Supplementary Online Materials for:

Malka, A., & Adelman, M. Expressive Survey Responding: A Closer Look at the Evidence and Its Implications for American Democracy. *Perspectives on Politics.*

Table of Contents

Section 1: Tables

Table S1: Sample Statements Defining Expressive Survey Responding Phenomena	
Paper	**Definition**
Prior, Sood, & Khanna, 2015	"Partisans give answers that reflect well on their party even when they have information that is less flattering to their party, or could have easily inferred such information." (p. 492).
	"Our main contribution is to show that, under typical survey conditions, answers to factual questions with partisan implications are contaminated by partisans' motivation to give answers that reflect well on their party. Deliberately or not, some partisans treat factual questions with political relevance as an opportunity to root for their team. Deep-seated perceptual differences between partisans are substantially smaller than past work suggests." (pp. 492-493).
	"Our results demonstrate that a significant portion of what scholars have called perceptual bias is in fact an artifact of partisan consistency pressures during the measurement of those perceptions." (p. 510).
Bulllock, Gerber, Hill, & Huber, 2015	"An alternative view is that survey responses are not entirely sincere. Instead, they may reflect the expressive value of making statements that portray one's party in a favorable light. . . Partisan divergence in surveys may therefore measure the joy of partisan "cheerleading" rather than sincere differences in beliefs about the truth." (p. 521).
	" 'congenial inference': when trying to answer a question under ordinary conditions, partisans are especially likely to call to mind those considerations that put their own party in a favorable light, and they infer the correct answer to the question at hand from this congenial set of considerations (e.g., Zaller, 1992, Chapter 5)".
Schaffner & Luks, 2018	"A recent challenge to the belief that misperceptions are truly held by individuals comes from Bullock et al. (2015) and Prior, Sood, and Khanna (2015), who argue that at least some of the misperceptions that respondents provide in response to survey questions are not genuinely held beliefs. This body of work proposes an alternative explanation, called expressive responding, whereby individuals intentionally provide misinformation to survey researchers as a way of showing support for their political viewpoint." (p. 136).
Berinsky, 2018	"I assess the extent to which subscription to political rumors represents genuine beliefs as opposed to expressive responses—rumor endorsements designed to express opposition to politicians and policies rather than genuine belief in false information." (p. 211).

Khanna & Sood, 2018	"differences in survey reports of factual beliefs do not always reflect differences in what people believe. Instead, they may be artifacts of the survey response process. Respondents sometimes give congenial but inaccurate answers in response to factual questions even when they have accurate but uncongenial facts at hand (Bullock et al. 2015; Prior et al. 2015). Other times, respondents are ignorant, having no relevant cognitions, and they offer a congenial answer as their best guess (Luskin et al. 2013). In both cases, the survey response process inflates estimates of bias in factual beliefs." (p. 80)
Bullock & Lenz, 2019	"cheerleading" defined as "When asked a factual question, respondents may believe one answer but give a different answer to support their party." (p. 327).
	"Directional motives, too, may operate in more than one way. If our Republican respondent is confident that unemployment rate declined under Obama, a directional motive may simply incline him to give a response that he believes to be false. This is insincere responding, or *cheerleading*. On the other hand, our Republican respondent may not be confident of the correct answer. In this case, he may canvas his memory for considerations related to the unemployment rate—but do so in a way that makes him especially likely to retrieve considerations that cast Obama in a negative light. This is biased consideration sampling. Alternatively, the Republican may rely on a pro-party heuristic (e.g., Khanna & Sood 2018) to determine his answer: for example, a heuristic that says "give the response that makes Obama look bad unless faced with incontrovertible contrary evidence." Following Prior et al. (2015, especially p. 494), we use the term *congenial inference* to refer to either biased consideration sampling or the use of pro-party heuristics" (p. 328-329)
	"Motivated responding concerns what survey respondents report on surveys, rather than what they have learned or know. . . it occurs when people with the same underlying beliefs give congenial answers more often than uncongenial answers when asked about their beliefs. In all, it contends that survey responses to factual questions reflect a mix of what people believe and what they wish to be true (Luskin et al. 2013; Prior et al. 2015)." (p. 83).
Yair & Huber, 2020	"Another view is that in answering survey questions, individuals provide responses that also communicate their partisan proclivities, such that partisan differences may indicate "expressive responding" (Bullock et al. 2015) rather than sincere differences." (p. 470)
Graham & Huber, 2020	"Expressive responding is the notion that individuals may answer questions not simply on the basis of what they truly believe, but also because they get more expressive benefit from some responses than others." (p. 1).

