency. First, researchers must be transparent with the scientific community about the data used, and posting data is part of this transparency. Second, researchers must inform participants about, among other things, the possibility of the public posting. These modes of transparency are required to meet both ethical and scientific guidelines. The goal of this paper's studies is to ask whether the joint presence of both modes of transparency can affect how people answer survey questions.

Our main analysis uses a survey experiment conducted with the GfK panel—a nationally representative survey sample (*N* = 767, fielded December 2015 to January 2016). We also conducted a study on Amazon's Mechanical Turk (MTurk; *N* = 1,546). We rely on MTurk to consider the generalizability of our results to another subject population (McDermott 2011), especially one researchers frequently use. In the manuscript, we focus on the GfK results, but we present all the MTurk studies in the Supplementary Online Materials, Appendix H.[3]

Many consider the GfK sample to be "the gold standard" because GfK uses random-digit-dialing techniques, callback strategies, and incentives to initiate and sustain contact with a nationally representative internet panel of American adults.[4] While GfK participants are sampled to participate in specific researcher-commissioned surveys,[5] GfK also fields its own questionnaires in order to maintain a frequently updated "profile" on panel members. This is beneficial for our study, as three of the measures we ask (household income level, media attention, and turnout in the 2014 midterm elections) were previously asked of our participants as part of their GfK profile. This allows us to see if respondents are *changing* their answers after the public posting notification. That GfK participants are part of a panel does create some limitations. These participants may be more familiar with consent forms and may be more likely to participate in any given survey due to their commitment to GfK. Thus, while we can consider changes in response behavior due to data-sharing information, we cannot track the effect of this information on the initial willingness to participate in a survey.

As Time-Sharing Experiments for the Social Sciences (TESS) funded the GfK study, we include our original TESS application in Appendix I so that readers can view our study questions, pre-data collection predictions, and analysis plans.

DESIGN

The basic design features are shown in figure 1; this design approach builds on prior research on informed consent procedures (Singer 1978). First, all

3. All appendices are contained in the accompanying Supplementary Online Materials.
4. Per GfK documentation, as of 2014, 40 percent of the panel was recruited via Random Digit Dialing and 60 percent via Address Based Sampling. Recruited households who have no internet access are provided with free internet service by GfK to ensure panel representativeness.
5. Participants for each specific survey are selected via a methodology patented by GfK for using "multiple sequential samples from a finite panel membership" (GfK methodology document 2012, 3). The survey's completion rate was 64.2 percent. The cumulative response rate is 5.4 percent, considering panel recruitment and retention (Callegaro and DiSogra 2008).

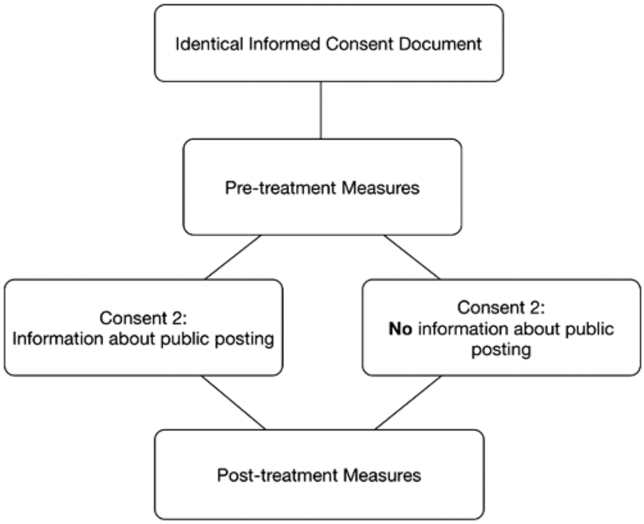

Figure 1. Basic experimental design of both studies.

subjects are presented with a basic consent form, which *does not* mention anything about the data post-collection. Respondents then complete a self-monitoring scale (Berinsky and Lavine 2012), a news interest question, and one question about internet privacy.

After these initial questions, our subjects are presented with the experimental manipulation: we present subjects with a second "consent form." In the control groups, subjects are told, "Next we will ask you for your opinions on a series of political questions. Do you wish to continue with this study?"

There are two statements that subjects receive in the treatment groups. Some subjects are randomly assigned to the following prompt: "Next we will ask you for your opinions on a series of political questions. Please note that if this research is published, your responses will be made publicly available to other researchers. You will never be asked for your name. Only the responses you give to the questions that follow will be available for download. Do you wish to continue with this study?" Other subjects are given a reason for the public posting. After noting that the information will be made public, we insert the line, "This will be done so that other scientists can reduce risk of scientific fraud."[6] All of our subjects chose to continue.

It is important to note that subjects in the treatment groups are reminded that even if their responses are posted, they remain de-identified. Also, providing

6. During debriefing, control group subjects are given information about public posting.

a reason for the public posting has no discernible effect, so we combine the public posting treatment groups in our analyses.[7]

MEASURES

Our outcome measures (listed in table 1) are a series of questions that previous research has demonstrated are likely to have different levels of potential for socially desirable responding. The dependent variables include post-treatment measures that scholarship considers less likely to produce socially desirable responding, such as political knowledge (Fiske, Lau, and Smith 1990; Zaller 1992; Mondak 2000); and those that are considered more sensitive and thus more likely to produce socially desirable responding—abortion (Singer and Couper 2014; Liu 2018) and racial attitudes (Berinsky 2004; Clifford and Jerit 2015).

