

Expressive Survey Responding: A Closer Look at the Evidence and Its Implications for
American Democracy

Ariel Malka and Mark Adelman

Yeshiva University

To appear with minor edits in *Perspectives on Politics*

Last updated December 7, 2021

Abstract

Concerns about public opinion-based threats to American democracy are often tied to evidence of partisan bias in factual perceptions. However, influential work on *expressive survey responding* suggests that many apparent instances of such bias result from respondents insincerely reporting politically congenial views in order to gain expressive psychological benefits. Importantly, these findings have been interpreted as “good news for democracy” because partisans who knowingly report incorrect beliefs in surveys can act on their correct beliefs in the real world. We presently synthesize evidence and commentary on this matter, drawing two conclusions. First, evidence for insincere expressive responding on divisive political matters is limited and ambiguous. Second, when experimental manipulations in surveys reduce reports of politically congenial factual beliefs, this is often because such reported beliefs serve as flexible and interchangeable ways of justifying the largely stable allegiances that guide political behavior. So when circumstances render it costly to endorse a partisan belief, assessments of that belief become less diagnostic of the political predispositions that matter most, not more diagnostic of sincere views that will override partisan commitments. The expressive value of acting on political commitments should be viewed as a central feature of the American political context rather than a methodological artifact of surveys.

Concerns about the health of American democracy are often tied to notions of an acrimonious partisanship that powerfully structures political thinking and behavior within the public. Much of the evidence for this form of partisanship comes from political surveys, including findings that Democrats and Republicans have diverged on many partisan issues over the last several decades (Baldassarri & Park, 2020), have become more inclined to dislike out-partisans (Iyengar, Lelkes, Levendusky, Malhotra, & Westwood, 2019), and have revealed tendencies to follow partisan and other political cues when making political judgments (Arceneaux & Vander Wielen, 2017). Among the most troubling manifestations of this form of partisanship is the tendency of many partisans to report believing pieces of factual misinformation that are propagated by, or reflect positively on, their favored political elites (e.g., Jerit & Barabas, 2012; Kuklinski, Quirk, Jerit, Schwieder, & Rich, 2000; Nyhan, 2020). This suggests a relatively strong potential for elites to manipulate their followers and serious limitations on partisans' ability to retrospectively vote on the basis of objective social and economic conditions (Bartels, 2002; Bisgaard, 2019; Flynn, Nyhan, & Reifler, 2017).

Over the last several years, however, an influential program of research has cast doubt on some of the more troubling interpretations of this survey evidence. This work has centered around an inter-related set of survey response phenomena termed “expressive survey response,” “motivated responding,” “partisan cheerleading,” and “congenial inference” (Bullock, Gerber, Hill, & Huber, 2015; Prior, Sood, & Khanna, 2015; Schaffner & Luks, 2018). According to this general viewpoint, a substantial proportion of partisans who report belief in politically congenial misinformation do not sincerely believe the incorrect views that they report. Rather, they are motivated to either misrepresent their factual beliefs in order to gain expressive psychological

benefits from doing so (sometimes called “cheerleading”) or default to an expressively rewarding partisan response in the absence of certainty or after only a brief, biased sampling of considerations from long-term memory (sometimes called “congenial inference”) (Bullock et al. 2015; Prior et al., 2015). These two types of biased survey response – referred to collectively as *expressive survey responding* – are said to be an artifact of the survey context that produces exaggerated estimates of partisan division and misinformation.

Crucially, expressive survey responding has been regarded as more than a matter of methodological interest about the survey response process. Specifically, claims and suggestions have been made about the normative significance of expressive responding for American democracy. First of all, this evidence has been taken to mean that the inability of partisans to retrospectively vote on the basis of economic and societal conditions has been exaggerated. This is because “uncongenial information that respondents are reluctant to reveal may still affect their judgments. Partisans who withhold inconvenient information during a survey interview can draw on it when they develop policy preferences and make voting decisions” (Prior et al., 2015, p. 513; see also Khanna and Sood, 2018, p. 84). This, Prior et al. (2015) argue, “is bad news for survey research, but good news for democracy” (p. 490). As for those partisans who default to an expressively rewarding pro-partisan response when they do not know the answer, their “aware[ness] of their own ignorance... may make it easier to inform them of the facts and in turn change their votes” (Bullock et al., 2015, pp. 561). “In either of these cases,” Bullock et al. (2015, p. 521) note, “partisan differences in factual assessment would be of less concern than is suggested by prior work, because survey response would not reveal actual beliefs about factual matters.”

Second, this evidence has been taken to suggest that partisan divergence in *attitudes* (as opposed to factual beliefs) and partisan willingness to follow attitudinal cues from political leaders have also been exaggerated. Though careful to acknowledge that they did not empirically address self-reported attitudes, Bullock et al. (2015, p. 523) speculated that “if respondents misstate their factual beliefs in surveys because of their partisan leanings, they may misstate their attitudes in surveys for the same reason” (Bullock et al., 2015, p. 523). Indeed, Bullock et al. (2015) argue that “efforts to assess the dynamics of public opinion should grapple with the possibility that over-time changes in partisans’ expressed attitudes do not reflect changes in real beliefs” but might instead “reflect changes in the social returns to cheerleading” or “the degree to which different responses are understood to convey support for one’s party” (Bullock et al., 2015, pp. 561). As for experimental research on cues, Bullock et al. (2015) note the possibility “that partisan cues merely remind participants about the expressive utility that they gain from offering partisan-friendly survey responses” and that this research “may not be showing that partisanship alters actual attitudes or beliefs” (p. 523). Thus, “when survey reports of attitudes have expressive value, they may be inaccurate measures of true attitudes” (pp. 561).