Peterson & Iyengar, 2021a	"A contrasting interpretation is that these divides reflect insincere partisan cheerleading in which partisans knowingly distort their responses to survey questions to signal support for their side (Bullock et al. 2015; Prior,Sood, and Khanna 2015; Schaffner and Luks 2018). In this view, partisans are in fact well informed, but they prefer to act misinformed in surveys. In essence, the cheerleading account implies that the partisan divides over factual evidence measured in surveys are largely illusory." (pp. 133-134).
	"Proponents of cheerleading argue that polarization creates short-term psychic rewards for partisans to offer knowingly incorrect responses to knowledge questions, even as it does not induce genuine belief in these party-congenial falsehoods." (p. 138)

Table S2: Summaries of Methods and Key Findings About Prevalence of Expressive Survey Responding				
Study	**Sample**	**Field dates**	**Method Summary**	**Summary of Key Findings**
Prior, Sood, & Khanna (2015), Study 1	Partisans from national probability sample (Knowledge Networks), N = 471	October - November, 2004	Respondents answered five factual economic questions and were randomly assigned to receive $1 per correct answer or to a control condition with no financial reward for answering correctly. Responses were coded as correct, overstating of economic problems, or understating of economic problems. Partisan bias was operationalized as the difference between the percentage of party congenial and party uncongenial errors.	Across all questions, partisan bias averaged 12.9 percentage points in the control condition and was near-significantly reduced using a one-tailed test ($p < .10$) to 8.1 percentage points in the monetary incentives condition.
Prior, Sood, & Khanna (2015), Study 2	Partisans from national probability sample (Knowledge Networks), N = 660 (excludes respondents in the Bush reference treatment condition that was omitted from main analyses)	March - April, 2008	Respondents answered five factual economic questions and were randomly assigned to either receive $2 per correct answer, to receive an appeal to be accurate, or to a control condition with no financial reward for answering correctly and no accuracy appeal. Respondents were also randomly assigned to either receive or not receive explicit information that the questions referred to economic changes during the G.W. Bush Presidency. Partisan bias was operationalized in the same way as in Study 1.	Across all questions when G.W. Bush was referenced, partisan bias was not reduced by either an accuracy appeal or monetary incentives. Across all questions when no mention was made of G.W. Bush, partisan bias averaged 9.9 percentage points in the control condition and was significantly reduced to 3.4 percentage points in the accuracy appeal condition and 3.8 percentage points in the monetary incentives condition.

Bulllock, Gerber, Hill, & Huber, 2015, Study 1	Partisans from demographically representative sample selected from opt-in internet panel for 2008 CCES (YouGov), N = 419 (Excludes respondents in accuracy appeal condition)	October, 2008	Respondents answered ten political questions and were randomly assigned to receive an entry into a drawing for a $200 Amazon gift card for each correct answer, to receive an appeal to be accurate, or to a control condition with no financial reward for answering correctly and no accuracy appeal. Partisan divergence was operationalized as between-party difference in scale scores (coded so higher score means a more Democrat-friendly response), without regard for correctness. Only the 8 questions on which near-significant (one-tailed $p<.10$) partisan difference emerged were included in analyses.	The accuracy appeal condition was not included in the main analyses and its associated findings were not reported in detail, but the authors stated that it did not impact partisan divergence (footnote 28, p. 559). Across the eligible questions, partisan divergence averaged 11.8% of the scale range in the control condition and was significantly reduced to 5.3% of the scale range in the monetary incentives condition.

Bulllock, Gerber, Hill, & Huber, 2015, Study 2	Partisans from Amazon Mechanical Turk sample, N = 795 (Excludes respondents in the unanalyzed second control group).	March - April, 2012	Prior to the manipulation, respondents answered five political questions (randomly selected from a pool of 11), and responses were used to identify items with partisan division. Then respondents answered seven political questions (two new and the initial five again) under one of four conditions: payment for each correct answer (amounts randomly varied from $0.10 to $1.00), payment for each correct answer (amounts randomly varied from $0.10 to $1.00) and for 'don't know' responses (amounts randomly varied from 20-33% of correct response payment), control (no payment) condition without a 'don't know' option, and (not included in the analyses) control (no payment) condition with a 'don't know' option. Partisan divergence was operationalized as between-party difference in scale scores (coded so higher score means a more Democrat-friendly response), without regard for correctness and with 'don't know' responses coded at the average pre-treatment position (i.e., non-polarized).	The analyses excluded the control condition with a 'don't know' option. When paid for 'don't know' answers, 46%, 47%, and 50% selected 'don't know' when paid 20%, 25%, and 33% of the correct answer amount, respectively. Across the eligible questions, partisan divergence averaged 14.5% in the control condition without a don't know option, which was significantly reduced to 5.8% of the scale range in the paid correct condition and 2.8% in the paid for correct and 'don't know' condition (in which large numbers selected 'don't know' and were placed in non-polarized positions).

