Public Opinion Quarterly (2025) **Vol 00 No 00,** 1–25

A Measurement Gap? Effect of Survey Instrument and Scoring on the Partisan Knowledge Gap

Lucas Shen[1], Gaurav Sood[2], Daniel Weitzel[3,*]

[1]Senior Scientist, Institute for Human Development and Potential, Agency for Science, Technology and Research, Singapore, Singapore
[2]Independent Researcher, Seattle, WA, US
[3]Assistant Professor, Department of Political Science, Colorado State University, Fort Collins, CO, US

Abstract Research suggests that partisan gaps in political knowledge with partisan implications are wide and widespread in the United States. Using a series of experiments, we estimate the extent to which the partisan gaps in commercial surveys reflect differences in confidently held beliefs rather than motivated guessing. Knowledge items on commercial surveys often have guessing-encouraging features. Removing such features yields scales with greater reliability and higher criterion validity. More substantively, partisan gaps on scales without these "inflationary" features are roughly 40 percent smaller. Thus, contrary to some prior research, which finds that the upward bias is explained by the knowledgeable deliberately marking the wrong answer (partisan cheerleading), our data suggest that partisan gaps on commercial surveys in the United States are strongly upwardly biased by motivated guessing by the ignorant. Relatedly, we also find that partisans know less than what toplines of commercial polls suggest.

Wide and widespread partisan gaps in political knowledge challenge the idea that citizens can hold representatives accountable (Hochschild and Einstein 2015; Bailey 2021). Hence, the alarm over research that finds as much (Campbell et al. 1980; Bartels 2002; Jerit and Barabas 2012) (though see Roush and Sood 2023). However, an emerging line of research argues that a large fraction of the partisan knowledge gap is an artifact of the survey response process (Bullock et al. 2015; Prior, Sood, and Khanna 2015; Yair and Huber 2020; Graham and Yair 2025) (though see Berinsky 2018; Peterson and Iyengar 2021; Malka and Adelman 2023). In this paper, we extend this investigation.

*Corresponding author: Daniel Weitzel, Department of Political Science, Colorado State University, Fort Collins, CO, US; email: daniel.weitzel@colostate.edu.

https://doi.org/10.1093/poq/nfaf044

Our starting point is commercial polls in the United States. In particular, we examine how common features of knowledge items on commercial polls, for example, presenting social proof about the less socially desirable option, including information about the topic, not including a "Don't know" option, providing a partisan cue in the question stem, and so on, affect estimates of partisan gaps in knowledge of facts with political implications. Removing these guessing-encouraging features yields knowledge scales with greater reliability and higher criterion validity. When we measure partisan gaps using these scales, they are 14 percentage points smaller (figure 1). To further ablate response biases, we use an instrument and scoring scheme inspired by Pasek, Sood, and Krosnick (2015) and Graham (2023) that considers respondents' confidence in their answers. Using the scoring scheme that

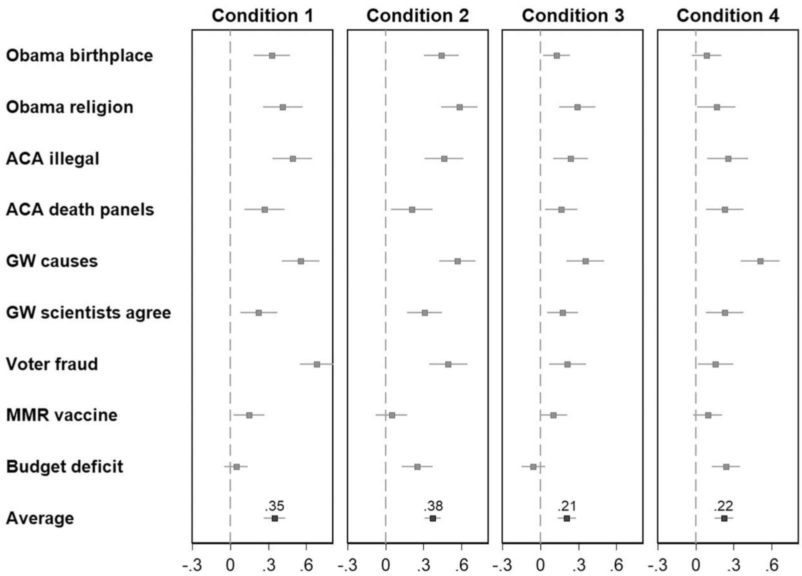

Figure 1. Partisan gap by treatment arm (MTurk 1). The figure shows the estimated partisan gap in each of the nine knowledge items (see Supplementary Material section 3 for the details of the items) and the average partisan gap across the four conditions (table 1). The partisan gap is estimated using the linear model Correct response$_i = \alpha + \beta$congenial$_i + \varepsilon_i$, where "congenial" is a dummy variable that takes the value 1 when the correct response is congenial to the respondent's party. All four columns have the same horizontal axis scale. Horizontal bars are 95 percent confidence intervals constructed from robust standard errors.

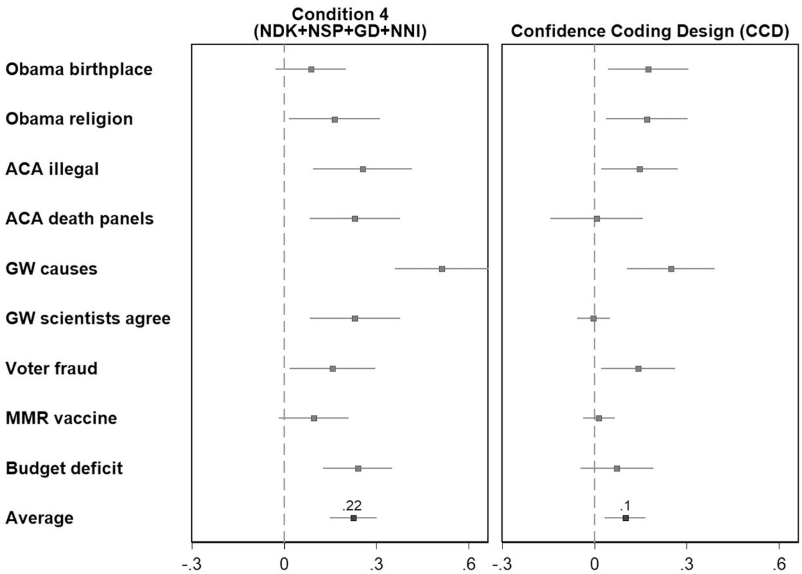

Figure 2. Partisan gaps in knowledge in different question designs. The figure shows the estimated partisan gaps in knowledge from MTurk 1 for two different survey conditions. The CCD condition only considers selecting the right answer with complete confidence as evidence that the respondent knows the answer (see Supplementary Material section 5). See tables 2.1 to 2.5 in Supplementary Material section 2 for the regression estimates of the multiple-choice conditions to the confidence coding condition. See Supplementary Material figure 2.6 for the same analysis with all four multiple-choice conditions pooled together. Supplementary Material figure 6.1 implements a robustness check, setting the relative scoring threshold c to 8. Horizontal bars are 95 percent confidence intervals constructed from robust standard errors.

credits only confidently held beliefs as knowledge, we find that partisan gaps are 50 percent smaller still (see figure 2).