We also present measures considered likely to produce socially desirable responding that subjects answered at two points in time: income (Yan, Curtin, and Jans 2010) and turnout (Belli, Traugott, and Beckmann 2001). Using these variables, we can compare respondents' answers in previous iterations of the GfK panel to post-treatment responses. We note that in the case of the turnout variable, in contrast to the GfK turnout measure, our turnout measure typically reduces overreporting (Duff et al. 2007). While the questions differ, random assignment to treatment allows us to rule out differences due to question wording.[8]

Our key independent variable (see table 1) is our experimental treatment: *Public Posting*. Our other main independent variable is the self-monitoring scale adapted by Berinsky and Lavine (2007, 2012), where larger values indicate higher propensity toward self-monitoring. The variable is positively skewed with a mean of 3.7. To ensure that subjects very high on the self-monitoring scale do not have undue influence on the results, we construct a logged self-monitoring variable.[9]

In some of our analyses, we include an interaction between the treatment dummy variable and logged self-monitoring, following our theoretic argument that if information about posting leads individuals to perceive the data as more

7. We note that we only tested one kind of reason. Fobia and her colleagues (2019) report that type of reason matters for consent to data linkage with less favorable attitude toward consent when the benefit was an increase in knowledge, but more favorable when framed in terms of better use of taxpayer money. Future research on the effect of transparency disclosures should examine multiple frames.

8. Due to question-wording changes, we also consider the post-treatment measure alone (Appendix C).

9. Cronbach's $\alpha = 0.72$, which compares favorably with Berinsky and Lavine (2007), who report $\alpha = 0.6$. We also consider the reliability of the scale using a parallel analysis scree plot, which demonstrates that the scale has a single dimension. See full results as well as a discussion of how the scale was constructed in Appendix A.3.1.

Table 1. Summary of measures

Variables	Question structure	Variable coding	Wording source
		Outcome variables, measured post-treatment only	
Knowledge	Three objective knowledge items. Full item wording Appendix A2.1	Count of correct answers scaled 0 = lowest to 1 = highest	ANES
Abortion	Four items Full item wording Appendix A2.2	Index scaled 0 = most anti to 1 = most pro	GSS, see Alvarez and Brehm (1995)
Racial resentment	Four items Full item wording Appendix A2.3	Index scaled 0 = lowest to 1 = highest	Sears et al. (1997)
		Outcome variables, measured pre- and post-treatment only	
Income	Single item Full question wording for both pre- and post- in Appendix A2.4	Scaled 0 = lowest to 1 = highest	Pre: GfK panel; Post: ANES
Turnout	Single item Full question wording for both pre- and post- in Appendix A2.5	0 = did not vote, 1 = voted	Pre: GfK panel; Post: ANES
		Main independent variables	
Public posting	Randomly assigned experimental treatment about data-posting Full wording in Appendix A.1	0 = no posting info. 1 = posting info.	—
Self-monitoring	Three items Full item wording in Appendix A.3	Items added, higher values = higher self-monitoring, variable logged	Berinsky and Lavine (2007)

Continued

Table 1 (*continued*)

Variables	Question structure	Variable coding	Wording source
		Control variables	
Party ID	Two branched questions		
	Full item wording in Appendix A.4	7-point scale, 7 = most Dem.	GfK Panel
Education	Single item		
	Full item wording in Appendix A.4	0 = lowest to 1 = highest	GfK Panel
Gender	Single item		
	Full item wording in Appendix A.4	0 = woman, 1 = man	GfK Panel
Race/Ethnicity	Single item		
	Full item wording in Appendix A.4	Three dummy variables: African American, Hispanic, White	GfK Panel

public, people who differ in their sensitivities to changes in social context—that is, high and low self-monitors—should also vary in their response to the treatment.

In the statistical models, we include control variables (following Kam and Trussler 2017). These were asked of respondents as part of their GfK panel participation at a time point separate from our study. Full summary statistics and question wording for measures are shown in Appendix A; measures for news attention are in Appendix D.

Between-Subjects Results

MAIN EFFECT OF THE EXPERIMENTAL TREATMENT

Although our theoretical approach suggests an interaction between self-monitoring and the treatment, we first present the bivariate results. We adjust our *p*-values for multiple comparisons because we examine treatment effects on multiple dependent variables within the same data sets. We use the method Benjamini and Hochberg (1995) describe, which controls the "false discovery rate" (FDR)—the expected proportion of false positives (Type I errors).

The results are in table 2. We see statistically significant differences at the 0.05 level (two-tailed tests) in both abortion and turnout, and at the 0.10 level (two-tailed tests) in knowledge. That means there is evidence in most of these dependent variables that the public posting notification can lead to different responses,

Table 2. Main effect of experimental treatment in GfK samples, bivariate analyses

	Control		Treatment		Benjamini–Hochberg
Variable	Mean	Std. err.	Mean	Std. err.	*p*-value
Knowledge (0, low–1, high)	0.502	0.032	0.428	0.022	0.082
Abortion (0, most anti–1, most pro)	0.395	0.03	0.494	0.021	0.021
Racial resentment (0, low–1, high)	0.629	0.02	0.651	0.013	0.379
Income (0, low–1, high)	0.581	0.017	0.562	0.013	0.379
Turnout (0, vote or 1, no vote)	0.739	0.028	0.614	0.022	0.004

NOTE.—*P*-values are for two-tailed tests. Abortion: $n = 748$; Racial Resentment: $n = 571$ (white respondents only); Knowledge: $n = 760$; Income: $n = 764$; Turnout: $n = 764$.

though the directionality of these differences is not systematic. We also want to note that we conduct similar bivariate analyses for racial resentment and abortion attitudes using the MTurk data and find null results (see Appendix H).

Our theoretic arguments suggest, however, that if the treatment is affecting perceptions of the social context of the survey, high and low self-monitors should respond differently. In particular, we should observe more socially desirable responding among high self-monitors and less among low self-monitors. This directionality is something that the results in table 2 cannot capture, and thus we turn to a set of analyses that rely on an interaction between the treatment and self-monitoring.

POLITICAL KNOWLEDGE

Self-monitoring should influence the way in which the public posting notification affects survey responses. In this section, we present the results of models with *Posting Treatment*, *Logged Self-Monitoring*, and an interaction between these two as the key independent variables. Because self-monitoring is not randomly assigned, we include control variables to avoid omitted variable bias with respect to the effect of self-monitoring.