Bulllock, Gerber, Hill, & Huber, 2015, one-item replication of Study 2	Partisans from demographically representative sample selected from opt-in internet panel for 2012 CCES (YouGov), N = 593	October-November, 2012	Single-item replication of Study 2 involving a rating of change in unemployment rate during President Obama's first term. A 'don't know' option was offered in the no incentive condition but respondents who selected it were treated as missing data, while those who selected 'don't know' in the condition that incentivized 'don't know' responses were placed in the non-polarized position of the scale mean.	Partisan divergence averaged 36.6 percent of the scale range in the control condition, which was near-significantly reduced (one-tailed p-value <.10) by 13.2 percentage points in the paid correct condition and significantly reduced by 22.2 percentage points in the paid for correct and 'don't know' condition (in which 9.2% selected 'don't know' and were placed in non-polarized positions, compared to 6.7% who selected don't know in the no incentive condition who were treated as missing data).
Khanna & Sood, 2018, Study 1	Amazon Mechanical Turk sample, N = 686 individuals who reported a position on concealed carry ban and who performed well on an initial numeracy test	December-January, 2013-14	Respondents viewed data in a table indicating the results of a study addressing whether concealed carry bans are associated with higher or lower crime, with results (higher vs. lower crime) randomly varied. Respondents were then asked whether the results showed higher or lower crime and were randomly assigned to be paid $0.10 for a correct answer or to a control condition with no financial reward for answering correctly.	Results of three studies reported together (see Study 3)

Khanna & Sood, 2018, Study 2	Amazon Mechanical Turk sample, N = 604 who indicated positions on the relevant issues.	March-April, 2013	The concealed carry experiment from Study 1 was repeated with a new sample with changes in the visual presentation to improve understandability. Respondents were also administered a second experiment of the same design dealing with the association between minimum wage increases and employment.	Results of three studies reported together (see Study 3)
Khanna & Sood, 2018, Study 3	Qualtrics sample, N = 1,055 who indicated positions on the relevant issues	August, 2016	The concealed carry and minimum wage experiments were repeated with a new sample, but only the uncongenial information conditions were run in the concealed carry experiment.	Results for Studies 1-3 were presented together. Incentives never impacted percentage correct when results were politically congenial, as expected. For the concealed carry experiments, pooling results from Studies 1-2, concealed carry supporters answered correctly 42.6% of the time without incentives and 40.1% of the time with incentives (a non-significant difference), while concealed carry opponents answered correctly 41.0% of the time without incentives but 57.0% of the time with incentives (a significant increase). For the concealed carry experiment in Study 3, concealed carry supporters answered correctly 33.6% of the time without incentives and 33.3% of the time with incentives (a non-significant difference), while concealed carry opponents answered correctly 25.6% of the time without incentives and 32.5% of the time with incentives (a significant increase). For

				the minimum wage experiments, pooling across Studies 2-3, minimum wage opponents answered correctly 62.0% of the time without incentives and 70.4% of the time with incentives (a significant increase) while minimum wage supporters answered correctly 87.0% of the time without incentives and 86.3% of the time with incentives (a non-significant difference).
Schaffner & Luks, 2018	Demographically representative sample selected from opt-in internet panel (YouGov), N = 651	January, 2017	Respondents were presented with a photo of the crowd from Obama's 2009 inauguration alongside one from the crowd from Trump's 2017 inauguration. Photos were labeled as Image A and Image B. Obama's inauguration clearly involved a substantially larger crowd. Respondents were randomly assigned to either be asked which photo showed a larger crowd or which photo went with which inauguration.	Among respondents asked which photo contained the larger crowd, 3% of non-voters, 2% of Clinton voters, and 15% of Trump voters selected the obviously wrong answer. Among respondents asked which photo went with which inauguration, 21% of non-voters, 8% of Clinton voters, and 41% of Trump voters selected the incorrect answer. This latter finding, as the authors note, could reflect sincerely believed misinformation. In both conditions, Trump voters were significantly more likely to select the wrong answer than the other groups.