The claims and suggestions from the expressive responding literature have not only had a strong scholarly impact, they have also influenced media coverage of polling since the rise of Donald Trump. Commentators have at times dismissed public opinion findings that are normatively undesirable or difficult to believe on the basis of this type of argument, sometimes going beyond the claims made in the scholarly literature. In one example, articles in the *Atlantic*, *Business Insider*, and the *National Review* cited research on expressive responding when skeptically covering a 2017 poll (Malka & Lelkes, 2017) suggesting that a high proportion of Republicans would be receptive to a flagrantly authoritarian action if Trump supported it

(Barro, 2017; Bloom, 2017; Graham, 2017). This literature was also invoked when interpreting early polling evidence that strong majorities of Republicans believed Donald Trump's claims of rampant fraud in the 2020 election and supported his efforts to overturn the election results. In a New York Times article from late November, 2020, titled, "Most Republicans Say They Doubt the Election. How Many Really Mean It?," the author noted that "Research has shown that the answers that partisans (on the left as well as on the right) give to political questions often reflect not what they know as fact, but what they wish were true. Or what they think they should say." (Badger, 2020). Normative implications drawn from the expressive responding literature suggested a more benign interpretation of evidence that most Republicans believed the incorrect factual bases of Trump's efforts to overturn a democratic election. Specifically, this might have represented an artifact of pressure to respond to survey questions in a party congenial fashion rather than an indication that Republicans would get on board with, and incentivize among their political elites, brazenly authoritarian behavior in the real world.

To summarize, the expressive responding literature has posited that substantial parts of the partisan belief gaps demonstrated in survey research are artifacts of the survey context and has suggested that citizens who are motivated to report insincerely held politically congenial beliefs in the survey context will – to an extent that matters normatively – act on their true politically uncongenial beliefs (or their acknowledgement of their ignorance) outside of the survey context.

The Present Argument

We believe that the expressive responding literature has made a valuable contribution by encouraging scholars, journalists, and others to explicitly consider the survey response process, and the potential for measurement errors associated with this, when interpreting results that

suggest partisan bias. In particular, this literature highlights how the intention to convey a political allegiance may sometimes be a more powerful motivator of survey response than the intention to provide a straightforward answer to a question whose meaning has been taken at face value.

That said, in this paper we push back on key claims associated with the expressive responding literature and describe a perspective on expressive motivation and partisan factual beliefs that we believe to be more consistent with the evidence. Our argument draws on ideas and findings that are scattered across the recent literature (Berinsky 2018; Bisgaard, 2019; Bullock & Lenz, 2019; Flynn, Nyhan, & Reifler 2017; Green, Kingzette, Minozzi, & Neblo, 2020; Iyengar et al., 2019; Kahan, 2015; Khanna & Sood, 2018; Nyhan, 2020; Peterson & Iyengar, 2021a; 2020b; Robbett & Matthews, 2018), including often neglected insights from the seminal expressive responding papers (Bullock et al., 2015; Prior et al., 2015). Our novel contribution is a synthesis of this theory and evidence distilled to two simple points that we hope will advance scholarly debate on this topic and provide insight that is useful for interpreting survey evidence for partisan division and misinformation.

Our first point is that the actual evidence that a large number of partisans are being insincere when they report politically congenial factual misperceptions on matters of strong partisan dispute is quite limited. In particular, incentives tend to reduce gaps to a far lesser extent – sometimes not at all – on issues of strong (as opposed to weak) partisan valence, and expressive responding findings in general have been inconsistent and ambiguous. Meanwhile, there is no convincing evidence for insincerity in reports of party congenial *attitudes* in politically salient issue domains.

Our second point pertains to the normative implications of expressive responding. Even though some partisans are insincere in their reports of politically congenial factual beliefs, we contend that there is little reason to expect that their privately held correct factual beliefs (or awareness of their ignorance) will overcome the political predispositions that gave rise to the expressive survey responding in the first place. Rather, evidence is more consistent with political predispositions – such as party identity, ideological identity, identification with a revered leader, or a broader identity encompassing a range of political and social self-representations (Mason & Wronski, 2018) – being the key drivers of political behavior, and incorrect factual beliefs serving as malleable and substitutable tools for justifying these political commitments in ways that are compatible with current circumstances. As we will show, evidence suggests that when circumstances render it costly or untenable to express one set of politically congenial beliefs, the partisan can readily adjust by adopting or emphasizing other politically congenial beliefs that support the underlying political commitment. Thus, these types of circumstances have the consequence of making self-reports of factual beliefs *less diagnostic* of the underlying political commitments that influence political behavior, not more diagnostic of sincere beliefs that will override partisan consistency pressures in the real world. The expressive value of acting on political commitments should be viewed as a central feature of the American political context, not a mere artifact of the survey context.

Defining Expressive Survey Responding

It is first of all important to establish clearly what has been meant by “expressive survey responding” and related terms. Despite definitional inconsistency, such responding is typically regarded as having two features, as we show in Table S-1 of the Supplementary Materials. One is represented in the term; an expressive survey response is motivated by the desire to express

support for a political team, usually a party but potentially a candidate, an ideological label, or some mix of mutually aligned identities. But the second feature is left out of the term: insincerity. Insincerity in a survey response entails privately believing one thing but saying another. For example, one might know that Barack Obama was born in the United States, or one might be aware of one's lack of knowledge about where he was born, but still respond to a survey question about this with an assertion that he was born outside of the United States.

It is important to note this explicitly, because of the confusion that might arise from terms such as "expressive" and "insincere." If a political attitude, belief, or behavior is "expressive" in nature, this does not necessarily mean that it is insincere (Hamlin & Jennings, 2011). One might be motivated to reach a conclusion because it is identity-consistent or bolsters one's self-representation as having a particular trait or belonging to a particular group, and still endorse this conclusion with sincerity. But expressive survey response is defined as both expressive and insincere.

Meanwhile, "insincerity" as defined in this literature refers to misrepresentation of a private psychological state. Insincerity in a survey response creates a disjunction between reported belief and one's private reckoning of what is true or that one does not know the answer. Hence the consistent depiction of expressive responding as an artifact of the survey context that distorts the picture of real-world partisan division and misinformation. In sum, the criteria for regarding a response as "expressive" have typically been (a) motivation to express support for one's side, combined with (b) misrepresentation of privately acknowledged belief, attitude, or uncertainty.

Finally, it is widely recognized that when responding to a political survey question, some (perhaps most) respondents retrieve considerations from long-term memory and use these to

construct a response in a ‘top-of-the-head’ fashion (Zaller, 1992). Oftentimes they are uncertain of their answer to factual questions (Graham, 2020), and might apply a heuristic (i.e., judgmental shortcut) in which trust of their own party (or distrust of the other party) guides their perception of reality (Bullock & Lenz, 2019). A discussion of how expressive responding relates to these phenomena is presented in Section 2 of the Supplementary Material.