Our results contribute to a growing literature that suggests that a large fraction of partisan gaps are artifacts of survey design. The results further clarify the source of bias in estimates of partisan gaps. While some previous research shows that the partisan gap is due to partisan cheerleading—deliberate selection of congenial incorrect answers by the knowledgeable (Prior, Sood, and Khanna 2015)—our data suggest that the bias in the estimate of the partisan gap is primarily a result of partisan guessing by the ignorant (see also Bullock et al. 2015, who reach similar conclusions).

Our results suggest that some concerns about democratic health are overstated and some are underappreciated. Reducing guessing-related error reveals that partisan gaps on partisan knowledge items are not as wide, but also that partisans know less about politics than what the toplines of commercial polls suggest.

Theory and Motivation

> "Has unemployment increased, decreased, or stayed the same since President Joe Biden took office in 2021?"

How knowledge about this fact and other such politically consequential facts is distributed across the population is relevant to the health of a democracy. If there are wide gaps in partisans' knowledge of politically relevant facts, citizens' ability to hold politicians accountable might be limited.

Concerningly, a large body of research finds that partisan gaps in political knowledge with partisan implications are both wide and widespread (Bartels 2002; Jerit and Barabas 2012; Lodge and Taber 2013; LaLoggia 2018; though see Roush and Sood 2023). Some recent research, however, shows that a large part of the partisan gaps stems from partisan responding rather than differences in what partisans know to be true about the world (Bullock et al. 2015; Prior, Sood, and Khanna 2015; Yair and Huber 2020; Graham and Yair 2025) (though see Peterson and Iyengar 2021 and Malka and Adelman 2023).

More generally, researchers argue that partisan gaps in political knowledge with partisan implications are inflated by:

1. **Partisan Cheerleading:** Partisans who know the right uncongenial answer deliberately pick the wrong partisan congenial answer to register their support for their party or to influence the survey results (Prior, Sood, and Khanna 2015).
2. **Partisan Guessing:** Partisans who don't know the answer offer substantive responses congenial to their party (Bullock et al. 2015; Graham and Yair 2025). For instance, when asked about what happened to the unemployment rate during the Biden administration, Republicans ignorant about the unemployment rate may still respond that unemployment rose during the Biden administration because they viscerally dislike Democrats or because they believe that Democrats mishandle the economy.

In this paper, we interrogate the latter explanation in the context of commercial polls. An analysis of 180 media polls by Luskin et al. (2018) found that guessing-encouraging features were exceedingly common. For instance, less than 9 percent of the surveys offered an explicit "Don't Know" or "Not Sure" option, which causes a positive bias in the estimates of political knowledge (Luskin and Bullock 2011; Cor and Sood 2016). And about half of the items offered only two choices, a design choice that dramatically inflates estimates of knowledge (Fortin-Rittberger 2016; Bullock and Rader 2022). An overwhelming majority of the items (168) also included wording encouraging guessing by framing the factual question as a "matter of opinion." They also found that the scoring rules used by analysts treated all correct responses—even when the respondent is unconfident about their answer—as evidence of knowledge. Doing so conflates guesses and on-the-spot inferences with knowledge (Pasek, Sood, and Krosnick 2015). Other research finds that the partisan context of the survey cues directional motivations and increases the partisan gap (Prior, Sood, and Khanna 2015; Bailey 2021).

Guessing vs. Diffidence

Survey measurement of political knowledge using closed-ended items faces two competing challenges: respondents marking a substantive option when they don't know the correct answer and respondents not marking the correct answer when they do know. If you take each incorrect answer in a multiple-choice item as evidence of a random guess by a respondent who doesn't know the answer, the percentage of correct answers that stem from guessing on political knowledge questions is 22 percent (Luskin and Bullock 2011).[1] Guessing-related error doesn't just distort the description of how much people know but also abrades correlations with the latent construct. The rationale is as follows: the less you know, the more items you must guess on, and the more the positive error in your score absent guessing adjustment in scoring (Cor and Sood 2016). In effect, the guessing-related error is negatively correlated with the latent construct (knowledge). These concerns motivate many researchers to look for a solution. Some researchers argue that offering a "Don't Know" option is a reasonable solution. They point out that "Don't Know" responses hide very little knowledge (Sanchez and Morchio 1992; Sturgis, Allum, and Smith 2008; Luskin and Bullock 2011; Jessee 2017); the tendency to mark "don't know" doesn't vary by gender (Ferrín, Fraile, and García-Albacete 2017) or personality (Jessee 2017), and hence produces more descriptively and correlationally valid estimates (Luskin and Bullock 2011; Jessee 2017). Others, however, contend that the cure is worse than the problem. They implicitly argue that Don't Know responses hide a lot of knowledge and that the amount of knowledge that Don't Know responses hide varies by the type of person (Mondak 1999; Mondak and Anderson 2004; Dolan and Hansen 2020; Kraft and Dolan 2023).

Don't Know is but one way to reduce guessing-related errors. Other ways include removing social proof and neutral information from the question

1. In the DK Discouraging condition, the percentage correct is 60.7. Correcting for guessing reduces the number to 47.1.

stem, using self-assessed confidence, and increasing the number and difficulty of options (Fortin-Rittberger 2016; Bullock and Rader 2022).

We hypothesize that guessing-encouraging features inflate partisan gaps and yield measures with lower correlational and descriptive validity.[2]

Empirical Strategy

To study the effect of “inflationary” features of survey and knowledge items on the partisan knowledge gap, we conduct a series of survey experiments that modify various guessing-encouraging features. To study the effect of taking respondents’ confidence into account, we draft an instrument and scoring rule inspired by Pasek, Sood, and Krosnick (2015), which uses self-assessed confidence to rescore the answers, taking only correct answers that respondents are confident about as evidence that the respondent knows the fact. We also analyze which item formats produce measures with greater reliability and higher criterion validity. In all, we use data from four surveys. The results of these four surveys are presented as part of three studies:

- In Study 1, we use data from a survey experiment conducted on Amazon Mechanical Turk (MTurk) (MTurk 1) to examine how guessing-encouraging features affect the partisan gap.
- In Study 2, we use survey experiments conducted on a YouGov and a telephone survey (Texas Lyceum) to examine the effect of partisan cues on the partisan gap.
- Finally, in Study 3, we use data from MTurk 1 and another survey fielded on MTurk (MTurk 2) to study the impact of taking respondents’ confidence in their answers into account on the partisan gap.

Before we proceed further, we would like to note that many of our questions are on topics on which people can be misinformed—and know the wrong thing confidently. This includes partisan retrospection items like those used by Bartels (2002). However, on all of these “misinformation” items, we can also ask how many people know the correct answer. Like Bartels (2002) and Prior, Sood, and Khanna (2015)—and for much the same reasons—we are interested in measuring the partisan gap in knowledge, though we believe that it would also be useful to study partisan gaps in misinformation.

2. In Supplementary Material 8, we analyze the impact of removing social proof, neutral information about the topic, and so on, on the gender gap. In Supplementary Material 10, using data from Bullock and Rader (2022), we assess the impact of increasing the difficulty and number of choices on the gender gap. The results are inconsistent, suggesting that the effect is unlikely to be large and consistent.

Study 1: The Effect of Guessing-Encouraging Features

The first study focuses on four survey design features that we suspect inflate the partisan gap. These features are:

1. the absence of a "Don't Know" (DK) option
2. including additional neutral information in the question stem
3. providing social proof for an answer
4. the absence of a guessing discouraging preamble

Research Design and Data

We conducted a survey experiment on MTurk on July 9, 2017, in which we randomly assigned 1,253 respondents to one of four conditions (see table 1).[3] The data are unweighted. In each condition, respondents answered nine misinformation items, ranging from President Obama's citizenship to whether global warming is happening or not. (For exact question wording, see Supplementary Material section 3.) The conditions reduce the number of inflationary features from four to zero and are labeled with the abbreviated features.