The first models examine the GfK political knowledge questions. Across both conditions, 45.3 percent of respondents answered all three questions correctly.[10] For all three questions, the majority of respondents who failed to provide a correct answer replied, "I don't know." As a result, we conduct logit models with two different dependent variables: (1) the dependent variable as 1 if the respondent answered all questions correctly and 0 if the respondent failed to answer any question correctly; (2) the dependent variable as 1 if the respondent answered "I don't know" to any question and 0 if the respondent provided answers (correct or incorrect) to all questions.[11]

The full logit models are in Appendix E. Given the nature of our models, limited dependent variables with interaction effects, we present marginal effects plots to demonstrate the effects of the *Public Posting Treatment*. As Berry, DeMeritt, and Esarey (2010) note, a statistically significant product term is neither necessary nor sufficient for variables to interact in how they affect probabilities in these models. It is necessary to plot the marginal effects to test hypotheses related to the outcome probability.

In figure 2, we present the posting treatment's marginal effect for various levels of self-monitoring. Three things are important to notice. First, the treatment notification only affects those *low* in self-monitoring. Second, the treatment causes those lowest in self-monitoring to say "don't know" more

10. This may seem high; however, this compares to the 50 percent of respondents who answered three different knowledge questions correctly in the Indianapolis–St. Louis Study, a widely used dataset for public opinion in social networks (e.g., Huckfeldt, Sprague, and Levin 2000; Ryan 2011; Djupe, McClurg, and Sokhey 2018).

11. We also consider success rates on individual questions; see Appendix B.

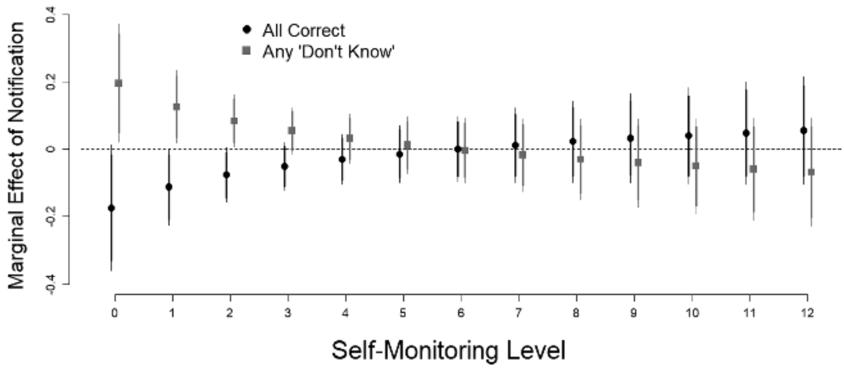

Figure 2. Marginal effect of posting treatment on objective knowledge measures by levels of self-monitoring. Marginal effects calculated based on logit models available in Appendix E.1. Thick lines represent 90 percent confidence intervals; thin lines represent 95 percent confidence intervals.

frequently (at least one answer more). Finally, as a result, low self-monitors are less likely to answer all questions correctly in the treatment condition than in the control condition.

These results follow from previous research. First, political science research commonly considers how high self-monitors are more likely to behave in a socially desirable manner (e.g., Weber et al. 2014), yet studies suggest that in certain contexts low self-monitors will more openly state their beliefs (Day and Schleicher 2006). In this case, the public posting notification may cause low self-monitors to simply state "don't know" instead of guessing—even if that guess is likely correct. Second, self-monitoring was positively correlated with education, meaning that high self-monitors were more likely to know the answers and thus to not be worried about public posting (see Appendix E).

As a consequence of these two factors, the public posting notification influences those who did not attend college. In the control group, 43 percent of non-college graduates answered all three questions correctly, whereas in the *Posting Treatment* group, only 32 percent answered all three questions correctly (difference is $p = 0.02$).[12] Depending on one's concept of political knowledge, the public posting could prompt more accurate measurement. If the researcher only wants respondents to answer the question when they truly "know," then the increased number of "don't know" responses is meaningful (Luskin and Bullock 2011). If, on the other hand, the researcher wants to see if the respondent can guess the correct answer, then the public posting notification likely underestimates the levels of political knowledge (Mondak and Davis 2001).

12. In terms of the number of questions answered correctly, the difference is between 2.19 correct in the control group and 1.95 correct in the treatment group ($p < 0.01$).

RACIAL RESENTMENT AND ABORTION ATTITUDES

The next step is to determine if self-monitoring conditions the posting treatment effect on more sensitive issues: racial resentment and abortion. We construct our racial resentment measure from respondents' mean score on the four items, rescaled to 0–1, with larger values indicating greater racial resentment. Because the variable is bound at 0 and 1, we run a Tobit model and include the same previously used control variables (except respondent's race—we only include white respondents, per common practice; see Piston 2010; Banks and Valentino 2012).

In the abortion model, the dependent variable is the number of respondents' pro-choice responses, which range from 0 to 4 (with 47 percent of respondents at 0 and 40 percent at 4). We estimate a negative-binomial regression model because our dependent variable is an over-dispersed count of pro-choice responses. We add the same control variables to the model.

The full abortion and racial resentment models are in Appendix E. The two panels in figure 3 plot the posting treatment's marginal effect for different levels of self-monitoring. The racial resentment results are on the left, with the abortion results on the right. Again, we see that self-monitoring influences results, but this differs depending on the issue. Low self-monitors are influenced in the abortion model, while high self-monitors are influenced in the racial resentment model.