Inconsistent and Ambiguous Evidence for Expressive Survey Responding

We have identified twelve papers as of this writing that have reported original findings that potentially inform the amount and nature of expressive survey responding. Seven of these papers made use of incentives for correct answers and the other five took different approaches. Given the claims and normative implications drawn from studies on expressive responding, it is worthwhile to take stock of this evidence and consider what conclusions are warranted. We do so in this section, beginning with the seminal papers by Prior et al. (2015) and Bullock et al. (2015). We also provide summaries of the twelve papers in Table S-2 of the Supplementary Materials. In some cases we report more details about prior studies than is customary in a review of this sort. We do so because we believe these sometimes overlooked details are important for making accurate inferences about the nature and extent of expressive responding. Finally, some of the information we report goes beyond what is reported in the papers reviewed, and was generously provided by a number of authors in response to our queries.

The Seminal Studies. Across two studies, Prior et al. (2015) had respondents provide estimates for pieces of factual economic information on which mild partisan gaps existed. In one study, respondents were randomly assigned to be paid or not paid for correct answers; in the other study they were randomly assigned to be paid (payment condition), encouraged to be accurate (accuracy appeal condition), or neither (control condition), and, independent of this,

were randomly assigned to either receive or not receive cues that the economic conditions were those under Republican President George W. Bush. In both studies, answers were coded as either accurate, overstating of economic problems, or understating of economic problems. Partisan bias was computed as percentage of questions on which party-congenial errors were made minus the percentage of questions on which party-uncongenial errors were made. In the first study, payment near-significantly (one-tailed p -value $< .10$) reduced partisan bias from 12.9% to 8.1%. In the second study, neither payment nor accuracy appeals reduced partisan bias when political cues were present, but bias was reduced in the absence of cues from 9.9% in the control condition to 3.8% in the payment condition and 3.4% in the accuracy appeal condition.

As for Bullock et al. (2015), across two studies respondents answered factual questions on which there were mild partisan gaps (e.g., amount of debt service spending). In the first study, respondents were randomly assigned to a control condition (no incentive offered), an “accuracy appeal” condition similar to that of Prior et al. (2015) (see footnote 28, page 559), or an incentivization condition in which correct answers were rewarded with entries in a drawing for a \$200 gift certificate. The second design was more complicated. It involved two no payment control groups, one without a ‘don’t know’ option and another – unanalyzed in the paper – with a ‘don’t know’ option. It also involved two incentivization groups – one where only correct answers were rewarded and one where both correct answers *and* ‘don’t know’ options were rewarded, with ‘don’t know’ selections (which were chosen close to 50% of the time in this condition) coded as non-polarized responses at the scale mean. In addition, incentive amounts were randomly varied. Unlike Prior et al. (2015), responses were not coded for accuracy, but rather were coded on a 0 to 1 scale with higher score meaning more of a Republican-friendly answer and lower meaning more of a Democratic-friendly answer, regardless of correctness.

In the first study, the average partisan gap across questions where significant partisan differences were observed was 11.8% of the total scale range. This gap was not significantly reduced in the accuracy appeal condition but was significantly reduced to 5.3% in the monetary incentive condition. In the second experiment, the average partisan gap with no incentives (and no ‘don’t know’ option) was 14.5% of the total scale range. This was reduced to 5.8% when respondents were paid for the correct answer and 2.8% when paid for both correct and “don’t know” answers with the latter placing partisans in a non-polarized position. In general, higher amounts of compensation reduced partisan gaps more. In addition, Bullock et al. (2015) reported a one-item replication of Study 2 in which partisan differences in belief about unemployment rate change during Obama’s first term averaged 36.6% in a no incentive condition with ‘don’t know’ selections treated as missing cases, which was near significantly (one-tailed $p < .10$) reduced to 23.4% of the scale range in the paid correct condition and significantly reduced to 14.4% in a paid for correct and 'don't know' condition with ‘don’t know’ responses coded at the non-polarized scale mean.

When considering in some detail the evidence from these seminal studies it would seem that strong conclusions about the prevalence of expressive responding are unwarranted. First of all, the types of factual beliefs sampled in the incentivization studies were not generally the subject of salient partisan dispute or were framed in a way that does not capture the essence of the partisan dispute (e.g., the precise average temperature increase between 1950-1980 and the year 2010) (Bullock & Lenz, 2019; Flynn et al., 2017; Nyhan, 2020; Peterson & Iyengar, 2021a; 2021b). This is reflected in the relatively small partisan belief gaps in the control conditions, ranging from 9 to 15 percentage points in the main studies. Incentives had the effect of making these already small partisan gaps even smaller.

Second, the incentivization findings were inconsistent across experimental comparisons. When a simple partisan cue was present in Prior et al.'s (2015) second study, incentives failed to reduce partisan bias. And although the accuracy appeal condition reduced partisan bias in Prior et al.'s (2015) no cues condition of Study 2 as expected, it did not reduce partisan gaps in factual beliefs in Bullock et al. (2015).

Third, although Bullock et al. (2015) speculated that their findings might have implications for insincerity of reports of partisan political *attitudes*, in addition to factual beliefs, their findings do not provide evidence for this.

Fourth, as Berinsky (2018), Bullock and Lenz (2019), and others have pointed out, the widely cited evidence in the incentivization studies is, in some cases, subject to multiple interpretations. Bullock et al. (2015) did not code for correct vs. incorrect answers in their main analyses, but in supplementary analyses they found that incentives for correct answers did not increase accuracy. Across the two papers then, there is rather limited evidence for incentives increasing the accuracy of partisans on factual questions. Furthermore, when incentives for correct responses reduce partisan gaps it is important to consider that certain incentivized partisans might be motivated to answer certain specific questions in the way they expect the researcher considers to be correct rather than the way they personally believe to be correct (Berinsky, 2018, pp. 217-218). Bullock and Lenz (2019, p. 332-333) argue that this is unlikely to explain partisan gap findings with balanced numbers of pro-Democrat and pro-Republican questions; but it might in part explain the reduction in partisan gap found in the single-item replication of Bullock et al. (2015) Study 2.