The four conditions are:

1. **Condition 1 (NDK+SP+GE+NI):** The design includes four features that encourage guessing. It serves as our baseline condition. The items in this condition include all the common features of commercial polls. In this design, the "Don't Know" option is never presented (NDK; prefix N indicates "Not presented"), so the respondents must guess if they don't know. The questions also include social proof about the incorrect answer (SP). By social proof, we mean information about what other people

Table 1. Experimental treatments.

		Treatments				
Condition	Label	Don't know	Social proof	Guessing encouraged	Neutral information	Inflationary features
1	NDK+SP+GE+NI	No	Yes	Yes	Yes	4
2	NDK+NSP+GE+NI	No	No	Yes	Yes	3
3	DK+NSP+GD+NI	Yes	No	No	Yes	2
4	DK+NSP+GD+NNI	Yes	No	No	No	0

3. For generalizability of effects in studies conducted on MTurk, see Mullinix et al. 2015 and Coppock, Leeper, and Mullinix 2018. Balance tests suggest that the randomization was successful (see Supplementary Material figures 1.1 to 1.4).

believe. Seeing that some people believe in an option can cause more people to select that option (see Sherif 1935; Cialdini 2009). For instance, on the question about where Obama was born, we add, "Some people believe Barack Obama was not born in the United States but was born in another country." In other cases, we provide some neutral information (NI) about the topic, like "According to the Constitution, American presidents must be natural-born citizens." Finally, the preamble to the knowledge questions is neutral and doesn't discourage guessing or cheating (GE). (Please see Luskin and Bullock 2011 for data on how the DK neutral preamble has the same effect as a DK discouraging one.) The preamble simply reads: "Now here are some questions about what you may know about politics and public affairs … "

2. **Condition 2 (NDK+NSP+GE+NI):** By removing social proof (SP), we arrive at a very commonly used design in commercial polling. Like the baseline condition, the questions in this design do not feature a "Don't Know" option (NDK) but include neutral information (NI) in the question stem.
3. **Condition 3 (DK+NSP+GD+NI):** The next design removes two inflationary features. First, the preamble now discourages blind guessing and cheating (GD). The preamble reassures respondents that it is okay not to know the answers to these questions, asks respondents to commit not to look up answers or ask anyone, and asks respondents to mark don't know when they don't know the answer. Second, the items now include a DK option (see, e.g., Luskin and Bullock 2011; Bullock et al. 2015). (We code DK the same as an incorrect answer.)
4. **Condition 4 (DK+NSP+GD+NNI):** Our final version offers respondents a DK option, discourages guessing and cheating (GD) and does not include neutral information (NI) or social proof (NSP) in the question stem. (We code DK the same as an incorrect answer.)

Measures

Since our study was fielded in the United States, we measure partisanship using the conventional branched seven-point partisan self-identification scale. Respondents are first asked if they identify as Republicans, Democrats, or Independents. If respondents pick a party, they are asked about the strength of their attachment to the party. Independents are asked if they lean toward one party or the other. In our study, Independents who lean toward one of the two major parties are coded as supporters of that party. A knowledge item is coded as congenial if the correct answer is congenial to the partisanship of the respondent.

Results

We start by summarizing the average partisan gap on each survey item in each treatment arm (see figure 1). In the baseline condition (Condition 1) with all inflationary attributes, when the correct response is congenial to the respondents' party, respondents are 35 percentage points more likely to choose the correct response. As figure 1 shows, the partisan gap is unresponsive to the removal of social proof in the question stem (Condition 2). However, the estimates in Condition 3 and 4 are approximately 14 percentage points lower than in the baseline condition. The 14-percentage-point reduction, stemming from the inclusion of guessing discouraging text and the exclusion of neutral information in the question stem, translates to a 40 percent relative drop ($100*\frac{.35-.21}{.35}$).

To formally test our hypothesis, we estimate the following equation:

$$\begin{aligned}\text{Correct}_{\text{ijk}} = \alpha + \beta\ \text{Congenial}_{\text{i}} \\ + \sum\nolimits_{k=1}^{4} \gamma\text{Condition}_{\text{k}} \\ + \sum\nolimits_{k=1}^{4} \delta_k(\text{Congenial}_{\text{i}}\ x\ \text{Condition}_{\text{k}}) \\ + \ \text{Question}_{\text{j}} + \varepsilon_{\text{ijk}}\end{aligned} \tag{1}$$

In the above equation, *i* iterates over respondents, *j* over survey item, and *k* over treatment arms. β captures the difference in the proportion of correct responses when the answer is congenial to the respondent's party in the baseline condition. δk's capture how Conditions 2–4 affect the partisan knowledge gaps versus the baseline condition. Table 2 reports the results. Column (1) includes just the congenial variable, which has a *p*-value < 0.001 and is consistent with conventional wisdom about gaps in partisan knowledge (e.g., Bullock et al. 2015; LaLoggia 2018).

Column (2) includes only the survey treatments. The negative coefficients on Conditions 3 (DK+NSP+GD+NI) and 4 (DK+NSP+GD+NNI) show that respondents' estimated knowledge is sharply lower in the two conditions, which remove first any encouragement to guess and then any neutral information, compared to the baseline. In column (3), we include the interaction between the congenial dummy and the three conditions (baseline is Condition 1 with all inflationary features). Now, the congenial variable captures the knowledge gap in the baseline condition (corresponding to column (1) of figure 1). The congenial and survey condition interactions reveal the extent to which partisan knowledge gaps change across the different survey conditions. The gap drops from 0.35 and 0.38 in the more inflationary designs to 0.21 and 0.22 in designs with fewer problematic features.

Table 2. The effect of various treatments on the partisan gap (MTurk 1).

	(1)	(2)	(3)	(4)	(5)	(6)
Congenial	0.281		0.351	0.284		0.353
	(0.017)		(0.035)	(0.017)		(0.034)
	[0.000]		[0.000]	[0.000]		[0.000]
Condition 2		0.010	0.000		0.011	0.002
		(0.028)	(0.022)		(0.028)	(0.021)
		[0.722]	[0.985]		[0.687]	[0.934]
Condition 3		−0.064	0.000		−0.063	−0.001
		(0.024)	(0.019)		(0.024)	(0.019)
		[0.009]	[0.993]		[0.010]	[0.964]
Condition 4		−0.080	−0.023		−0.079	−0.021
		(0.025)	(0.019)		(0.025)	(0.019)
		[0.002]	[0.245]		[0.002]	[0.281]
Congenial*Cond. 2			0.024			0.024
			(0.046)			(0.045)
			[0.605]			[0.601]
Congenial*Cond. 3			−0.173			−0.163
			(0.046)			(0.045)
			[0.000]			[0.000]
Congenial*Cond. 4			−0.132			−0.136
			(0.048)			(0.048)
			[0.006]			[0.005]
Constant	0.179	0.306	0.184	0.050	1.331	0.227
	(0.007)	(0.020)	(0.014)	(1.056)	(1.255)	(1.010)
	[0.000]	[0.000]	[0.000]	[0.962]	[0.289]	[0.823]
R^2	0.315	0.234	0.328	0.324	0.243	0.337
Survey item FE	Yes	Yes	Yes	Yes	Yes	Yes
Demographic controls	No	No	No	Yes	Yes	Yes
Items	9	9	9	9	9	9
Respondents	628	628	628	627	627	627
Respondent-items	5,652	5,652	5,652	5,643	5,643	5,643

Note: All models are linear probability models where the dependent variable is whether the response is correct. See table 1 for the description of the four conditions. Condition 1, with four inflationary features, is the baseline. Demographic controls include age, gender, education, and race. Coefficients are unstandardized; standard errors (in parentheses) are clustered at the respondent level. *P*-values are in square brackets. Figure 1 visualizes partisan gaps by condition. Alternate visualization of results in Supplementary Material section 7, figures 7.1 and 7.2.

