An Unclear Gap: How Vague Response Options Produce Partisan Knowledge Gaps [*]

Carolyn E. Roush[†]   Gaurav Sood[‡]

June 27, 2023

Abstract

Roush and sood (2023) use a dataset of 162,083 responses to 187 items on 47 surveys to find that partisan gaps are smaller and less frequent than commonly understood. The average is a mere six and a half points and gaps' "signs" run counter to expectations roughly 30% of the time. However, one exception is the size of gaps on retrospection items on the ANES, which are considerably bigger. These retrospection items use vague response options, e.g., 'About the same.' Vague response options can inflate partisan gaps by offering partisans the opportunity to interpret the same data differently. We test this assumption with a novel survey experiment. We present partisans data indicating a small improvement in economic indicators and manipulate the partisan tint of the change by manipulating who is responsible for the change. We find that significantly fewer partisans pick the option that '[things] got better' when presented with an out-partisan cue than a co-partisan cue. Our findings suggest that vague options can induce knowledge gaps where none exist.

Data from Roush and sood (2023) suggest that partisan knowledge gaps are highly variable and that large differences in what Democrats and Republicans believe are less common than what many assume. If partisan gaps are small on average, why does the common wisdom that Democrats and Republicans differ substantially in political knowledge persist?

One explanation for this discrepancy is that conventional wisdom is largely based on studies using data from the American National Election Studies (N.d.). But there may be a good reason to think that the gaps in the ANES data are not representative of broader trends in partisan knowledge differences. Unlike most knowledge questions—which require partisans to identify an objectively correct answer—most ANES questions about party consequential "factual beliefs" do not ask respondents to do the same. Instead, these questions ask respondents to make subjective *assessments* about performance or policy over a certain time period. Canonical ANES questions, for example, ask people to gauge whether the budget deficit increased, decreased, or remained about the same over a president's tenure, or how the rate of inflation changed over the past year. Because the response options for these questions—"got better," "stayed about the same," or "got worse"—are imprecise, people have a greater opportunity to interpret the meaning themselves (e.g. Beyth-Marom 1982) using common heuristics, including partisanship (e.g. Sood and Guess 2017). As a result, a large partisan "knowledge" gap may reflect how partisans interpret response options rather than a true difference between what Democrats and Republicans know.

This is particularly problematic for cases where changes in inflation, unemployment, the deficit, or other performance items are marginal. While there are certain contexts—such as a stock market crash—where unambiguous evidence forces partisans to acknowledge the same economic reality (e.g. Bisgaard 2015; Parker-Stephen 2013), far more survey questions are asked in times when performance indicators change gradually over time. When researchers ask respondents to classify these changes in performance indicators using vague response options, it opens the door for partisan bias even if individuals know the same objective informa-

tion. Consider the case of two highly knowledgeable survey respondents (who perhaps work in the Bureau of Labor Statistics) who know definitively that the national unemployment rate in the United States grew from 4.0% to 4.2% over the past year, a time during which a Republican president occupied the White House. When the first respondent, a Democrat, is asked to evaluate how unemployment changed over the past year, she might (correctly) reason that unemployment "got worse" as the rate objectively increased over the previous 12 months. On the other hand, the second survey respondent, a Republican, might also (reasonably) conclude that 0.2 percentage points is a negligible change in unemployment, and might therefore be more liable to answer that the unemployment rate "stayed about the same" over the past year. In this situation, two people who know *the exact same fact* could plausibly choose two different response options and still be correct. The end result is that some "knowledge gaps" may be artificially large simply because respondents interpret the same response categories differently.

To test this hypothesis, we first examined the average size of partisan knowledge gaps that occur in ANES data and compared them to the average partisan gaps in other studies. To do so, we relied on the data compiled by Roush and sood (2023): all knowledge items that carry a partisan implication that appeared on ANES surveys over the past 32 years.