Fifth, it is important to highlight that Bullock et al. (2015) compared control conditions that did not offer “don’t know” options (Study 2) or that offered them but treated “don’t know”

selections as missing data (Study 2 replication) to incentive conditions in which ‘don’t know’ answers were offered, incentivized, and counted as non-polarized responses. In Study 2, it is possible that incentives increased ‘don’t know’ reports – and thereby decreased partisan gaps – among uncertain respondents who would have given a sincere best-guess based on party heuristics in the control condition (Berinsky, 2018; Bullock & Lenz, 2019). This would constitute a reduction in accuracy-motivated heuristic use, not a reduction in expressive survey responding. Furthermore, in the Study 2 replication, treating “don’t know” selections as non-polarized responses in the incentive condition but as missing data in the no incentive condition exaggerates the effect of incentives on partisan gap reduction.

Finally, as we will address in the discussion of normative implications, incentives for correct answers render the survey context less reflective of the real-world political environment, with its very different incentive structure.

Subsequent Studies Using Financial Incentives. A handful of subsequent studies have also incentivized correct responses to gauge the extent of insincere expressive responding. Taken together, these studies do not provide convincing evidence that much of the partisan gap in politically salient belief reports results from expressive survey responding. Take, for instance, Khanna and Sood’s (2018) paper addressing whether politically biased inferences from quantitative information reflect expressive responding. Respondents were presented with numeric data bearing on whether or not particular policies had been associated with good outcomes. After viewing a quantitative summary of evidence that was either congenial or uncongenial to their pre-existing attitudes, respondents answered a factual question about what the data showed, either with or without a financial incentive to answer correctly. As the authors predicted, the incentives never impacted the percentage giving the correct answer in response to

congenial information. As for the impact of incentive vs. no incentive on correct responses to uncongenial information, across the six comparisons reported, three showed significant increases and the other three showed no significant changes. The average increase in correct responses with incentives across these six comparisons was 4.6 percentage points. Overall, then, incentives did not reliably and substantially reduce incorrect inferences about politically uncongenial information.

So far we have considered studies that largely sampled content that is not the subject of strong partisan dispute or presented quantitative information in a way that does not match messages from the real-world political context (see Khanna & Sood, 2018, p. 98). Recent studies by Peterson and Iyengar (2021a; 2021b), however, overcome these limitations. In Peterson and Iyengar (2021a), partisans either did or did not receive financial compensation for correct answers to five questions that were the subject of strong partisan dispute, such as scientific consensus about human-caused global warming and millions of illegal votes having been cast in the 2016 election. Partisan gaps in the control condition were far larger than in the seminal studies, between 30 and 50 percentage points in seven out of ten cases. On average, providing an incentive reduced these partisan gaps by about one third, far less than the proportional reductions from the seminal studies. In the second paper (Peterson & Iyengar, 2021b), respondents were asked five factual questions about the origins, risk factors, and consequences of Covid-19. Partisan gaps in correct answers to these questions were 50, 39, 35, 18, and 0 percentage points. Partisan divisions were not diminished by either high (\$1.00) or low (\$0.25) incentives for correct answers. Meanwhile, partisan gaps for non-covid related factual questions (e.g., immigrant crime, climate change, and voter fraud) averaged 32

percentage points and were reduced by only 9 and 6 percentage points with low and high incentives, respectively.

A study by Allcott et al. (2020) also suggests that partisan gaps in Covid-related beliefs do not reflect expressive responding. They conducted a survey in April of 2020 and separately collected mobile device GPS data on social distancing behavior among millions of Americans. In the survey, respondents were asked to predict the number of confirmed Covid-19 cases that they expected in April of 2020, a clearly partisan matter at a time when a Republican President was minimizing the severity of the pandemic and touting his positive handling of it while receiving severe criticism from the Democratic opposition. Respondents answered the question with or without a financial incentive for accuracy. Democrats estimated that there would be more Covid cases than did Republicans, and incentives for getting it right did not decrease this gap, but non-significantly increased it. Moreover, GPS data revealed partisan gaps in social distancing, with Republican-voting counties doing notably less of it than Democratic-voting counties, even when controlling for likely confounds. As the authors succinctly put it, “[o]ur empirical results show that partisan gaps in beliefs and behavior are real.” (Allcott et al., 2020, p. 9). That said, Allcott et al. (2020) did find that incentives yielded a large reduction in the partisan divide on a single-item prediction of Trump’s future COVID handling approval rating.

Finally, a recent study by Robbett and Matthews (2018) financially incentivized correct answers to political questions for all respondents but varied whether respondents “voted” on these answers (with a correct majority required to earn the money) or acted alone as decisive individuals. As in the original incentivization studies, political questions were chosen on which partisan gaps were mild. However, these gaps were larger when respondents voted in groups (about 13 percentage points) than when they acted as decisive individuals (about 5 percentage

points). Furthermore, likelihood of giving a correct and politically uncongenial answer decreased by 12 percentage points when going from decisive individual to voter. The main conclusion from these findings, which we discuss more below, is that the closer the survey context comes to resembling the real world (election outcomes rather than individual citizen decisions impacting outcomes) the more expressive motivation matters.

Studies Taking Other Approaches to Gauging Insincere Expressive Responding.

Because of concerns raised about incentivization studies, other studies have taken different approaches to estimate the prevalence of insincere expressive responding. We briefly review these here and provide more detailed summaries in Table S-2 of the Supplementary Material.

Schaffner and Luks (2018) focused on responses to questions about photographs of inaugural crowds for Obama in 2009 and Trump in 2017, which provided unmistakable visual evidence that crowds were larger in Obama’s inauguration. In the key condition for gauging rates of expressive responding, respondents were shown both photos with labels “Image A” and “Image B,” and asked which photo displayed the larger crowd size. Fifteen percent of Trump voters gave the obviously wrong answer, compared to 2% of Clinton voters and 3% of non-voters. Thus, a strong majority of Trump voters did not report an obviously insincere Trump-congenial factual belief, and the overall partisan gap in doing so was 13 percentage points.

Connors (2020) asked respondents to answer questions about fraud in the 2020 election and whether the election results should be accepted, in one of two ways: as they thought a member of their party would respond to *impress* co-partisans or as they thought a member of their party would respond to *disappoint* co-partisans. The findings showed that partisans believed party congenial views would impress co-partisans while party uncongenial views would

disappoint co-partisans. While Connors (2020) rightly draws attention to the importance of social pressure in political belief expression, these particular findings would likely apply to any attitude or belief on which partisans differ (e.g., abortion) and do not speak to the sincerity of reports of partisan beliefs.