Columns (4)–(6) of table 2 show that including self-reported characteristics of respondents does not change the conclusion. Overall, Study 1 suggests that partisan gaps are much larger in surveys with guessing-encouraging features, like the absence of a “Don’t Know” option, the inclusion of social proof and neutral information in the question stem, and no discouragement to guess, all features that are common in commercial polls.

Study 2: The Effect of Partisan Cues on Partisan Gaps

In Study 2, we investigate the impact of partisan cues. We test this by manipulating the partisan cue in the question stem.

Research Design and Data

To answer the question, we leverage data from a national survey conducted by YouGov and a telephone survey in Texas. The YouGov survey includes data from 2,000 respondents interviewed between July 10 and 12, 2012. The Texas survey has data from 1,003 respondents who were interviewed between September 10 and 21, 2012. The data are unweighted.

In the YouGov survey, we asked respondents two retrospective economic evaluation questions: unemployment and the budget deficit. To manipulate congeniality, we randomly inserted a Republican or a Democratic cue into the question stem. In particular, we asked the following two questions:

1. Since the 2010 midterm elections ("when Republicans regained control of the US Congress" or "when Democrats retained control of the Senate"), the unemployment rate [had] gone up, down, or remained the same, or couldn't you say?
2. Since the 2010 midterm elections ("when Republicans regained control of the US Congress" or "when Democrats retained control of the Senate"), has the budget deficit gone up, gone down, remained the same, or couldn't you say?

For the unemployment rate question, we added a no partisan cue condition to establish a neutral baseline against which to measure the effects of partisan framing, with respondents in this condition seeing:

3. Since the 2010 midterm elections, has the unemployment rate gone up, gone down, or remained the same? Or couldn't you say?

The partisan cue was randomized between participants. If someone received the Republican cue in the unemployment question, they also received it in the deficit question.

To test whether our findings would generalize across different policy areas and to examine whether substantive response encouragement would increase or decrease partisan bias, we made two changes to the second question. We switched from budget deficits to federal tax rates and modified the experimental design to three conditions with different question stems:

1. No partisan cue: “Since January 2009, have federal taxes … ”;
2. Democratic cue: “Since Barack Obama took office, have federal taxes … ”; and
3. Democratic cue with substantive response encouragement: “Based on what you have heard, since Barack Obama took office, have federal taxes … ”

All three versions used the same response options: “increased,” “decreased,” “remained the same,” or “couldn’t say.”

For national macroeconomic retrospective evaluation questions like the ones we ask in these two surveys, we assume that the answer is congenial when the correct answer has positive implications for the respondent’s party. On such a question, we expect to obtain the largest estimate of the partisan gap when the partisan cue nudges the president’s copartisans toward the right answer and the president’s main opposing partisans toward the wrong answer. For instance, if, say, under President Obama, the unemployment rate went down over some years, we expect a cue that highlights the Democratic responsibility to lead Democrats to mark more correct answers and Republicans to mark fewer correct answers.

Conversely, a partisan cue that highlights Republican responsibility will lead Democrats toward the wrong answer and Republicans toward the right answer, attenuating the partisan gap. Conditional on there being a partisan cue in the question stem, theoretically, the best estimate of the partisan gap is the difference between proportion correct between groups when the partisan cue directs partisans to the wrong answer. The rationale is that adversarial cues yield estimates of knowledge that are least contaminated with partisan guessing. Partisan guessing is but one force affecting partisan gaps. Another is partisan cheerleading. Having a partisan cue increases partisan cheerleading—deliberately marking the wrong answer even when you know the correct answer. Hence, questions with a neutral stem likely yield a better estimate of partisan gaps than ones with a partisan cue, though we expect question stems with adversarial cues to do the best.

YouGov Results

We estimate the impact of randomly inserting partisan cues on the estimated partisan gap by regressing whether the response is correct on the interaction of partisan cue and whether the correct response is congenial (to party):

$$\text{Correct}_\text{i} = \alpha + \beta\text{Congenial}_\text{i} + \gamma\text{Dem.cue}_\text{i} + \delta\text{Congenial}_\text{i} \times \text{Dem.cue}_\text{i} + \varepsilon_\text{i}, \quad (2)$$

where the constant is the proportion of correct responses in the baseline condition. β captures the partisan gap in the baseline condition (Republican cue). We are interested in δ, which captures how randomly receiving the Democratic cue (which leads Democrats to mark the right answer and Republicans to mark the wrong one) widens the estimated partisan gap.

Table 3 reports the estimated coefficients for the two questions in YouGov. Randomly receiving a Democratic cue increases the probability of getting the correct response in the unemployment question by 16 percentage points ($p < 0.001$) and a seven-percentage-point increase in the gap on the federal deficit question ($p = 0.002$). Adjusting for demographics (see columns (2) and (4)) doesn't appreciably change the coefficients.

And, as we stated in the Research Design and Data section of Study 2, the best estimate of partisans' stores of knowledge is under an adversarial partisan cue. Compared to the theoretical maximal partisan gap (Democrats getting a cue that makes them more likely to get the correct answer and Republicans getting a cue that makes them more likely to get the wrong answer), the partisan gap in the unemployment question obtained under the adversarial cue is, on average, 14 percentage points smaller ($p < 0.001$).

Table 3. The impact of partisan cues on partisan gaps (YouGov).

	"Unemployment has gone down"		"Deficit has gone down"	
	(1)	(2)	(3)	(4)
Congenial	0.044	0.102	0.027	0.029
	(0.029)	(0.030)	(0.014)	(0.016)
	[0.129]	[0.001]	[0.051]	[0.072]
Democratic cue	−0.016	−0.015	−0.009	−0.006
	(0.029)	(0.029)	(0.012)	(0.012)
	[0.588]	[0.612]	[0.449]	[0.615]
Congenial × Democratic cue	0.155	0.147	0.067	0.061
	(0.041)	(0.041)	(0.021)	(0.022)
	[0.000]	[0.000]	[0.002]	[0.005]
Constant	0.297	0.017	0.041	−1.292
	(0.021)	(2.000)	(0.009)	(1.117)
	[0.000]	[0.993]	[0.000]	[0.247]
R^2	0.0273	0.0887	0.0214	0.0383
Demographic controls	No	Yes	No	Yes
Respondents	2,104	2,066	2,104	2,066

Note: Dependent variables indicate whether or not the respondent chose the correct answer. Demographic controls include age cohort, gender, education level, marital status, employment status, news interest, family income, and race. Coefficients are unstandardized; standard errors (in parentheses) are heteroskedasticity-robust. *P*-values are in square brackets. All models are linear probability models.