What causes these differences? We next consider these variables another way. For racial resentment, we run Tobit models separately for each question

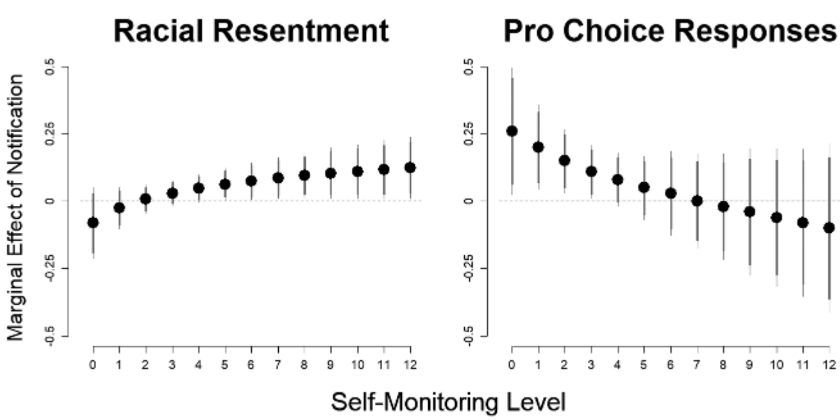

Figure 3. Self-monitoring conditions the effect of the posting treatment for abortion and racial resentment. Racial resentment marginal effects calculated based on Tobit model: available in Appendix E.2. Abortion marginal effects calculated based on negative binomial model: available in Appendix E.3. Thick lines represent 90 percent confidence intervals; thin lines represent 95 percent confidence intervals.

(available in Appendix E2.1). These show that the largest effects occur with two items: "It is really a matter of some people not trying hard enough; if blacks would only try harder they could be just as well off as whites," and "Irish, Italian, Jewish, and many other minorities overcame prejudice and worked their way up. Blacks should do the same without any special favors." Some have suggested these two questions are double-barreled because they measure both racial attitudes and attitudes toward individualism (Sniderman and Tetlock 1986; Feldman and Huddy 2005). When respondents answer these questions in the treatment group, they are less likely to strongly disagree with the first question and more likely to neutrally respond to the second question. This make sense if respondents did not want to appear racially prejudiced, but also did not want to reject the commonly held individualism value. As a result, average racial resentment levels increase in the treatment group even though high self-monitors attempt to not appear racist in both groups.

With abortion, we see an effect similar to the knowledge questions: low self-monitors are more influenced by the treatment, increasing their pro-choice responses. This, again, suggests that rather than hiding responses, these respondents are motivated to state their (potentially unpopular) opinions. Republicans, specifically, should be most likely to believe they should hide pro-choice attitudes (Adams 1997). Therefore, if low self-monitors are revealing unpopular pro-choice opinions, then the changes should occur among Republicans. Indeed, this is what we find. Democrats offer an average of 2.5 pro-choice responses—regardless of treatment—whereas Republicans offer an average of 0.76 in the control group and about 1.3 in the treatment group ($p < 0.01$). Again, perhaps this reflects increased honesty among Republicans with public posting.

The results in this section suggest that public posting can influence responses to a variety of measures, but that self-monitoring moderates this effect. Further, the results suggest that the measures of these concepts for certain subgroups change when respondents are informed their data will be publicly posted. It is unclear, however, if this leads to biased or more accurate responses—in the case of racial resentment, we likely see biased responses, but for knowledge and abortion we may see more accurate responses.

Respondents Changing Their Answers

INCOME

To this point we have used between-subjects analysis to suggest that respondents change their answers to some survey questions in response to a public posting notification. Yet, comparing subjects across treatment groups does not show us *movement*, necessarily. In this section, we rely on the information GfK collects, retains, and updates as part of their panel members' profile to see if treatment subjects were more likely to change responses.

Our post-treatment question is: "In 2014, what was your total family income (please approximate if you are not certain)?" with nine response categories and "don't know."[13] Since income is in the demographic profile GfK maintains about panel members, we can compare our post-treatment measure to the information GfK obtained from our participants prior to our study. GfK's profile information includes data imputed when missing, and presents respondent household income in 19 categories. We collapsed these to match our nine categories and then constructed a three-category dependent variable: (–1) reported less income post-treatment; (0) no change; (1) reported more income post-treatment. Of course, given that time passed between GfK's collection (or imputation) of the profile income measure and our post-treatment measure, it is possible that respondents' incomes changed. However, we have no reason to believe these changes would differ by treatment. That is, *even if* income levels did change, the treatment effect should be uninfluenced.[14]

Based on the results in table 3, it appears the posting treatment had no effect. Just under half of respondents changed their answers, and about three-quarters of those were to report less income.[15] As a next step, we consider self-monitoring.

In figure 4, we can again see that self-monitoring plays a role. Figure 4 presents the treatment's marginal effects using a multinomial logit that includes an interaction between the treatment and logged self-monitoring variable with controls of the previous models and for initial reported income.[16]

Table 3. The effect of a public posting notification on reported income

	Control	Treatment
Change to less income	34.6%	31.3%
No change	55.7%	56.9%
Change to more income	9.7%	11.8%
N	231	461
$\chi^2 = 1.17$; $p = 0.56$		

13. Respondents who said "don't know" to this post-treatment measure were set to their education level's mean income.

14. It is possible that GfK's imputation was inexact and our post-treatment measure is better. First, this is something that categorizing the 19-item GfK measure should correct (i.e., the imputation may have been less exact with 19 categories but should do better with nine). Second, there is no reason to believe that the relationship between imputation and the post-treatment measure should be affected by treatment.

15. It is possible that changes tended toward reporting less income because of changes in response options, but given the location of the midpoint on the two scales, research suggests this would be unlikely (Schwarz et al. 1985).