Table 1 compares partisan gaps on ANES items to those included in the other studies in our analysis. As expected, the mean and median gaps on ANES knowledge items are substantially larger than those in the other three studies. In fact, the mean partisan knowledge gap in the ANES data (17 percentage points) is more than 50% larger than the largest average gap in any other study (10 percentage points, from Bullock et al. (2015)). Furthermore, only two of the 28 items taken from ANES surveys produce negatively-signed partisan gaps, and only five of these items produce knowledge gaps that are not statistically significant at the 95% confidence level. While the sample size of questions taken from the ANES is small, it is clear that the partisan gaps produced from these items are markedly larger.

Our hunch is that these large gaps are a result of vague response categories that allow

Study	Mean Gap	Median Gap	SD	N(Items)
ANES	0.168	0.139	0.139	28
Bullock et al.	0.104	0.089	0.119	21
Jerit & Barabas	0.036	0.032	0.091	128
Prior et al.	0.063	0.035	0.093	10

Table 1: Partisan Gap by Study

partisans to classify the same information in different ways. To test this hypothesis, we conducted an original experiment using Amazon's Mechanical Turk (MTurk) in June 2020. In the study, we provided all respondents with a question prompt that featured real economic information about the change in the inflation and unemployment rates during 2016.[1] In addition to providing this information, we randomly assigned respondents to one of two treatments: one that attributed these changes to then-Democratic President Barack Obama and the other that attributed the changes to the Republican-controlled Congress. We then asked respondents to classify these changes using the canonical ANES response categories ("got worse," "stayed about the same," or "got better"). The specific treatment was as follows:

During 2016, (when Barack Obama was president | when Republicans were in control of both Houses of Congress), unemployment decreased from 5.0% to 4.8%, a change of 0.2 percentage points. How would you interpret this change? Would you say that unemployment got better, stayed about the same, or got worse?

In 2016, inflation also decreased from 2.1% to 1.9%, a change of 0.2 percentage points. How would you interpret this change? Would you say that inflation got better, stayed about the same, or got worse?

Since prior research demonstrates that partisans evaluate economic conditions favorably when their own party is in power and unfavorably when the other party is in power (e.g.

---

[1]This information was collected from the U.S. Bureau of Labor, available at `https://data.bls.gov/timeseries/lns14000000`.

Bartels 2002; Bisgaard 2015), we expected respondents to classify objective economic information differently depending on the partisan cue they received. Specifically, we expected that partisans would be more likely to classify a 0.2 reduction in the unemployment or inflation rates as having "got[ten] better" under co-party leadership and as having "stayed about the same" (or "got[ten] worse") under the opposing party's leadership. For ease of interpretation, we recoded the data so that treatments and respondents are characterized in relation to one another: we classified Democrats who saw the President Obama cue and Republicans who saw the Republicans-in-Congress cue as receiving an *In-party cue* and Democrats who saw the Republicans-in-Congress cue and Republicans who saw the President Obama cue as having received an *Out-party cue*).[2]

Figure 1 presents the distribution of responses by experimental conditions for both dependent variables. As we can see, partisans classify a small, 0.2 percentage point change in unemployment or inflation very differently depending on the party to which the change was attributed. Respondents who received the *Out-party cue* were, on average, 21.5 percentage points less likely than those receiving the *In-party cue* to view the reduction in the unemployment rate as having "got[ten] better" during 2016. Respondents who received the *Out-party cue* were about 18 percentage points less likely than *In-party cue* respondents to classify the reduction in unemployment as "stayed about the same." We find similar results for inflation, albeit with somewhat smaller effects: those who received the *Out-party cue* were approximately 16 percentage points less likely than those who received the *In-party cue* to classify the 0.2 reduction in inflation as having "got[ten] better;" those who received the *Out-party cue* were also 13.6 percentage points more likely than *In-party cue*-receivers to classify the change as "stayed about the same."

---

[2]Consistent with previous research (e.g., Keith et al. 1992), we classify Independent leaners as partisans.