Texas Lyceum Results

Table 4 shows that the pattern we saw in YouGov (table 3) still holds when we include a neutral cue. Here, we let the Republican cue be the baseline condition and examine the extent to which a neutral cue and a Democratic cue change the estimated partisan gap (see the Research Design and Data section of Study 2). The specification is otherwise similar to Equation (2) (with an additional neutral cue arm). Consistent with the theory, randomly receiving the Democratic cue increases the gap by 20 percentage points ($p = 0.013$) (column (1) in table 4). Again, little changes when we add the baseline demographics (column (2) of table 4). A meta-analysis of the Texas

Table 4. Impact of partisan cue on proportion correct on unemployment (Texas Lyceum).

	"Unemployment has gone down"	
	(1)	(2)
Congenial	0.089	0.048
	(0.054)	(0.073)
	[0.103]	[0.103]
Neutral cue	0.013	0.017
	(0.046)	(0.048)
	[0.771]	[0.725]
Democratic cue	−0.053	−0.047
	(0.044)	(0.045)
	[0.233]	[0.299]
Congenial × Neutral cue	0.072	0.074
	(0.078)	(0.084)
	[0.356]	[0.376]
Congenial × Democratic cue	0.196	0.216
	(0.079)	(0.082)
	[0.013]	[0.009]
Constant	0.189	−0.159
	(0.032)	(0.181)
	[0.000]	[0.380]
R^2	0.0502	0.123
Demographic controls	No	Yes
Respondents	758	752

Note: The dependent variable is whether or not the respondent got the answer correct. Demographic controls include age cohort, gender, education level, marital status, number of children, children's school enrollment, family income, religion, liberalism/conservatism, and race. Coefficients are unstandardized; standard errors (in parentheses) are heteroskedasticity-robust. All models are linear probability models. *P*-values are in square brackets.

Lyceum estimates (table 4) and the YouGov estimate (table 3) suggests a 16-percentage-point increase in the partisan gap when individuals randomly receive a Democratic cue (Supplementary Material figure 7.3).

Finally, we examine the federal tax rate question in the Texas Lyceum survey in table 5. In this question, the baseline condition is without any partisan cue, and we compare how the estimated partisan gap changes with a Democratic cue and a Democratic cue with wording that encourages guessing (see the Research Design and Data section of Study 2). The estimates in column (1) of table 5 have large standard errors that mean we cannot confidently conclude that much about the effect of Democratic cue on the partisan gap. Overall, survey experiments show that partisan cues dramatically affect

Table 5. Impact of partisan cue and guessing encouraging wording on proportion correct on federal taxes (Texas Lyceum).

	"Federal taxes remained the same"	
	(1)	(2)
Congenial	0.098	0.105
	(0.061)	(0.080)
	[0.109]	[0.188]
Democratic cue	−0.050	−0.049
	(0.053)	(0.054)
	[0.344]	[0.364]
Democratic cue w/guess cue	0.070	0.071
	(0.056)	(0.058)
	[0.214]	[0.226]
Congenial × Democratic cue	0.112	0.133
	(0.088)	(0.093)
	[0.202]	[0.154]
Congenial × Democratic cue w/guess cue	−0.058	−0.037
	(0.086)	(0.091)
	[0.499]	[0.687]
Constant	0.303	0.131
	(0.038)	(0.211)
	[0.000]	[0.535]
R^2	0.0214	0.0847
Demographic controls	No	Yes
Respondents	758	752

Note: The dependent variable is whether or not the respondent got the answer correct. Demographic controls include age cohort, gender, education level, marital status, number of children, children's school enrollment, family income, religion, liberalism/conservatism, and race. Coefficients are unstandardized; standard errors (in parentheses) are heteroskedasticity-robust. All models are linear probability models. *P*-values are in square brackets.

the size of partisan gaps. If partisan gaps reflected only partisans' existing stores of knowledge, the gaps would be unresponsive to these cues. Thus, the experiments highlight the role of partisan inference and cheerleading in the estimates of partisan gaps that we see.

Study 3: The Effect of the Scoring Method on Partisan Gaps

Finally, we examine the consequences of scoring decisions on partisan gaps. We introduce a scoring scheme that considers respondents' confidence in their answers. We propose scoring only confidently held correct beliefs about political facts as knowledge.

Research Design and Data

Knowledge questions are commonly offered as multiple-choice items, and conventionally, if a respondent marks the correct answer, it is taken as evidence that the respondent knows the answer. Such scoring does not differentiate between confidently held beliefs, hunches, inferences, blind guesses, and expressive responses. To distinguish between hunches, guesses, and confidently held beliefs, we use the design from studies like Pasek, Sood, and Krosnick (2015). In our confidence coding design (CCD), respondents rate claims on a Likert scale, going from "definitely false" (0) to "definitely true" (10).

To estimate the impact of the question and scoring design that considers respondents' confidence in their answers, we use data from two separate survey experiments. Our first survey experiment is the one underlying Study 1 (*MTurk 1*). The survey had a fifth condition in addition to the four above-mentioned closed-ended multiple-choice conditions. The fifth condition offered the same questions, except respondents were asked to respond on a Likert scale ranging from 0 (definitely not true) to 10 (definitely true). The CCD condition builds on the first four conditions. It discourages guessing and features no social proof or neutral information in the question stem. (See Supplementary Material section 3 for the question text.) Since the items in the multiple-choice questions are dichotomous choice, we offer only a true and an incorrect response; the confidence coding is straightforward. One of the response options from the multiple-choice question becomes its own Likert scale item. Respondents can indicate whether they think the statement (e.g., "Barack Obama was born in the US" or "Barack Obama is a Muslim") is "definitely false (0)" or "definitely true." We scored respondents who marked "definitely true" for the right answer or "definitely false" for the wrong answers as knowledgeable. (See Supplementary Material section 3 for the question text.)

For the second survey experiment, we turn to another MTurk survey fielded on March 27, 2017 (*MTurk 2*). In this survey, we randomly assigned 1,059 respondents to two conditions (multiple choice and CCD). The preamble, topics, and answer options of these questions were identical to the first survey and included questions about the Affordable Care Act (two questions), the effect of greenhouse gases (one question), and the consequences of Trump's executive order on immigration (one question). In the multiple-choice version of the item, participants received three options. In two of the four conditions, respondents also saw a "Don't Know" option. (See Supplementary Material section 5 for the question text.) The data are unweighted.

The scoring for this survey is more nuanced, as the multiple-choice questions had four potential response options. In the confidence coding treatment, survey participants see the same question as in the multiple-choice treatment but have to score the correctness of all the n answer options from the multiple-choice treatment. Broadly, we code an answer as correct if the respondent indicates that they are confident that the correct answer is correct and when they do not indicate that any of the incorrect options might also be correct. More precisely, we code a response as correct if four conditions are met:

1. The respondent is most confident about the correct answer. For instance, it shouldn't be the case that the respondent is more confident about an incorrect answer.
2. The respondent is not as confident about the correct answer as another option. For instance, it cannot be that the four options are all rated 10.
3. The respondent must have at least c level of confidence in the correct answer. We use a c of 10 in the main text, but in Supplementary Material section 6, we try less stringent criteria.
4. The confidence in the incorrect answers cannot be above the threshold t. We use a t of 0 in the main text, but in Supplementary Material section 6, we try less stringent criteria.