Robbett & Matthews, 2018	Partisans from Amazon Mechanical Turk sample, N = 627	June, 2016	Respondents answered six factual political questions (from a pool of 10) and one factual neutral (i.e., not controversial or political) question (from a pool of 3) in a 3 X 3 between-subjects design. For the first independent variable, respondents were randomly assigned to a group size of 1, 5, or 25 and were incentivized to answer factual political questions correctly. Respondents assigned to a group size of 1 received $1 for every correct answer they offered individually. Those assigned to groups of either 5 or 25 "voted" for a correct answer, only receiving $1 when a majority of the group answered correctly. As for the second independent variable, respondents either received no information to help them answer correctly, had free information made available, or had costly information (costing $0.50 per question deducted from final earnings) made available. Partisan divide was operationalized as the between-party difference in proportion of responses that were pro-Republican (anti-Democrat).	When no information was provided, the partisan gap was about 5% of the scale range in the individual condition, and was significantly increased to about 13% of the scale range in both the 5-person and 25-person group vote conditions. Likelihood of correct politically uncongenial answers was 32% in the individual condition and was significantly reduced by about 10 percentage points in the 5-person group condition and by about 14 percentage points in the 25-person group condition. Individuals voting in groups were less likely to obtain costly information than individuals voting alone, and respondents were more likely to obtain information (costly and free) for neutral questions than political questions. Information availability did not reduce the partisan gap among individuals but reduced it by about half among voters.

Berinsky, 2018, Study 1	Partisans from demographically representative sample selected from opt-in internet panel (YouGov), N=1,000	January, 2012	Respondents were asked whether they believed the 9-11 attacks were an inside job by the G.W. Bush administration and whether they believed Barack Obama is a Muslim. Respondents were randomly assigned to either receive standard instructions (control condition) or instructions to exclude their personal feelings about people and policies, and simply indicate what they believe to be true (exclusion condition).	With regard to the 9-11 question, 22% of Democrats reported belief in the rumor in the control condition and 19% did so in the exclusion condition, and 22% reported "unsure" in the control condition compared to 23% in the exclusion condition (both non-significant differences). With regard to the Obama religion question, 32% of Republicans reported belief in the rumor in the control condition and 34% did so in the exclusion condition, and 35% reported "unsure" in the control condition compared to 37% in the exclusion condition (both non-significant differences).
Berinsky, 2018, Study 2	Partisans from demographically representative sample selected from opt-in internet panel (Survey Sampling International) , N = 843	March, 2014	Respondents were asked the same rumor questions as in Study 1, but were randomly assigned to either standard instructions (control condition) or instructions stating, "We sometimes find that people who do not believe these statements say they do believe them so they can say something bad about the people and policies mentioned in these statements" and urging respondents to answer truthfully (subtle pipeline condition).	With regard to the 9-11 question, 21% of Democrats reported belief in the rumor in the control condition and 20% did so in the subtle pipeline condition, and 18% reported "unsure" in the control condition compared to 14% in the subtle pipeline condition (both non-significant differences). With regard to the Obama religion question, 34% of Republicans reported belief in the rumor in the control condition and 34% did so in the subtle pipeline condition, and 26% reported "unsure" in the control condition compared to 21% in the subtle pipeline condition (both non-significant differences).

Berinsky, 2018, Study 3a	Republicans from demographically representative sample selected from opt-in internet panel (Survey Sampling International), N = 430	August, 2013	Respondents were asked the Obama Muslim rumor question from Studies 1-2 and were randomly assigned to answer the question with standard instructions (control condition), with instructions indicating there would be 5 minutes of additional questioning if rumor is endorsed (encouraged to reject rumor condition), or with instructions indicating there would be 5 minutes of additional questioning if rumor is rejected (encouraged to accept rumor condition).	44.4% endorsed the rumor in the control condition, compared to 47.7% in the encouraged to accept condition and 41.6% in the encouraged to reject condition, neither constituting a significant difference from the control condition. Based on particular assumptions, estimated that 3.3% (47.7-44.4) of respondents were willing to say anything to receive the incentive, and subtracting this from the 2.8% effect of being encouraged to reject the rumor (44.4-41.6) estimated that all of the apparent expressive responders were actually just willing to say anything.
Berinsky, 2018, Study 3b	Democrats from demographically representative sample selected from opt-in internet panel (Survey Sampling International), N = 713	August, 2013	Same design as Study 3a, but using a question about the 9-11 rumor and Democratic respondents.	35.7% endorsed the rumor in the control condition, compared to 41.3% in the encouraged to-accept condition and 26.7% in the encouraged to reject condition, neither constituting a significant difference from the control condition. Based on particular assumptions, estimated that 5.5% (41.3-35.7, with rounding error) of respondents were willing to say anything to receive the incentive, and subtracting this from the 9.0% effect of being encouraged to reject the rumor estimated that 3.5% of Democrats engaged in expressive responding.