16. The full multinomial logit model is in Appendix E.4 and includes the same control variables as before.

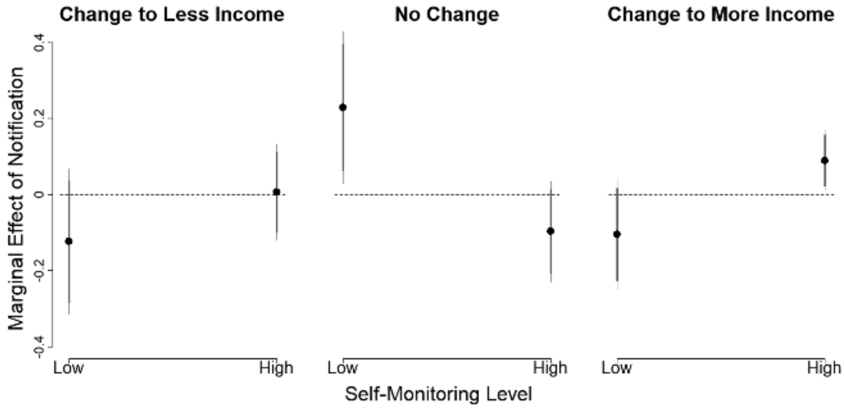

Figure 4. How self-monitoring affects changes to reported income. Marginal effects calculated based on multinomial-logit models available in Appendix E.4. Low self-monitors are the 5th percentile of self-monitoring (0 out of 12) while high self-monitors are the 95th percentile of self-monitoring (8 out of 12). Thick lines represent 90 percent confidence intervals; thin lines represent 95 percent confidence intervals.

Importantly, we see differing treatment effects by self-monitoring levels. The treatment increases the probability that low self-monitors will not change their reported income. At the same time, though, the treatment raises the probability that high self-monitors will report *more* income. We do not know the respondent's actual household income, but the patterns suggest that low self-monitors provide more accurate answers—they typically report the same income at both time points, with enough time between them to make the recall of a lie unlikely. High self-monitors, on the other hand, consistent with their goal of making the best impression, typically report more income when given the public posting treatment (see also Hall et al. 2010). Again, the public posting notification produces different qualities of data depending on whether the respondent is a low or high self-monitor.

REPORTED TURNOUT

As we noted, it is possible that an individual's actual income changed between the GfK panel and our survey. On the other hand, whether or not respondents actually voted in a particular election that took place in the past does not change regardless of when subjects are asked.

As part of GfK panel participation, all subjects were asked, "Did you happen to vote in the 2014 elections for the U.S. Congress?" with response options "Yes," "No," and "Not Sure." In our study, subjects were asked, "In asking people about the 2014 election, surveys often find that a lot of people were not able to vote because they weren't registered, they were sick, or they just didn't have time. Which of the following statements best describes you?" They could respond, "I did not vote (in the election in November 2014)," "I thought about voting this time, but didn't," "I usually vote, but didn't this time," "I am sure I voted," or "Don't know."

We use these two questions to create a three-category variable, coded 1 if respondents change from either "No" or "Not sure" in the GfK panel to "I am sure I voted" in our post-treatment measure; –1 if they switched from saying they did vote to saying they did *not* vote; and 0 otherwise. Hence, the measure is coded 1 if they moved toward the socially desirable response, 0 if they did not change, and –1 if they changed in a non-socially desirable direction (Karp and Brockington 2005).

Our post-treatment question differs from the original question, and this provides a tougher test for our argument. We are examining if people overreport turnout in response to the public posting notification, and the question we used was specifically designed to mitigate socially desirable overreporting (Duff et al. 2007). Hence, if we find the treatment induces overreporting, it is possible this would have been even greater if we had used the question wording from the GfK panel. While not all errors in voter self-reports are due to deliberate overreporting, even people who initially cannot remember whether they voted are often "motivated to infer that they did through source monitoring processes that are influenced by social desirability or self-presentation concerns" (Belli, Traugott, and Beckmann 2001, 495).[17]

Again, it does not appear that the treatment has any effect when we look at the bivariate results in table 4. We see less people changing than we observed with income—though, that is to be expected given that turnout in a specific

Table 4. The effect of a public posting notification on reported turnout

	Control	Treatment
Change to not voting	5.2%	6.9%
No change	84.4%	80.5%
Change to voting	10.4%	12.6%
N	231	461
$\chi^2 = 1.65$; $p = 0.44$		

17. Further, for memory issues to explain our data patterns, at least one of two conditions would have to hold. First, the treatment would have to affect individual abilities to remember whether they voted. Second, there would have to be a relationship between self-monitoring and ability to remember. At this time, there is no empirical research to suggest either of these relationships.

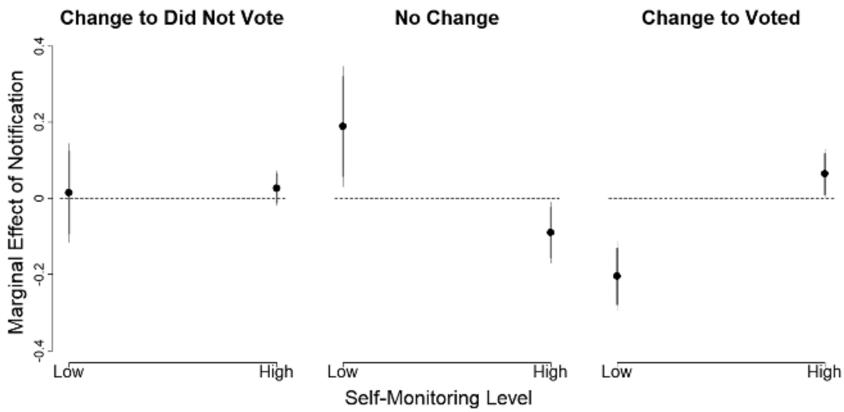

Figure 5. How self-monitoring affects changes to reported turnout. Marginal effects calculated based on multinomial-logit models available in Appendix E.6. Low self-monitors are the 5th percentile of self-monitoring (0 out of 12) while high self-monitors are the 95th percentile of self-monitoring (8 out of 12). Thick lines represent 90 percent confidence intervals; thin lines represent 95 percent confidence intervals.

election *cannot* change. When there are changes, it is no surprise that they generally are in the socially desirable direction.

Once again, self-monitoring conditions the effect of the public posting notification, as figure 5 shows. We again run a multinomial logit with the interaction between the treatment and self-monitoring and include the same previously used controls and their panel turnout response. The full model is in Appendix E.6.