Other studies using non-incentivization methods also do not provide much evidence for expressive survey responding on matters of strong partisan dispute. Berinsky (2018), for example, conducted a series of studies involving non-financial interventions to induce sincere responding to questions about federal government involvement in the 9-11 attacks and Barack Obama's religion. These treatments did not impact proportions giving party-congenial incorrect answers. Yair and Huber (2020), meanwhile, examined the extent to which partisan bias in attractiveness ratings reflects expressive responding by adapting procedures developed by Gal and Rucker (2011) to study "response substitution," or the tendency to use questions about specific evaluations (like the quality of a restaurant's service) to express a broader opinion that one considers more important (e.g., one's overall opinion of the restaurant). Yair and Huber (2020) gave some respondents, but not others, a treatment that would obviate the motive to respond in this way (e.g., giving them a chance to express their partisanship before rating the target's attractiveness). Republicans did not display the predicted partisan bias in attractiveness ratings. Democrats, however, did, and this bias was reduced by half when given one of the treatments. However, as the authors note, this study is quite limited in its ability to inform the extent to which expressive responding underlies partisan belief gaps on divisive issues. Most importantly, "assessments of physical attractiveness do not evoke deep partisan feelings" (p. 25); which is perhaps why partisan bias was not observed for Republicans. Moreover, the authors

queried “attractiveness,” not “physical attractiveness” per se, so the treatment might have altered which considerations Democrats drew upon when evaluating the target’s attractiveness.

Finally, Graham and Huber (2020) took a creative approach to estimating the proportion of American respondents who derive expressive value from answering political survey questions. After answering initial questions, respondents were given an option to either answer five extra questions before the final question, or to simply skip to the final question. A large proportion of respondents overall (64%) opted to answer extra questions, and the proportions were highest when respondents were made to think they would be asked about politically salient partisan and rumor questions (almost reaching 80%) and among the most partisan and politically engaged respondents. Though insightful in many ways, these findings do not provide evidence for the degree to which partisan gaps in reported factual beliefs are due to expressive responding. First, that many partisans derive psychological value from answering political questions does not mean that the answers to these questions are insincere. Moreover, the extent to which those opting to answer extra questions are deriving psychological value from the act of self-expression, as opposed to satisfaction of other psychological needs (such as having curiosity satisfied) is questionable (see Table S-2 in the Supplementary Material).

Summary. We have delineated the evidence for expressive responding in some detail. We conclude that there does not exist convincing evidence that large portions of the partisan factual belief gaps on divisive matters are attributable to this. That said, it remains very likely that *some* portion of the partisan bias demonstrated in surveys reflects insincere expressive responding. In the next section we address the potential normative implications of this matter.

Normative Implications of Expressive Responding for American Democracy

A key normative inference drawn from the expressive responding literature is that partisans who report incorrect politically congenial factual beliefs in surveys might still draw on accurate factual knowledge, or acknowledgement of their own ignorance, when voting and making other important political choices. There is little doubt that some portion of the partisans who report incorrect politically congenial beliefs are making assertions in surveys that they do not sincerely believe in order to express support for their side. And even if just a small number of partisans act on politically uncongenial beliefs that they are reluctant to reveal in surveys, this could have electorally and politically significant consequences in the United States.

In this section, however, we will argue that existing evidence undermines this key normative conclusion. In particular, we will make the case that the very political commitments – the mix of partisan, ideological, and other identities – that make it psychologically rewarding to express politically congenial factual beliefs in surveys also make it psychologically and socially rewarding to engage in identity-expressive political behavior in the real world. When features of the survey context cause a change in politically congenial belief expression, this is often because such beliefs serve as substitutable tools for rationalizing political commitments in a way that is attuned to momentary situational affordances. If one politically congenial factual belief becomes costly to uphold in the moment, politically biased reasoning allows the partisan to selectively emphasize other considerations, deflect from the current argument, or engage in other strategies that maintain the integrity of the political predisposition. As Khanna and Sood (2018) felicitously put it, biased belief expression has a “whack-a-mole nature” (p. 79).

Expressive Survey Responding Reflects Political Orientations with Real-World Relevance

Do incentives for correct responding yield answers that reflect a more honest reckoning that will override political commitments in the real world outside of the survey context? We do

not think so. In fact, evidence suggests that it is more likely that incentives in the survey context artificially render the response process less diagnostic of real-world political motives. In the real world, there are strong expressive and social incentives to toe the party line (Connors, 2020; Kahan, 2015). Position on a political divide can be central to one's identity and integral to one's belonging and standing among important others (Stern & Ondish, 2018). Though politically congenial belief expression in surveys is sometimes dismissed as "cheap talk," for a partisan it is more likely that *defiance* of the party line under incentives in a survey is cheap talk. Private doubts about the party line can be more safely expressed in a private survey context for a few dollars than in the outside world in which there would be psychological and social costs. Expressive motivation is a prominent factor in actual political behavior, and experimental treatments (like financial incentives) that reduce its role in belief reports should therefore yield information that is less diagnostic of the predispositions that guide actual political behavior.

In fact, Bullock et al. (2015) reported evidence consistent with this view in their seminal paper on expressive responding. They found that holding factual beliefs that were favorable to the Republican vs. Democratic Party based on normal unincentivized reports was a strong correlate of vote choice. However, this correlation was substantially reduced when factual beliefs were measured with incentives for correct responses. As the authors note, this is consistent with factual beliefs themselves having a weaker causal influence on vote choice than might be assumed and suggests caution in inferring that partisanship impacts vote choice via biasing factual beliefs (pp. 556-558; see also Prior et al., 2015, p. 514). In our view, this finding also suggests that an absence of incentive to produce a correct answer yields an indicator that is more reflective of a predisposition that underlies real-world political choices (see Green et al., 2020).