Berinsky, 2018, Study 4	Partisans from national probability sample (Knowledge Networks), N=4,770	December-January, 2011-12	Conducted a list experiment (unmatched count technique) with respondents randomly assigned to one of four conditions. In the base-list condition, respondents indicated the number of statements they agreed with out of four. In the additional treatment item condition, a fifth statement was added, "I believe Barack Obama is a Muslim." In the remaining two conditions, respondents were directly asked if they believed Barack Obama was a Muslim, either with or without a "not sure" response option (only the direct ask without unsure option was included in the analyses).	Among Republicans, the mean number of items agreed with was 2.02 in the base-list condition and 2.40 in the additional treatment item condition, yielding an estimate that 37.3% of Republicans indicate belief Obama is a Muslim when they have privacy in doing so. In the direct ask condition, 44.0% of Republicans endorsed the rumor, yielding an estimate that 6.7% of Republicans are expressively responding. Among Democrats, the mean number of items agreed with was 2.09 in the base-list condition and 2.14 in the additional treatment item condition, yielding an estimate that 4.2% of Democrats indicate belief Obama is a Muslim when they have privacy in doing so. In the direct ask condition, 12.6% of Democrats endorsed the rumor, yielding an estimate that 8.4% of Democrats (surprisingly) gain expressive benefits from saying Obama is a Muslim.

Allcott, Boxell, Conway, Gentzkow, Thaler, & Yang, 2020	Demographically representative sample selected from opt-in internet panel (Prime Panels from CloudResearch), N = 2,000	April, 2020	Respondents were asked to predict the number of new COVID-19 cases that would be confirmed in the United States in April of 2020 and the percentage of the population that would approve of Trump's handling of the COVID-19 pandemic in the last poll of April. Respondents were randomly assigned to either be informed that 10 respondents would be selected for each prediction to receive a payment of $100 minus the absolute percentage point difference between their guess and the correct value, or to a control condition with no financial reward for close predictions. Partisan difference was operationalized as the effect of party identification (7-point scale) on standard score estimates.	The effect of party identification (going from Strong Democrat to Strong Republican) on standard-scored COVID-19 prediction was -.14 without incentives and -.33 with incentives, a non-significant difference in the opposite direction of what would be expected based on expressive responding. This effect on standard-scored Trump approval prediction was -1.14 without incentives and was significantly reduced to -.59 with incentives.

Yair & Huber, 2020, Study 1	Partisans from Amazon Mechanical Turk Sample, N = 502	July, 2017	Respondents viewed a photo of an opposite-sex person with "About me" information presented underneath, including "friendly," "smart," and "runner." For one independent variable, respondents were randomly assigned to a control condition in which this was the only information presented, or to one of four experimental conditions in which "Republican," "supported Trump in the 2016 election," "Democrat," or "supported Clinton in the 2016 election" was presented. For the other independent variable, respondents were randomly assigned to simply rate the attractiveness of the individual, first "blow-off-steam" by expressing partisan sentiment about the person before making the attractiveness rating, or first receive a "warning" that they would have an opportunity to express a partisan message later before making the attractiveness rating.	Results of three studies reported together (see Study 3)