As with the income results, low self-monitors are less likely to change their reported turnout when they are given the public posting notification. They are also less likely to change their response to state that they voted. Again, low self-monitors are more likely to stick to their original response, even if that response is not socially desirable.

This is not the case with high self-monitors. When told their data will be posted, high self-monitors are more likely to change their response to the socially desirable response. This result, while consistent with theory, is still somewhat surprising given that many high self-monitors likely gave socially desirable responses the first time they were asked.[18] The first time they were asked, 66 percent of respondents reported voting in the 2014 midterm election—a notable election, as it marked the lowest percentage turnout since

18. Even though some control variables are associated with overreporting turnout (Belli, Traugott, and Beckmann 2001; Duff et al. 2007), only one is statistically significant. This is likely because these variables resulted in overreporting in the first wave but did not result in additional overreporting in wave two.

World War II.[19] As with income, finding out that their information may be publicly posted heightens reputational concerns among high self-monitors.

We unfortunately do not have validated turnout data for our respondents. Berent, Krosnick, and Lupia (2016) show, however, that "validated" turnout estimates are not necessarily superior to self-reports because there are errors in matching names to government records.[20] More importantly, though, our goal in this study is to see how public posting affects common measures, including self-reported turnout. Unlike some of the previous analyses, this pattern of results suggests that data gathered with the public posting treatment is of lower quality, as it heightens overreports due to social desirability.

Conclusion

Many journals, as part of an effort to increase transparency, are requiring authors to post their data. At the same time, new federal guidelines, and in turn university IRBs, require increased transparency with human subjects about what will happen to their (de-identified) data. In this manuscript, we show that the combination of these two important developments can have implications for survey research. Informing subjects their de-identified data will become part of a publicly posted data set may cause some respondents to change survey responses. Our results signal a warning to survey researchers to be mindful that increasing transparency on both ends of the research process can affect data. The results do *not* mean, however, that researchers should be *less* transparent, with either other scholars or their subjects.

We want to note several things about these results. First, one might make the (ethically dubious) counterpoint that typical participants do not thoroughly read consent forms. This relies on an outdated view of consent forms. New federal guidelines require consent forms that are more readable, with important information—like data-sharing—clearly marked. Because of this, it is difficult for researchers to hide the public posting notification even if they wanted to take this unethical step. Further, one of the authors asked participants in a recent study on MTurk if they knew what would happen to the data if the research was published. The public posting notification was in the middle of the study's consent form. The majority of participants knew the data would be publicly posted.

Second, notifying respondents of public posting did not always affect their answers, as table 2 suggests. Even in situations where we observed a

19. For our study, the reported turnout is 71 percent, while the turnout rate in 2014 was around 37 percent (United States Election Project 2015). This level of overreporting is consistent with studies like the Cooperative Congressional Election Study (aggregate reported turnout rate: 79.5 percent).

20. Berent, Krosnick, and Lupia (2016) say validated measures *appear* superior because aggregate turnout levels are lower, but that is only because matching errors cause researchers to underestimate registration levels.

statistically significant result, models that included self-monitoring demonstrate that the notification does not affect all respondents equally.

Third, our results suggest that the data one gets from a survey with a public posting notification may be superior in some cases, as the notification might encourage low self-monitors to answer truthfully. We caution, however, that improved data quality could make comparisons with data from earlier time periods more difficult. Fixing biased data is obviously useful, but researchers should be aware of changes in consent procedures before drawing conclusions about differences in attitudes at various time periods.

Fourth, even if data is biased by public posting—as is possibly the case with turnout—one can correct the bias. In this case, we have shown that self-monitoring moderates the effect of the public posting notification.

Fifth, while we present our MTurk studies in Appendix H due to space constraints, we want to underscore that the MTurk results present a series of null findings. We believe this is important to contextualizing our findings and the role pre-treatment may play.

Next, there are methods available for dealing with social desirability that can be used in these cases as well. The results of a list experiment we conduct on our MTurk sample suggest this method's potential; other methods for reducing socially desirable responding exist, and more will likely be developed.

Finally, we want to reinforce that this is only the first attempt to use data to consider the potential tension between informed consent and survey response. Indeed, our approach has some limitations. Critically, our participants were part of a panel that may make them more likely to agree to participate; it is possible, for example, that data about public posting may have a different effect for "fresh" participants. Moreover, our results rely on specific types of information about data-sharing and respondent privacy. We note these limitations to suggest that more research is necessary to fully discern the way data-sharing notices affect individual behavior in surveys.

Supplementary Data

Supplementary data are freely available at *Public Opinion Quarterly* online.

References

Adams, Greg D. 1997. "Abortion: Evidence of an Issue Evolution." *American Journal of Political Science* 41:718–37.

Alvarez, R. Michael, and John Brehm. 1995. "American Ambivalence Towards Abortion Policy: Development of a Heteroskedastic Probit Model of Competing Values." *American Journal of Political Science* 39:1055–82.

Banaji, Mahzrin R., and Deborah A. Prentice. 1994. "The Self in Social Contexts." *Annual Review of Psychology* 45:297–332.