MTurk 1 Results

The partisan gap on the best version of the dichotomous multiple-choice items (Condition 4; DK+NSP+GD+NNI) was 0.22 (see figure 1). As figure 2 shows, nearly half of the gap vanishes under confidence scoring. Furthermore, the number of items with no statistically significant gap between partisans doubles from two to four. In all, there is a nearly 11-percentage-point drop in the size of the partisan gap when we treat only confident correct answers as evidence that the respondent knows the answer.

MTurk 2 Results

To further illuminate how treating answers a respondent is confident about as evidence that the respondent knows the fact affects partisan gaps, we regress the dependent variable, an indicator of whether the response is correct, on the interaction between CCD (with conventional scoring serving as the baseline) and the congenial dummy:

$$\mathrm{Correct_{ij}} = \alpha + \beta \mathrm{Congenial_i} + \gamma \mathrm{CCD} + \delta (\mathrm{Congenial_i} \times \mathrm{CCD}) + \varepsilon_{\mathrm{ij}} \quad (3)$$

for respondents i and survey item j. In Equation (1), β captures the difference in the proportion of correct responses when the answer to the question is congenial to the respondent's party affiliation under the baseline conventional scoring condition. A positive estimate indicates that respondents are likelier to choose the correct response when it is congenial to their party affiliation in the multiple-choice treatment. γ captures the effect of relative scoring in the CCD scheme. A positive coefficient indicates that confidence scoring is associated with more correct responses, and a negative one with fewer. δ captures how the two scoring treatments, multiple choice, and CCD, affect the knowledge gaps across partisans for congenial questions. In the pooled equation, which includes all questions, we also include question fixed effects, $\mathrm{question_j}$.

Table 6 reports the results from Equation (3). Columns 1 through 4 report the question-specific estimates. Column 5 pools all questions and adds question fixed effects to the model. In this specification, the intercept term reports the proportion correct for uncongenial questions that were scored with multiple-choice rules. For β, we can see across all but one column (column 4, Donald Trump) that congenial questions in multiple-choice scoring are associated with a higher proportion of correct responses. In the multiple-choice treatment, partisans are more likely to get questions correct when answers are congenial to their partisanship. γ shows that the partisan gap in knowledge nearly disappears in the CCD. Figure 3 visualizes the predicted proportion of correct responses as reported in Table 6.

Pooling evidence across the two studies, it appears that treating only confident correct answers as evidence that the respondent knows the answer yields dramatically lower estimates of the partisan gap.

Validity and Reliability of Different Ways of Measuring Political Knowledge

We have shown that survey and item design choices that encourage guessing have larger partisan gaps than ones that discourage guessing and where the scoring scheme codes only confident correct answers as evidence of knowledge. But which item and survey design choices lead to "better" measures?

Table 6. Confidence scoring and knowledge gaps: MTurk 2.

	Individual survey question				
	Affordable Care Act (1)	Affordable Care Act 2 (2)	Greenhouse gases (3)	Donald Trump (4)	All (5)
Congenial	0.091	0.084	0.087	0.005	0.025
	(0.038)	(0.040)	(0.041)	(0.038)	(0.023)
	[0.018]	[0.036]	[0.033]	[0.895]	[0.270]
Confidence coding design (CCD)	−0.179	−0.201	−0.206	−0.737	−0.377
	(0.028)	(0.030)	(0.032)	(0.028)	(0.018)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Congenial × CCD	−0.071	−0.070	−0.098	0.031	0.024
	(0.039)	(0.041)	(0.041)	(0.046)	(0.026)
	[0.073]	[0.092]	[0.018]	[0.509]	[0.351]
Constant	0.179	0.207	0.217	0.794	0.376
	(0.028)	(0.030)	(0.030)	(0.024)	(0.017)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
R^2	0.119	0.128	0.149	0.528	0.305
Survey item FE	No	No	No	No	Yes
Items	1	1	1	1	4
Respondents	902	902	902	902	902
Respondent-items	902	902	902	902	3,608

Note: Dependent variables indicate whether the respondent answered the question(s) correctly. See Supplementary Material section 5 for the exact wording of the four questions. Columns (1)–(4) estimates by the individual survey questions. Column (5) includes all questions and adds the survey question fixed effects. All models are linear probability models. In the confidence coding scheme, a response is correct only if the correct answer is selected with complete confidence of $c = 10$ (see the Research Design and Data in the Study 3: The Effect of the Scoring Method on Partisan Gaps section). The baseline is the multiple choice designs. Supplementary Material table 6.2 implements a robustness check, setting the relative scoring threshold to $c = 8$. Coefficients are unstandardized. Standard errors (in parentheses) are clustered at the respondent level. *P*-values are in square brackets.

To answer that, we use data from the first MTurk survey to assess the reliability and criterion validity of different designs. Specifically, we use average inter-item correlation and Cronbach's α to measure the scale's reliability. To measure criterion validity, we use the correlation of the scale with three criteria thought to correlate heavily with political knowledge: education, political interest, and political participation (see Supplementary Material section 4 for the question text). We expect items that discourage guessing to have higher reliability and greater criterion validity.

Table 7 reports results for each of the four conditions (see table 1) and the confidence coding condition (CCD) that scores a response as correct when

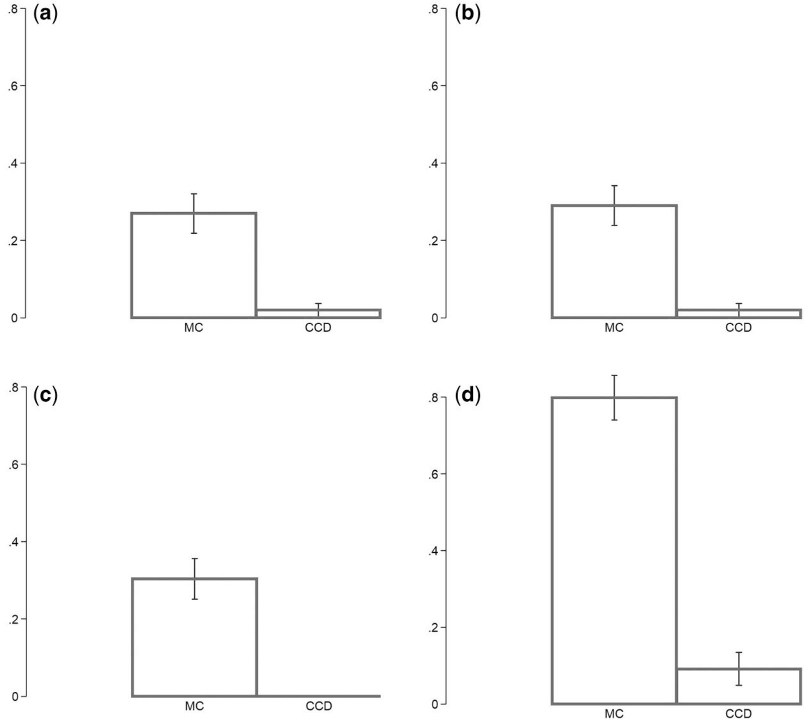

Figure 3. Predicted proportions of correct response by coding (MTurk 2). (a) Affordable Care Act, (b) Affordable Care Act 2, (c) Greenhouse gases, (d) Donald Trump. Bars indicate the predicted proportion of correct responses as reported in table 6. MC bar indicates the predicted proportion of correct responses for multiple choice with congenial responses. The CCD bar indicates the predicted proportion of correct responses for the confidence coding design with congenial responses. Note that the scales of the vertical axes vary. Capped vertical bars indicate 95 percent confidence intervals.

the respondent is completely confident about the correct answer. CCD (Condition 5) has better reliability than other versions. However, the picture is more mixed for the other conditions, with Conditions 1 (NDK+SP+GE+NI) and 3 (DK+NSP+GD+NI) having greater reliability than Conditions 2 (NDK+NSP+GE+NI) and 4 (DK+NSP+GD+NNI). One of the reasons for this mixed picture may be that partisan guessing increases reliability without increasing validity because it introduces correlated errors (a point we discussed in the Guessing vs. Diffidence section). A more diagnostic test for the quality of the instrument, hence, is criterion validity. As Panel A of table 7 shows, the average correlation between Condition 4 (multiple choice with no inflationary

Table 7. Validity and reliability.