The more general point here is that the type of expressive motivation that impacts normal survey responding also energizes real-world political behavior. Indeed, the findings of Robbett and Matthews (2018) suggest that when the survey context comes closer to resembling the real-world, expressive motivation exerts a greater impact on survey response. They found that when voting on correct answers to partisan factual questions (with a majority required for the incentive), as opposed to acting as decisive individuals (with one's answer alone determining one's receipt of the incentive), partisan gaps in factual beliefs and politically congenial incorrect responses both increased. The authors noted that their findings provided "strong evidence of expressive voting" and that "partisan bias is not an artifact of unincentivized questionnaires" (p. 3).

Politically Congenial Factual Beliefs as Flexibly Adjustable Political Self-Justification Mechanisms.

It is clear that sometimes partisans adjust their expressions of factual beliefs as a consequence of the survey context. Again, this has been taken to suggest that partisans who can be made to admit to politically uncongenial beliefs in a survey are, to an extent that matters normatively, likely to act on their sincere beliefs in the real world. In contrast, we argue that when partisan factual beliefs do prove malleable, it is because they are flexibly interchangeable ways of justifying one's stable political predisposition. When the survey context (or any other context for that matter) alters the costs and benefits of expressing one particular partisan factual belief, the partisan can simply tweak the belief system to maintain the political predisposition.

Recent experimental evidence shows that partisans adjust factual beliefs in order to rationalize political commitments in a way that is tailored to the affordances present in the survey context. Lauderdale (2016), for example, found that informing Americans of Obama's

favorable opinions about Egyptian democracy-promotion efforts and free trade agreements led to a partisan divergence in factual beliefs relevant to these matters. As a result of this information, Democrats and Republicans diverged in beliefs about Egyptians' attitudes toward the U.S. (with Democrats viewing these as less negative than Republicans) and manufacturing job losses as a consequence of free trade agreements (with Democrats viewing these as less severe than Republicans). When reasoning in response to shifting situational affordances, partisans seem to "find their way to factual beliefs that will not call their political commitments into question." (Lauderdale, 2016, p. 3).

When manipulations reduce the rate of politically congenial responding to a specific factual question, they are apparently raising the costs of providing a politically congenial answer to that particular question in that moment. A substantial amount of evidence now suggests that partisans are quite resourceful in tweaking their factual belief systems to accommodate an abandonment of one particular partisan belief necessitated by the survey context. Bisgaard (2019) addressed this matter in the context of uncongenial factual economic information. He found that when presented with such information, partisans often accepted it but then adjusted their perception of whether the incumbent was responsible for these conditions. If the partisan had to accept that a same-party incumbent presided over negative economic changes or an out-party incumbent presided over positive economic changes, they adjusted their belief systems by viewing incumbents as less responsible for the economy (see also Tilley & Hobolt, 2011). Similarly, in the study by Khanna and Sood (2018) summarized above, when partisans correctly acknowledged that empirical evidence was uncongenial to their attitudes, they adjusted by dismissing the credibility of this evidence. Other work shows that when induced to accept

negative information about their own party, partisans often adjust by further degrading the opposing party in a “lesser-of-two evils” political self-justification strategy (Groenendyk, 2013).

Indeed, a long line of psychological theorizing posits that certain beliefs can be flexibly adjusted to provide a sense of “balance” or psychological consistency, depending on the other beliefs one holds and the affordances in the situation (Heider, 1958; Shafer, 1981). This type of theorizing has informed discussions of motivated reasoning aimed at reaching conclusions that are compatible with one’s political identity (Druckman, Leeper, & Slothuus, 2018; Taber & Lodge, 2006). What is impressive about these psychological processes is the apparent scope of the toolkit partisans possess for motivated reasoning. For example, candidate supporters manage to maintain this support even when they are successfully induced to accept that their preferred candidate has made factually incorrect statements (Swire, Berinsky, Lewandowsky, & Ecker, 2017). In this case, partisans may adjust their view of the importance of veracity as a candidate quality depending on whether the context forces them to accept their candidate’s incorrect statements. As Effron (2018) has shown, simply imagining how one’s preferred candidate’s dishonest statements “could have been true” (e.g., Trump’s inauguration crowd could have been larger had the weather been nicer) leads candidate supporters to view dishonest statements as less condemnable.

This type of evidence suggests that a key feature of factual belief expression is that it often serves as a versatile strategy for justifying the identity commitments that guide political behavior, rather than an independent cause of political behavior. This aligns with the view that moral reasoning often serves the function of justifying conclusions reached for other (often identity-based) reasons (Haidt, 2001; Uhlmann, Pizarro, Tannenbaum, & Ditto, 2009) and that the causal influence of factual beliefs on political behavior might be smaller than assumed

(Bullock et al., 2015). When circumstances induce an adjustment away from politically congenial factual beliefs, this is not likely to yield a corresponding adjustment of vote intention or candidate support (Swire et al., 2017). In addition, partisans become more likely to accept what had been inconvenient facts when the implications of these facts for their political commitments are altered by the survey context. Campbell and Kay (2014), for example, found that conservative identifiers in the United States became more likely to accept the reality of human-caused climate change after being led to believe that solutions to the problem of climate change would be consistent with their conservative ideology. We certainly would not expect this belief change to endure when people leave the survey context and return to the experience of real-world political incentives (Nyhan, 2020; Swire et al., 2017). But that is precisely the point; political belief expression is often a flexibly adjustable and situationally attuned strategy for justifying stable political commitments based on what is happening in the momentary context.

Political Engagement is Linked with both Expressive Responding and Stable Political Allegiance

We argue that the Americans who insincerely report politically congenial views in surveys to gain expressive psychological rewards are likely to act in the real world on the political predispositions that caused them to respond expressively. This argument is further bolstered by evidence that the most politically engaged and partisan Americans – that is, those who are most likely to act on stable political allegiances (Zaller, 1992) – are also apparently the most likely to engage in expressive responding. In the seminal paper by Prior et al. (2015), for example, it was the most politically knowledgeable respondents who were most likely to give politically congenial wrong answers and to reduce their rate of politically congenial wrong answers when an incentive was on the line. In the paper that most effectively zeroes in on insincere expressive responding, Schaffner and Luks (2018) found that the gap between Clinton

and Trump voters in report of an obviously wrong pro-Trump answer went from 9% among those without a college education to 25% among those with a college education, and similar results were found for gaps across levels of political interest. Graham and Huber (2020) found that those most inclined to make a choice revealing that they find political survey responding to be psychologically rewarding were those who were most politically interested and partisan. All of this suggests that the type of person who would find it most rewarding to report partisan views in a survey is the type of person who is likely to stick to their partisan commitments in the real world.