Banks, Antoine J., and Nicholas A. Valentino. 2012. "Emotional Substrates of White Racial Attitudes." *American Journal of Political Science* 56:286–97.
Belli, Robert F., Michael W. Traugott, and Matthew N. Beckmann. 2001. "What Leads to Voter Overreports? Contrasts of Overreporters to Validated Voters and Admitted Non-Voters in the American National Election Study." *Journal of Official Statistics* 17:479–98.
Benjamini, Yoav, and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society, Series B (Methodological)* 57:289–300.
Berent, Matthew K., Jon A. Krosnick, and Arthur Lupia. 2016. "Measuring Voter Registration and Turnout in Surveys: Do Official Government Records Yield More Accurate Assessments?" *Public Opinion Quarterly* 80:597–621.
Berinsky, Adam J. 2004. "Can We Talk? Self-Presentation and the Survey Response." *Political Psychology* 25:643–59.
Berinsky, Adam J., and Howard Lavine. 2007. "Self-Monitoring and Political Attitudes: 2006 NES Pilot Study Report." In *NES Pilot Study Reports*, pp. 1–21.
———. 2012. "Self-Monitoring and Political Attitudes." In *Improving Public Opinion Surveys: Interdisciplinary Innovation and the American National Election Studies*, edited by John H. Aldrich and Kathleen M. McGraw, 27–45. Princeton, NJ: Princeton University Press.
Berinsky, Adam J., Michele F. Margolis, and Michael W. Sances. 2016. "Can We Turn Shirkers into Workers?" *Journal of Experimental Social Psychology* 66:20–28.
Berry, William D., Jacqueline H. R. DeMeritt, and Justin Esarey. 2010. "Testing for Interaction in Binary Logit and Probit Models: Is a Product Term Essential?" *American Journal of Political Science* 54:248–66.
Callegaro, Mario, and Charles DiSogra. 2008. "Computing Response Metrics for Online Panels." *Public Opinion Quarterly* 72:1008–32.
Clifford, Scott, and Jennifer Jerit. 2015. "Do Attempts to Improve Respondents Attention Increase Social Desirability Bias?" *Public Opinion Quarterly* 79:790–802.
Day, David V., and Diedre J. Schleicher. 2006. "Self-Monitoring in the Workplace: A Motive-Based Perspective." *Journal of Personality* 74:685–714.
DeBono, Kenneth. 1987. "Investigating the Social-Adjustive and Value-Expressive Functions of Attitudes: Implications for Persuasion Processes." *Journal of Personality and Social Psychology* 52:279–87.
Djupe, Paul A., Scott D. McClurg, and Anand E. Sokhey. 2018. "The Political Consequences of Gender of Social Networks." *British Journal of Political Science* 48:637–58.
Duff, Brian, Michael J. Hanmer, Won-Ho Park, and Ismail K. White. 2007. "Good Excuses: Understand Who Votes with an Improved Turnout Question." *Public Opinion Quarterly* 71:67–90.
Feldman, Stanley, and Leonie Huddy. 2005. "Racial Resentment and White Opposition to Race-Conscious Programs: Principles or Prejudice?" *American Journal of Political Science* 49:168–83.
Fiske, Susan T., Richard Lau, and Richard A. Smith. 1990. "On the Varieties and Utilities of Political Expertise." *Social Cognition* 8:31–48.
Fobia, Aleia Clark, Jessica Holzberg, Casey Eggleston, Jennifer Hunter Childs, Jenny Marlar, and Gerson Morales. 2019. "Attitudes Towards Data Linkage for Evidence-Based Policymaking." *Public Opinion Quarterly* 83:264–79.
Freese, Jeremy, and David Peterson. 2017. "Replication in Social Science." *Annual Review of Sociology* 43:147–65.
Galiani, Sebastian, Paul Gertler, and Mauricio Romero. 2017. "Incentives for Replication in Economics." *NBER Working Paper 23576*.
Gangestead, Steven W., and Mark Snyder. 2000. "Self-Monitoring: Appraisal and Reappraisal." *Psychological Bulletin* 126:530–55.

Goffman, Erving. 1955. "On Face Work: An Analysis of Ritual Elements in Social Interaction." *Psychiatry* 18:213–21.

Haley, Kevin J., and Daniel M. T. Fessler. 2005. "Nobody's Watching? Subtle Cues Affect Generosity in an Anonymous Economic Game." *Evolution and Human Behavior* 26:245–56.

Hall, Jeffrey A., Namkee Park, Hayeon Song, and Michael J. Cody. 2010. "Strategic Misrepresentation in Online Dating: The Effects of Gender, Self-Monitoring, and Personality Traits." *Journal of Social and Personal Relationships* 27:117–35.

Huckfeldt, Robert, John Sprague, and Jeffrey Levin. 2000. "The Dynamics of Collective Deliberation in the 1996 Election: Campaign Effects on Accessibility, Certainty, and Accuracy." *American Political Science Review* 94:641–51.

Jensen, Carlos, Colin Potts, and Christian Jensen. 2005. "Privacy Practices of Internet Users: Self-Reports versus Observed Behavior." *International Journal of Human-Computer Studies* 63:203–27.

Joinson, Adam N., Alan Woodley, and Ulf-Dietrich Reips. 2007. "Personalization, Authentication and Self-Discourse in Self-Discourse in Self-Administered Internet Surveys." *Computers in Human Behavior* 23:275–85.

Kam, Cindy D., and Marc J. Trussler. 2017. "At the Nexus of Observational and Experimental Research: Theory, Specification, and Analysis of Experiments with Heterogeneous Treatment Effects." *Political Behavior* 39:789–815.

Karp, Jeffrey A., and David Brockington. 2005. "Social Desirability and Response Validity: A Comparative Analysis of Overreporting Voter Turnout in Five Countries." *Journal of Politics* 67:825–40.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research.* Princeton, NJ: Princeton University Press.

Klar, Samara, and Yanna Krupnikov. 2016. *Independent Politics: How American Disdain for Parties Leads to Political Inaction.* New York: Cambridge University Press.

Krosnick, Jon A. 1991. "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys." *Applied Cognitive Psychology* 5:213–36.

Krupnikov, Yanna, Spencer Piston, and Nichole Bauer. 2016. "Saving Face: Identifying Voter Response to Black Candidates and Female Candidates." *Political Psychology* 37:253–73.

Lee, Carole J., and David Moher. 2017. "Promote Scientific Integrity via Journal Peer Review Data." *Science* 357:256–57.

Liu, Mingnan. 2018. "Data Collection Mode Effect on Abortion Questions: A Comparison of Face-to-Face and Web Surveys." *Gender and Women's Studies* 1:2.

Lupia, Arthur, and Colin Ellman. 2014. "Openness in Political Science: Data Access and Research Transparency." *PS: Political Science and Politics* 47:19–42.