	Conditions				
	No DK		With DK		
	Cond. 1 (1)	Cond. 2 (2)	Cond. 3 (3)	Cond. 4 (4)	Cond. 5 (5)
Panel A. Criterion correlational validity					
Political interest	0.115	0.278	0.271	0.412	0.379
Political participation	0.138	0.168	0.276	0.298	0.356
Education	0.077	0.167	0.230	0.180	0.302
Average of 3 items	0.110	0.204	0.259	0.297	0.346
Panel B. Inter-item correlations					
Average inter-item correlation	0.237	0.163	0.248	0.172	0.325
Panel C. Scale reliability					
Cronbach's α	0.737	0.637	0.748	0.652	0.812

Note: Panel A reports the correlation coefficient between each condition and the three criterion variables. Political interest and political participation (voting) are coded on an 11-point scale. Education is coded from 1 to 5 by education qualification. Panel B reports the inter-item correlation for the nine items (see figure 1). Panel C reports the Cronbach's alpha for the nine items. See table 1 for a brief description of the first four conditions and Study 3: The Effect of the Scoring Method on Partisan Gaps for the confidence coding design.

features) and the confidence coding design and criterion variables is markedly higher (0.35) than in Conditions 1–3. The baseline Condition 1 (0.011), Condition 2 (0.20), and Condition 3 (0.26) all score lower.[4]

The results above are consistent with those obtained by Graham (2023), which finds that the test-retest reliability of confident correct answers is much higher, and Vidigal (2025), which finds that "incorporating belief certainty results in a knowledge scale that displays theoretically expected relationships with a range of outcome variables while also having superior psychometric properties."

Discussion and Conclusion

Since at least the publication of Bartels (2002), the conventional wisdom has been that partisan gaps in confidently held beliefs about politically

4. We did one more test to get at the validity. We hypothesized that partisan guessing would lead to a greater negative correlation between congenial and uncongenial items on items that encouraged guessing. And indeed, the item-rest correlations between uncongenial and congenial items are the smallest for CCD.

consequential facts are wide and widespread. The conventional wisdom in academia has also become the received wisdom for the mass public—nearly 80 percent of Americans believe that Democrats and Republicans disagree on facts (LaLoggia 2018).

In line with some other research on this topic (Bullock et al. 2015; Prior, Sood, and Khanna 2015; Schaffner and Luks 2018, though see Berinsky 2018 and Peterson and Iyengar 2021), our results suggest that a big chunk of the partisan gap in the knowledge of politically consequential facts in the United States is not founded in differences in confidently held beliefs. We find that standard features of commercial polls, like not offering "don't know," inserting a partisan cue, and treating unconfident answers as knowledge, inflate the partisan gaps. Removing such "inflationary" features reduces the partisan gap dramatically.

The fact that partisan gaps are smaller may seem at odds with some political behavior research. For instance, selective exposure theory posits vast imbalances in the consumption of partisan news. However, recent studies show that most people consume scant political news (Prior 2007; Flaxman, Goel, and Rao 2016), and the news that they do consume is relatively balanced (Gentzkow and Shapiro 2011; Flaxman, Goel, and Rao 2016; Garz et al. 2020; Guess 2021). Other evidence points to the fact that Democrats and Republicans update similarly in light of events (Gerber and Green 1999; Kernell and Kernell 2021; Coppock 2023). In the end, the results of our studies paint a mixed picture of democratic competence.

Smaller partisan gaps partly result from the fact that the average respondent doesn't know the facts. It is primarily partisan guessing masquerading as partisan gaps. The upside is that partisan gaps are small; the downside is that people know even less than we thought.

Finally, while the data are from the United States, we think there is reason to believe that similar concerns vitiate partisan gaps in knowledge measured in other countries. Research focusing on multiparty systems around the world has found that they are increasingly presidentializing[5] and affectively polarizing (Hobolt, Leeper, and Tilley 2021; Wagner 2021, 2024). See also Bailey (2021), which shows that self-reported economic perceptions by British voters are shaped by political cues in surveys, and Bisgaard and Slothuus (2018), which shows how central partisan cues have become for economic perceptions in Denmark. Hence, we expect the conclusions from the paper to travel to other contexts.

5. See Poguntke and Webb (2005) for the general argument, Krauss and Nyblade (2005) for Japan, and Poguntke and Webb (2015) for Italy and Germany.

Supplementary Material

Supplementary Material may be found in the online version of this article: https://doi.org/10.1093/poq/nfaf044.

Data Availability

Replication data and documentation are available at https://doi.org/10.7910/DVN/KNN0GL.

References

Bailey, Jack. 2021. "Political Surveys Bias Self-Reported Economic Perceptions." *Public Opinion Quarterly* 85:987–1008.

Bartels, Larry M. 2002. "Beyond the Running Tally: Partisan Bias in Political Perceptions." *Political Behavior* 24:117–50.

Berinsky, Adam J. 2018. "Telling the Truth About Believing the Lies? Evidence for the Limited Prevalence of Expressive Survey Responding." *The Journal of Politics* 80:211–24.

Bisgaard, Martin, and Rune Slothuus. 2018. "Partisan Elites as Culprits? How Party Cues Shape Partisan Perceptual Gaps." *American Journal of Political Science* 62:456–69.

Bullock, John G., Alan S. Gerber, Seth J. Hill, and Gregory A. Huber. 2015. "Partisan Bias in Factual Beliefs About Politics." *Quarterly Journal of Political Science* 10:519–78.

Bullock, John G., and Kelly Rader. 2022. "Response Options and the Measurement of Political Knowledge." *British Journal of Political Science* 52:1418–27.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1980. *The American Voter*. Chicago, Illinois: University of Chicago Press.

Cialdini, Robert B. 2009. *Influence: Science and Practice*. Boston, Massachusetts: Pearson Education.

Coppock, Alexander. 2023. *Persuasion in Parallel: How Information Changes Minds about Politics*. Chicago, Illinois: University of Chicago Press.

Coppock, Alexander, Thomas J. Leeper, and Kevin J. Mullinix. 2018. "Generalizability of Heterogeneous Treatment Effect Estimates Across Samples." *Proceedings of the National Academy of Sciences of the United States of America* 115:12441–46.

Cor, M. Ken, and Gaurav Sood. 2016. "Guessing and Forgetting: A Latent Class Model for Measuring Learning." *Political Analysis* 24:226–42.

Dolan, Kathleen, and Michael A. Hansen. 2020. "The Variable Nature of the Gender Gap in Political Knowledge." *Journal of Women, Politics & Policy* 41:127–43.

Ferrín, Monica, Marta Fraile, and Gema García-Albacete. 2017. "The Gender Gap in Political Knowledge: Is it all about Guessing? An Experimental Approach." *International Journal of Public Opinion Research* 29:111–32.