Conclusion

On January 6, 2021, a mob of Trump supporters stormed the U.S. Capitol Building in an apparent attempt to thwart the certification of Joe Biden's electoral college victory. It would not be unreasonable to expect that for many participants in this action, the behavior of storming the Capitol was expressively motivated. That is, the psychological value gained from acting on one's political identity and sharing this experience with politically like-minded others is likely to have exceeded the expected instrumental value of an effort to prevent Biden from assuming the Presidency. But this was, nonetheless, real-world political behavior that matters.

One could certainly question the depth of sincerity with which some Trump supporters endorse the "big lie." However, the present review provides two insights that should be kept in mind when evaluating this and related matters. First, evidence from studies of expressive responding do not show that substantial parts of partisan gaps in expression of politically salient factual beliefs are insincerely reported. Thus, for example, there is not a strong basis for expecting that most Trump supporters who report belief in a stolen election are misrepresenting their private beliefs. Second, evidence from studies of political rationalization and motivated

reasoning suggest that those who do report insincere politically congenial beliefs in surveys are unlikely to act on their correct beliefs in the real world. Thus, there is not a strong basis for expecting that those Trump supporters who say, but do not really mean, that they endorse the “big lie” will do anything but act on the political commitments that motivated them to cheerlead in the first place.

Indeed, evidence abounds that partisan responses to survey questions reflect beliefs or dispositions that matter in the real world. For one thing, elites pay attention to polls, so even insincere expressively motivated responses influence the political incentives that elites perceive and respond to. For another, evidence has accumulated that self-reported partisan beliefs and attitudes are reflected in real-world behavior, including policy uptake (Lerman, Sadin, & Trachtman, 2017), behavioral discrimination against out-partisans (Iyengar & Westwood, 2015), dating choices (Huber & Malhotra, 2017), public health behavior in the context of the Covid-19 pandemic (Allcott et al., 2020), and online sharing of politically congenial headlines whose inaccuracy one is aware of (Pennycook et al., 2021). As our review suggests, when an American partisan reports a factually incorrect but politically congenial belief, it is very likely that they either really believe it or are nonetheless inclined to behave as though they do.

Our analysis also has implications for the view that survey questions about factual political matters should avoid the use of partisan or candidate cues (Prior et al., 2015). The argument is that such cues are likely to heighten the salience of expressive, as opposed to accuracy, motivation in answering questions. While we agree that political cues in factual questions are likely to influence the distribution of responses (e.g., Graham, 2020), our analysis suggests that the presence of such cues in factual questions can be advantageous. Specifically, the expressive motivation that sometimes underlies responses to factual questions is a powerful

influence on real-world political behavior. It follows, then, that when factual answers are strongly influenced by desire to express political allegiance – a phenomenon that is more likely in the presence of cues – they will also be more likely to reflect the important predispositions that guide political behavior (see Green et al., 2020). In addition, inclusion of cues in factual questions will enable some uncertain respondents to construct a sincerely held – albeit not well considered – belief, via accuracy-motivated heuristic use. Accuracy-motivated heuristic use, self-expressive behavior, and the political cues that enable their application are important features of the American political context. Therefore, making them operative in the survey context has advantages.

Finally, we offer two related recommendations for future research on expressive motivation and survey response. The first is that research on cues, incentives, and other question features that alter the motivation underlying survey response should focus on *differences in the type of construct* gauged by factual questions when different question features are present. Rather than simply assuming that features that enhance expressive motivation are adding “error,” researchers should be open to the possibility that they are altering the construct being assessed to one that is perhaps more important (see Green et al., 2020). The second recommendation is that attention be devoted to the possibility that factual beliefs might not necessarily be an important part of the process by which partisanship impacts political behavior (Bullock et al., 2015; Prior et al., 2015). Factual perceptions might often be side effects of the political commitments that make it rewarding to act in a partisan matter, with little independent causal influence on political behavior. The methods pioneered in the expressive responding literature have potential to inform this important normative matter.

References

- Allcott, H., Boxell, L., Conway, J.C., Gentzkow, M., Thaler, M., & Yang, D.Y. (2020). Polarization and Public Health: Partisan Differences in Social Distancing during the Coronavirus Pandemic. NBER Working Paper 26946.
- Arceneaux, K., & Vander Wielen, R. J. (2017). Taming intuition: How reflection minimizes partisan reasoning and promotes democratic accountability. Cambridge: *Cambridge University Press*.
- Badger, Emily. 2020. "Most Republicans Say They Doubt the Election. How Many Really Mean It?." *New York Times*, November 30.
- Baldassarri, D., & Park, B. (2020). Was there a culture war? Partisan polarization and secular trends in US public opinion. *The Journal of Politics*, 82(3), 809-827.
- Barro, J. (2017, August 10). *Trump produces enough real risks, so stop imagining fake ones*. Business Insider Nederland.
- Bartels, L. M. (2002). Beyond the running tally: Partisan bias in political perceptions. *Political behavior*, 24(2), 117-150.
- Berinsky, A. J. (2018). Telling the truth about believing the lies? Evidence for the limited prevalence of expressive survey responding. *Journal of Politics*, 80, 2011– 224.
- Bisgaard, M. (2019). How Getting the Facts Right Can Fuel Partisan-Motivated Reasoning. *American Journal of Political Science*.
- Bloom, M. (2017, August 15). *Polls Don't Measure What You Think They Measure*. National Review.
- Bullock, J. G., A. S. Gerber, S. J. Hill, & G. A. Huber. (2015). Partisan Bias in Factual Beliefs about Politics. *Quarterly Journal of Political Science* (10): 519–78.