Luskin, Robert C., and John G. Bullock. 2011. "'Don't Know' Means 'Don't Know': DK Responses and the Public's Level of Political Knowledge." *Journal of Politics* 73:547–57.

McDermott, Rose, 2011. "Internal and External Validity." In *Cambridge Handbook of Experimental Political Science*, edited by James N. Druckman, Donald Green, James H. Kuklinski, and Arthur Lupia, 27–40. New York: Cambridge University Press.

Merton, Robert. 1949. "The Role of Applied Social Science in the Formation of Policy: A Research Memorandum." *Philosophy of Science* 16:161–81.

Mondak, Jeffrey J. 2000. "Reconsidering the Measurement of Political Knowledge." *Political Analysis* 8:57–82.

Mondak, Jeffrey J., and Belinda Creel Davis. 2001. "Asked and Answered: Knowledge Levels When We Will Not Take 'Don't Know' for an Answer." *Political Behavior* 23:199–224.

Mueller, Karsten, Tammo Straatman, Kate Hattrup, and Marco Jochum. 2014. "Effects of Personalized versus Generic Implementation of an Intra-Organizational Online Survey on Psychological Anonymity and Response Behavior: A Field Experiment." *Journal of Business and Psychology* 29:169–81.

Nelson, Leif, Joseph Simmons, and Uri Simonsohn. 2018. "Psychology's Renaissance." *Annual Review of Psychology* 69:511–34.
Paulhus, Delroy L. 1991. "Measurement and Control of Response Bias." In *Measures of Personality and Social Psychology*, edited by John P. Robinson, Philip R. Shaver, and Lawrence S. Wrightsman, 17–59. New York: Academic Press.
Paulhus, Delroy L., Peter D. Harms, M. Nadine Bruce, and Daria C. Lysy. 2003. "The Over-Claiming Technique: Measuring Self-Enhancement Independent of Ability." *Journal of Personality and Social Psychology* 84:890–904.
Piston, Spencer. 2010. "How Explicit Racial Prejudice Hurt Obama in the 2008 Election." *Political Behavior* 32:431–51.
Premeaux, Sonya Fontenot, and Arthur G. Bedeian. 2003. "Breaking the Silence: The Moderating Effects of Self-Monitoring in Predicting Speaking Up in the Workplace." *Journal of Management Studies* 40:1537–62.
Rigdon, Mary, Keiko Ishii, Motoki Watabe, and Shinobu Kitayama. 2009. "Minimal Social Cues in the Dictator Game." *Journal of Economic Psychology* 30:358–67.
Ryan, John Barry. 2011. "Accuracy and Bias in Perceptions of Political Knowledge." *Political Behavior* 33:335–56.
Schuman, Howard, Stanley Presser, and Jacob Ludwig. 1981. "Context Effects on Survey Responses to Questions about Abortion." *Public Opinion Quarterly* 45:216–23.
Schwarz, Norbert, Hans-J. Hippler, Brigitte Deutsch, and Fritz Strack. 1985. "Response Scales: Effects of Category Range on Reported Behavior and Comparative Judgements." *Public Opinion Quarterly* 49:388–95.
Sears, David O., Colette Van Laar, Mary Carrillo, and Rick Kosterman. 1997. "Is It Really Racism? The Origins of White Americans' Opposition to Race-Targeted Policies." *Public Opinion Quarterly* 61:16–53.
Singer, Eleanor. 1978. "Informed Consent: Consequences for Response Rate and Response Quality in Social Surveys." *American Sociological Review* 43:144–62.
———. 1993. "Informed Consent and Survey Response: A Summary of the Empirical Literature." *Journal of Official Statistics* 9:361–75.
———. 2003. "Exploring the Meaning of Consent: Research and Beliefs About Risks and Benefits." *Journal of Official Statistics* 19:273–85.
———. 2011. "Toward a Benefit-Cost Theory of Survey Participation: Evidence, Further Tests, and Implications." *Journal of Official Statistics* 27:379–92.
Singer, Eleanor, and Mick P. Couper. 2014. "The Effect of Question Wording on Attitudes Toward Prenatal Testing and Abortion." *Public Opinion Quarterly* 78:751–60.
Sniderman, Paul M., and Philip E. Tetlock. 1986. "Symbolic Racism: Problems of Motive Attribution in Political Analysis." *Journal of Social Issues* 42:129–50.
Snyder, Mark. 1974. "Self-Monitoring of Expressive Behavior." *Journal of Personality and Social Psychology* 30:526–37.
———. 1979. "Self-Monitoring Processes." *Advances in Experimental Social Psychology.* 12:85–128.
Sudman, Seymour, Norman M. Bradburn, and Norbert Schwarz. 1996. *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology.* San Francisco: Jossey-Bass.
Tourangeau, Roger. 2004. "Survey Research and Societal Change." *Annual Review of Psychology* 55:775–801.
Vasilevsky, Nicole A., Matthew H. Brush, Paddock Holly, Laura Ponting, Shreejoy J. Tripathy, Gregory M. LaRocca, and Melissa A. Haendel. 2013. "On the Reproducibility of Science: Unique Identification of Research Resources in the Biomedical Literature." *PeerJ* 1:e148. doi:10.7717/peerj.148

Weber, Christopher, Howard Lavine, Leonie Huddy, and Christopher Federico. 2014. "Placing Racial Stereotypes in Context: Social Desirability and the Politics of Racial Hostility." *American Journal of Political Science* 58:63–78.

Yan, Ting, Richard Curtin, and Matthew Jans. 2010. "Trends in Income Nonresponse Over Two Decades." *Journal of Official Statistics* 26:145–64.

Zaller, John. 1992. *The Nature and Origins of Mass Opinion.* New York: Cambridge University Press.

Zaller, John, and Stanley Feldman. 1992. "A Simple Theory of Survey Response: Answering Questions versus Revealing Preferences." *American Journal of Political Science* 36:579–616.