Yair & Huber, 2020, Study 2	Partisans from non-representative Lucid sample, N = 766	August, 2017	Same design as Study 1	Results of three studies reported together (see Study 3)
Yair & Huber, 2020, Study 3	Partisans from demographically representative sample selected from opt-in internet panel (Lucid), N = 1,006	September-October, 2017	Same design as Studies 1 and 2, except two additional treatments were fielded but not reported in the results because they did not produce baseline differences in attractiveness ratings (see footnote 5, p. 476)	The results of the three studies were pooled. When just making attractiveness ratings, Democrats rated matched profiles (Democrat or Clinton supporter) as no more attractive than profiles with no political information, but mismatched profiles (Republican or Trump supporter) as less attractive (just under one scale point on a 7-point scale) than profiles with no political information. Republicans' attractiveness ratings were not influenced by partisan match or mismatch. The blow-off-steam and warning treatments, combined, reduced Democrats' partisan bias (matched vs. mismatched) by 48% and did not significantly affect Republicans' partisan bias.
Connors, 2020	Partisans from non-representative Prolific sample, N = 939	December, 2020	Respondents answered questions about belief that there was a lot of fraud in the 2020 election and belief that the election results should be accepted. Respondents were randomly assigned to be instructed to either answer these questions in the manner that they thought a co-partisan would answer when trying to impress	Republicans indicated that they would expect Republicans trying to impress co-partisans to report belief in sizable fraud and non-acceptance of election results to a greater extent than Republicans trying to disappoint co-partisans. Democrats indicated that they would expect Democrats trying to impress co-partisans to report belief in sizable fraud and non-acceptance of election results to a lesser

			co-partisans or in the manner they thought a co-partisan would answer when trying to disappoint co-partisans.	extent than Democrats trying to disappoint co-partisans
Graham & Huber, 2020	Demographically representative sample selected from opt-in internet panel (Lucid), N = 3,813	December - January, 2019-2020	After answering initial questions about demographics, political attitudes, and other topics, respondents were given an option to answer five extra questions or to proceed to the last question. Respondents were randomly assigned to a description of the extra questions as a "National Poll," intended to increase the value of self-expression, or a "Test Survey," to reduce expressive rewards. Respondents were also randomly assigned to view a "teaser" question that was either neutral (about obscure matters difficult to connect to partisanship), partisan (provided clear party or candidate cues), or about a political rumor. For respondents who opted to answer five extra questions before the final question, the last six questions included the teaser	Overall, 64% of respondents opted to answer extra questions. Percentages opting to do so were 51 and 59% among those receiving one of the two neutral teasers, 62 and 77% among those receiving one of the two partisan teasers, and between 58-70% among those receiving one of the four rumor teasers. Various measures of strong partisanship, inclination to give partisan answers to rumor questions, and political engagement were positively associated with opting to answer extra questions. Description of the extra questions as a "National Poll" vs. "Test Survey" did not impact the proportion opting to answer extra questions. By far the most commonly selected reason for answering extra questions (64%) was "I was curious about what you would ask." The next most common answers also did not have to do with self-expression per se but with enjoying taking surveys in general, getting extra pay, being helpful, and

			question, three additional rumor questions, and two debriefing questions about why they chose to answer more questions. Respondents who opted to skip to the final question received a single question about a rumor.	enjoying answering questions like the one's shown (33-37%). Two clearly expressive options, standing up for what one believes and wanting to let other people know what one believes, were selected by 26 and 19% of respondents, respectively.

Peterson & Iyengar, 2021a, Study 1	Partisans from demographically representative sample selected from opt-in internet panel (YouGov), N=875	September, 2018	Respondents answered five factual questions that were mostly about matters of substantial partisan disagreement and were randomly assigned to receive $0.50 per correct answer or to a control condition with no financial reward for answering correctly. Before answering each question, respondents were allowed to read information about the topic from one of five sources -- one right-leaning, one left-leaning, one with specialized expertise, and two mainstream. Respondents then answered the question and rated the certainty of their answer. A subset of respondents had their web-browsing activity tracked. Partisan divide was operationalized as the between-party difference in proportion of responses that were pro-Democrat (anti-Republican).	Results of two studies reported together (see Study 2)
Peterson & Iyengar, 2021a, Study 2	Partisans from demographically representative sample selected from opt-in internet panel (YouGov), N=1,507	March, 2019	Same design as Study 1	Results were pooled across the two studies. With no incentives, pooled across all items, the partisan gap averaged 32% of the scale range. This was significantly reduced to 22% of the scale range in the incentives condition. Incentives for correct answers did not reduce selection of co-partisan news

				sources relative to the unincentivized condition or relative to real-world web-browsing behavior.