- Bullock, J. G., & Lenz, G. (2019). Partisan bias in surveys. *Annual Review of Political Science*, 22(1), 325–342.
- Campbell, T. H., & Kay, A. C. (2014). Solution aversion: On the relation between ideology and motivated disbelief. *Journal of Personality and Social Psychology*, 107, 809–824.
- Connors, E. C. (2020, December 22). *Do Republicans really believe the election was stolen – or are they just saying that?* Washington Post.
- Druckman, J. N., Leeper, T.J., & Slothuus, R. (2018). Motivated responses to political communications: Framing, party cues, and science information. In H. G. Lavine & C. S. Taber (Eds.), *The feeling, thinking citizen: Essays in honor of Milton Lodge* (pp. 125-150). New York, NY: Routledge.
- Effron, D. A. (2018). It could have been true: How counterfactual thoughts reduce condemnation of falsehoods and increase political polarization. *Personality and Social Psychology Bulletin*, 44(5), 729-745.
- Flynn, D. J., Nyhan, B., & Reifler, J. (2017). The Nature and Origins of Misperceptions: Understanding False and Unsupported Beliefs about Politics. *Advances in Political Psychology* 38(S1): 127–50
- Gal, D. & Rucker, D.D. (2011). Answering the unasked question: Response substitution in consumer surveys. *Journal Marketing Research* 48(1), 185–195.
- Graham, D. A. (2018, June 19). *Do Republicans Want to Postpone the 2020 Election?* The Atlantic.
- Graham, M. H. (2020). Self-awareness of political knowledge. *Political Behavior*, 42, 305-326.

- Graham, M. H., & Huber, G. A. (in press). The expressive value of answering survey questions. Forthcoming in *The Politics of Truth in a Polarized Era*, D. C. Barker & Suhay, E. (Eds.). Oxford University Press
- Green, J., Kingzette, J., Minozzi, W., & Neblo, M. A Speech Act Perspective on the Survey Response. Unpublished manuscript.
- Groenendyk, E. W. (2013). Competing motives in the partisan mind: How loyalty and responsiveness shape party identification and democracy. New York, NY: Oxford University Press.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108, 814-834.
- Hamlin, A., Jennings, C. (2011). Expressive Political Behaviour: Foundations, Scope and Implications, *British Journal of Political Science*, 41, 645-70.
- Heider, F. (1958). The Naive Analysis of Action. *The Psychology of Interpersonal Relations*, Hoboken, NJ: Wiley, 79–124.
- Huber, G. A., & Malhotra, N. (2017). Political homophily in social relationships: Evidence from online dating behavior. *Journal of Politics*, 79(1), 269–283.
- Iyengar, S., & Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science*, 59, 690–707.
- Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N. & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annu. Rev. Polit. Sci.* 22, 129–146.
- Jerit, J., & Barabas, J. (2012). Partisan perceptual bias and the information environment. *Journal of Politics*, 74(3), 672–684.

- Kahan, D. M. (2015). The expressive rationality of inaccurate perceptions. *Behavioral & Brain Sciences*, 40, 26-28.
- Khanna, K., & Sood, G. (2018). Motivated responding in studies of factual learning. *Political Behavior*, 40(79), 1-23.
- Kuklinski, J. H., Quirk, P. J., Jerit, J., Schweider, D., & Rich, R. F. (2000). Misinformation and the currency of democratic citizenship. *The Journal of Politics*, 62(3), 790–816.
- Lauderdale, B. E. (2016). Partisan disagreements arising from rationalization of common information. *Political Science Research and Methods*, 4(3), 477-492.
- Lerman, A.E., Sadin, M.L., & Trachtman, S. (2017). Policy Uptake as Political Behavior: Evidence from the Affordable Care Act. *American Political Science Review* 111 (4): 755–70.
- Malka, A., Lelkes, Y. (2017, August 10). *In a new poll, half of Republicans say they would support postponing the 2020 election if Trump proposed it.* Washington Post.
- Mason, L., & Wronski, J. (2018). One Tribe to Bind Them All: How Our Social Group Attachments Strengthen Partisanship. *Political Psychology* 39: 257–77.
- Nyhan, B. (2020). Facts and myths about misperceptions. *Journal of Economic Perspectives*, 34(3), 220-36.
- Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A.A., Eckles, D., & Rand, D.G. (2021). Shifting Attention to Accuracy Can Reduce Misinformation Online. *Nature* 1–6.
- Peterson, E., & Iyengar, S. (2021a). Partisan Gaps in Political Information and Information-Seeking Behavior: Motivated Reasoning or Cheerleading?. *American Journal of Political Science*, 65(1), 133-147.

- Peterson, E., & Iyengar, S. (2021b). Partisan Reasoning in a High Stakes Environment. Manuscript in preparation.
- Prior, M., Sood G., & Khanna, K. (2015). You Cannot Be Serious: The Impact of Accuracy Incentives on Partisan Bias in Reports of Economic Perceptions. *Quarterly Journal of Political Science* 10(4): 489–518.
- Robbett, A., & Matthews, H.P. (2018) Partisan bias and expressive voting. *Eur. J. Pol. Econ.* 157, 107–120.
- Schaffner, B. F., & Luks, S. (2018). Misinformation or expressive responding? What an inauguration crowd can tell us about the source of political misinformation in surveys. *Public Opinion Quarterly*, 82(1), 135-147.
- Shaffer, S. D. (1981). Balance theory and political cognitions. *American Politics Quarterly*, 9(3), 291-320.
- Stern, C., Ondish, P. (2018). Political aspects of shared reality. *Current Opinion in Psychology*, 23, 11–14.
- Swire, B., Berinsky A.J., Lewandowsky, S., Ecker, U.K.H. (2017). Processing political misinformation: comprehending the Trump phenomenon. *Royal Society Open Science* 4(3):160802.
- Taber, C.S., & Lodge, M. (2006). Motivated Skepticism in the Evaluation of Political Beliefs. *American Journal of Political Science* 50(3): 755–69.
- Tilley, J., & Hobolt, S. B. (2011). Is the government to blame? An experimental test of how partisanship shapes perceptions of performance and responsibility. *Journal of Politics*, 73(2), 316-330.

- Uhlmann, E., Pizarro, D., Tannenbaum, D., & Ditto, P. (2009). The motivated use of moral principles. *Judgment and Decision Making*, 4, 476 – 491.
- Yair, O., & Huber, G.A. (2020). How Robust is Evidence of Partisan Perceptual Bias in Survey Responses? A New Approach for Studying Expressive Responding. *Public Opinion Quarterly* 84(2): 469–492.
- Zaller, J. (1992). *The nature and origins of mass opinion*. New York, NY: Cambridge University Press.