Flaxman, Seth, Sharad Goel, and Justin M. Rao. 2016. "Filter Bubbles, Echo Chambers, and Online News Consumption." *Public Opinion Quarterly* 80:298–320.

Fortin-Rittberger, Jessica. 2016. "Cross-National Gender Gaps in Political Knowledge: How Much is Due to Context?"*Political Research Quarterly* 69:391–402.

Garz, Marcel, Gaurav Sood, Daniel F. Stone, and Justin Wallace. 2020. "The Supply of Media Slant Across Outlets and Demand for Slant Within Outlets: Evidence from US Presidential Campaign News." *European Journal of Political Economy* 63:101877.

Gentzkow, Matthew, and Jesse M. Shapiro. 2011. "Ideological Segregation Online and Offline." *The Quarterly Journal of Economics* 126:1799–839.

Gerber, Alan, and Donald Green. 1999. "Misperceptions About Perceptual Bias." *Annual Review of Political Science* 2:189–210.
Graham, Matthew H. 2023. "Measuring Misperceptions?"*American Political Science Review* 117:80–102.
Graham, Matthew H, and Omer Yair. 2025. "Less Partisan but no More Competent: Expressive Responding and Fact-Opinion Discernment." *Public Opinion Quarterly* 1:7–30.
Guess, Andrew M. 2021. "(Almost) Everything in Moderation: New Evidence on Americans' Online Media Diets." *American Journal of Political Science* 65:1007–22.
Hobolt, Sara B, Thomas J Leeper, and James Tilley. 2021. "Divided by the Vote: Affective Polarization in the Wake of the Brexit Referendum." *British Journal of Political Science* 51:1476–93.
Hochschild, Jennifer, and Katherine Levine Einstein. 2015. "It isn't what we don't know that gives us Trouble, it's What we know that ain't so': Misinformation and Democratic Politics." *British Journal of Political Science* 45:467–75.
Yair, Omer, and Gregory A. Huber. 2020. "How Robust is Evidence of Partisan Perceptual Bias in Survey Responses? A New Approach for Studying Expressive Responding." *Public Opinion Quarterly* 84:469–92.
Jerit, Jennifer, and Jason Barabas. 2012. "Partisan Perceptual Bias and the Information Environment." *The Journal of Politics* 74:672–84.
Jessee, Stephen A. 2017. "Don't Know" Responses, Personality, and the Measurement of Political Knowledge." *Political Science Research and Methods* 5:711–31.
Kernell, Georgia, and Samuel Kernell. 2021. "Monitoring the Economy." *Journal of Elections, Public Opinion and Parties* 31:199–219.
Kraft, Patrick W., and Kathleen Dolan. 2023. "Asking the Right Questions: A Framework for Developing Gender-Balanced Political Knowledge Batteries." *Political Research Quarterly* 76:393–406.
Krauss, Ellis S., and Benjamin Nyblade. 2005. "Presidentialization'in Japan? The Prime Minister, Media and Elections in Japan." *British Journal of Political Science* 35:357–68.
LaLoggia, J. 2018. Republicans and Democrats agree: They can't agree on basic facts, Pew Research Center. United States of America. Retrieved from https://policycommons.net/artifacts/617199/republicans-and-democrats-agree/1597955/ on 14 Aug 2024. CID: 20.500.12592/txbsxj.
Lodge, Milton, and Charles S. Taber. 2013. *The Rationalizing Voter*. New York: Cambridge University Press.
Luskin, Robert C., and John G. Bullock. 2011. "Don't Know" Means "Don't Know": DK Responses and the Public's Level of Political Knowledge." *The Journal of Politics* 73:547–57.
Luskin, Robert C., Gaurav Sood, Yul Min Park, and Joshua Blank. 2018. "Misinformation about Misinformation? Of Headlines and Survey Design." Working paper. https://gsood.com/research/papers/misinformation_misinformation.pdf
Malka, Ariel, and Mark Adelman. 2023. "Expressive Survey Responding: A Closer Look at the Evidence and Its Implications for American Democracy." *Perspectives on Politics* 21:1198–209.
Mondak, Jeffery J. 1999. "Reconsidering the Measurement of Political Knowledge." *Political Analysis* 8:57–82.
Mondak, Jeffery J., and Mary R. Anderson. 2004. "The Knowledge Gap: A Reexamination of Gender-Based Differences in Political Knowledge." *The Journal of Politics* 66:492–512.
Mullinix, Kevin J., Thomas J. Leeper, James N. Druckman, and Jeremy Freese. 2015. "The Generalizability of Survey Experiments." *Journal of Experimental Political Science* 2:109–38.

Pasek, Josh, Gaurav Sood, and Jon A. Krosnick. 2015. "Misinformed about the Affordable Care Act? Leveraging Certainty to Assess the Prevalence of Misperceptions." *Journal of Communication* 65:660–73.

Peterson, Erik, and Shanto Iyengar. 2021. "Partisan Gaps in Political Information and Information-Seeking Behavior: Motivated Reasoning or Cheerleading?" *American Journal of Political Science* 65:133–47.

Poguntke, Thomas, and Paul Webb. 2005. "The Presidentialization of Politics in Democratic Societies: A Framework for Analysis." In *The Presidentialization of Politics: A Comparative Study of Modern Democracies*, edited by Thomas Poguntke and Paul Webb, 1–25. Oxford, UK: Comparative Politics, Oxford University Press.

———. 2015. "Presidentialization and the Politics of Coalition: Lessons from Germany and Britain." *Italian Political Science Review/Rivista Italiana di Scienza Politica* 45:249–75.

Prior, Markus. 2007. *Post-Broadcast Democracy: How Media Choice Increases Inequality in Political Involvement and Polarizes Elections*. New York: Cambridge University Press.

Prior, Markus, Gaurav Sood, and Kabir Khanna. 2015. "You Cannot Be Serious: The Impact of Accuracy Incentives on Partisan Bias in Reports of Economic Perceptions." *Quarterly Journal of Political Science* 10:489–518.

Roush, Carolyn E., and Gaurav Sood. 2023. "A Gap in Our Understanding? Reconsidering the Evidence for Partisan Knowledge Gaps." *Quarterly Journal of Political Science* 18:131–51.

Sanchez, Maria Elena, and Giovanna Morchio. 1992. "Probing "dont know" answers: Effects on survey estimates and variable relationships." *Public Opinion Quarterly* 56:454–74.

Schaffner, Brian F., and Samantha Luks. 2018. "Misinformation or Expressive Responding? What an Inauguration Crowd Can Tell Us About the Source of Political Misinformation in Surveys." *Public Opinion Quarterly* 82:135–47.

Sherif, Muzafer. 1935. "A study of some Social Factors in Perception." Archives of Psychology (Columbia University). https://psycnet.apa.org/record/1936-01332-001

Sturgis, Patrick, Nick Allum, and Patten Smith. 2008. "An Experiment on the Measurement of Political Knowledge in Surveys." *Public Opinion Quarterly* 72:90–102.

Vidigal, Robert. 2025. "Measuring Belief Certainty in Political Knowledge." Political Behavior 47:529–51.

Wagner, Markus. 2021. "Affective Polarization in Multiparty Systems." *Electoral Studies* 69:102199.

——— 2024. "Affective Polarization in Europe." *European Political Science Review* 16:378–92.

Public Opinion Quarterly, 2025, 00, 1–25
https://doi.org/10.1093/poq/nfaf044
Original article