Peterson & Iyengar, 2021b	Partisans from demographically representative sample selected from opt-in internet panel (YouGov), N=1,447	July, 2020	Respondents answered five factual questions about the COVID-19 pandemic and were randomly assigned to receive either $0.25 or $1.00 per correct answer, or to a control condition with no financial reward for answering correctly. Before answering each question, respondents were allowed to read information about the topic from one of five sources -- one right-leaning, one left-leaning, one with public health expertise, and two mainstream. Respondents then answered the question. Partisan divide was operationalized as the between-party difference in proportion of responses that were pro-Democrat (anti-Republican). In addition, respondents answered five factual questions about matters of partisan disagreement unrelated to COVID, under the same experimental treatment as with the COVID questions. Partisan divide was operationalized in the same way.	With no incentives, pooled across all COVID items, the partisan gap averaged 21 percent of the scale range. This was reduced to 19% in the low incentive condition and 18% in the high incentive condition, neither constituting a significant reduction. The authors also reported effects of incentives on correct responses to COVID questions. With no incentives, the partisan gaps in correct responses to the three COVID questions whose answers favored Democrats were 50, 39, and 35 percentage points; and on the two COVID questions whose answers favored Republicans these gaps were 18 and 0 percentage points. Respondents in the low incentive condition became one percentage point likelier to answer correctly (a non-significant change) and respondents in the high incentive condition became three percentage points likelier to answer correctly (a significant change). Next, with no incentives, pooled across all non-COVID items, the partisan gap averaged 32 percent of the scale range. This was reduced to 23 percentage points in the low incentive condition and to 26 percentage points in the high incentive condition, both significant reductions. The authors also reported effects of incentives on correct responses to non-COVID questions. With no incentives,

				the partisan gaps in correct responses to the four non-COVID questions whose answers favored Democrats were 48, 45, 33, and 11 percentage points; and on the one non-COVID question whose answer favored Republicans this gap was 24 percentage points. Respondents in the low incentive condition became 5 percentage points likelier to answer correctly and those in the high incentive condition became 4 percentage points likelier to answer correctly, both significant increases. Neither low nor high incentives reduced selection of co-partisan news sources relative to the unincentivized condition.

Section 2
Additional Discussion of Expressive Responding, Accuracy-Motivated Heuristic Use, Top-of-the-Head Models of Survey Response, and Uncertainty

As described in the main text, we define expressive survey responding as report of expressively rewarding politically congenial answers to survey questions when one has privately constructed a different answer (cheerleading), or when one does not know the answer and defaults to the expressively rewarding congenial answer or samples considerations from memory in a way that is biased toward expressing support for one's side (congenial inference). This is consistent with the core features of the definitions that have appeared in the expressive responding literature, although in some instances in the literature definitions of congenial inference also seem to include accuracy-motivated party-heuristic use (see Table S-1).

As Bullock and Lenz (2019, pp. 327-328) note in their recent review, both cheerleading and congenial inference constitute "directionally motivated" forms of survey responding that are aimed at expressing support for one's side in a survey in a way that misrepresents one's privately held belief or awareness of one's own ignorance:

> Directional motives . . . may operate in more than one way. If our Republican respondent is confident that the unemployment rate declined under Obama, a directional motive may simply incline him to give a response that he believes to be false. This is insincere responding, or *cheerleading*. On the other hand, our Republican respondent may not be confident of the correct answer. In this case, he may canvas his memory for considerations related to the unemployment rate—but do so in a way that makes him especially likely to retrieve considerations that cast Obama in a negative light. This is biased consideration sampling. Alternatively, the Republican may rely on a pro-party heuristic (e.g., Khanna & Sood 2018) to determine his answer: for example, a heuristic that says "give the response that makes Obama look bad unless faced with incontrovertible contrary evidence." Following Prior et al. (2015, especially p. 494), we use the term *congenial inference* to refer to either biased consideration sampling or the use of pro-party heuristics.

Our conceptualization of expressive responding encompasses these two forms of directionally motivated misrepresentation in the survey response process. This is consistent with the central notion of expressive responding as a phenomenon that creates a disjunction between an expressively rewarding survey report and privately acknowledged belief, attitude, or ignorance.

This type of directionally motivated responding may be contrasted with accuracy motivated responding. Accuracy motivated responding may involve faithfully reporting a confidently and sincerely held belief. But accuracy motivated responding may also involve use of pro-party heuristics, in a way that differs subtly from directionally motivated pro-party heuristic use. As Bullock et al. (2019, p. 329) note:

> Given low levels of knowledge and confidence, it is even possible that accuracy motivations lead to partisan differences. This possibility arises when people have

> so little confidence in the correct answers to questions that a party default provides their best guess. When answering a question about how inflation changed under President Reagan, for instance, the only consideration that may come to mind is that Reagan was a Republican. Republican respondents may then report that inflation fell under Reagan because they think Republican officeholders generally perform well, while Democrats may report that inflation rose because they believe the opposite about Republican officeholders.

We do not regard application of an accuracy-motivated party heuristic as part of the expressive responding phenomenon. This is because relying on such a heuristic when forming a survey response entails responding on the basis of sincere – albeit not well-considered or long-held – belief with the goal of accuracy, not misrepresenting a privately held view in order to gain expressive benefits.

Placed in a more general context, it is important to consider how expressive responding relates to "top-of-the-head" models of the survey response process (Tourangeau, Rips, & Rasinski, 2000; Zaller, 1992). This influential view of the survey response process depicts respondents as forming their attitudes and beliefs on the fly while they are answering a survey question, based on considerations that they can retrieve from memory while they are in the process of answering that question. To the extent that the response process unfolds in this way, it is not obvious what it would mean for a respondent to believe one thing but say another. We have two things to say about this. First, if a respondent's political bias has caused them to internalize a one-sided set of *sincerely held* considerations that align with their political allegiance, then a response constructed from such considerations would not qualify as an expressive response. The same may be said when an uncertain respondent applies an accuracy motivated pro-party heuristic based on a sincere belief that their own party and its elites are trustworthy and competent. Second, in other circumstances the respondent's goal of expressing support for their side might guide the process of interpreting the question, retrieving considerations from memory, and using considerations to form a judgment (Green, Kingzette, Minozzi, & Neblo, 2020). This might occur while the respondent is fully aware of not sincerely believing the answer they have assembled for its expressive value, while the respondent is uncertain of the correct answer and opts to default to the expressively rewarding politically congenial view, or while the respondent truncates the information search and rapidly settles on politically congenial view because doing so is expressively rewarding. In all cases, this would fall within the domain of expressive survey responding even though the reported "beliefs" are assembled on the fly.

Much of the above discussion focuses on how expressive response relates to the phenomenon in which a partisan is uncertain of the correct answer and reports a politically congenial factual misperception. The vast majority of studies on factual beliefs do not gauge respondents' certainty about their answers. Graham (2020) recently reported the most detailed test to date of the relationship between certainty and accuracy of factual answers, and how this differs depending on whether the questions are politically neutral, favorable to one's party, or unfavorable to one's party. In general, people who were less certain of their answers were more likely to get those answers wrong. This relationship between certainty and accuracy was very strong when the questions were either neutral or favorable to one's party, but weaker and less reliable when correct answers were unfavorable to one's party. That said, even on party-

unfavorable questions, incorrect politically congenial answers were more often rated as low-certainty guesses than high certainty beliefs.

In and of itself, level of certainty of one's reported politically congenial error says little about whether the reported wrong answer is an expressive response. Take a respondent who offers a politically congenial incorrect answer in which they have low certainty. This may be because they do not know the answer and wish to take the opportunity to express support for their team, perhaps after a biased party-serving information search. Or this may be because they have applied a pro-party heuristic in an attempt to answer correctly with a best guess but are uncertain of whether they have succeeded. Now consider a respondent who reports a politically congenial incorrect answer with high certainty. This respondent might know that the politically congenial answer is wrong or that they do not really know the answer, but report this politically congenial answer with high certainty as a way of enthusiastically expressing support for their side. Or they might hold the politically congenial incorrect answer with great sincerity and certainty.

Overall, then, the presence of low certainty answers to factual questions, on-the-fly construction of survey responses, and accuracy-motivated party heuristic use make it difficult to empirically discern expressive responding from report of sincere but poorly considered or low-confidence belief (Bullock & Lenz, 2019). Indeed, we will echo arguments that certain findings that have been interpreted as evidence of expressive responding might in part reflect the presence of this form of belief. In addition, we will argue that if an individual is motivated to express support for their side in a factual report, even if they might acknowledge the uncertainty of this belief, the way they act in the real world will likely be influenced by the very predisposition that made the survey report rewarding in the first place.

Section 3: 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.

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.

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.

Connors, E. C. (2020, December 22). *Do Republicans really believe the election was stolen – or are they just saying that?* Washington Post. https://www.washingtonpost.com/politics/2020/12/22/do-republicans-really-believe-election-was-stolen-or-are-they-just-saying-that/

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.

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.

Tourangeau, R., L. Rips, J., & Rasinski, K. (2000). *The Psychology of Survey Response.* Cambridge: Cambridge University Press.

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*, 469–492.