Figure 1: *Distribution of Dependent Variables by Experimental Condition*
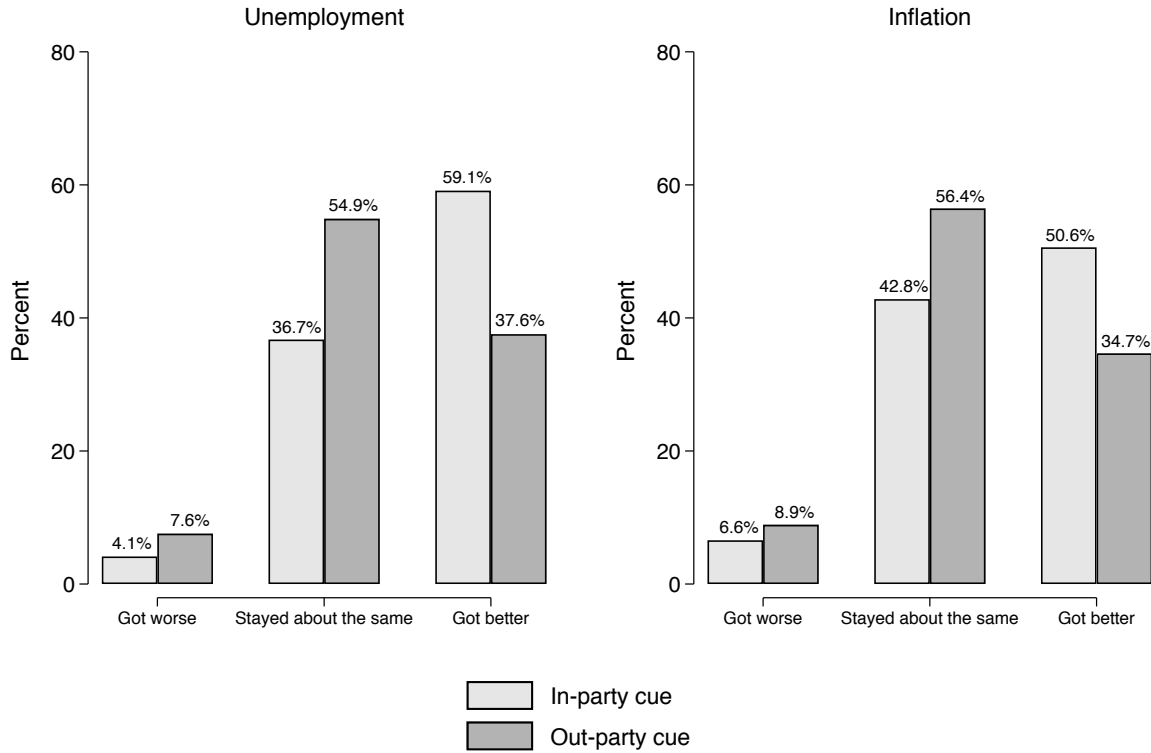
We also estimated two regression models predicting evaluations of unemployment and inflation (where 1 = "got better," 0.5 = "stayed about the same," and 0 = "got worse") as a function of whether or not respondents received the *Out-party cue*. The results of this analysis can be found in Table 2.[3] Consistent with the results in Figure 1, partisans who received the

---

[3]In recent years, researchers have noted that significant portions of the data collected on MTurk is of questionable quality, provided either by respondents who provide misleading information regarding the location from which they are completing a HIT or who provide humorous or insincere responses to survey questions (e.g., Ahler, Roush and Sood 2020; Amazon Mechanical Turk 2019; Bai 2018; Dreyfuss 2018; Kennedy et al. 2020; Ryan 2018). As demonstrated by Ahler, Roush and Sood (2020) and Kennedy et al. (2020), these bad actors can attenuate treatment effects by introducing noise into the data. Accordingly, we followed the recommendations of Ahler, Roush and Sood (2020) and Kennedy et al. (2020) to iden-

Out-party cue were less likely than respondents who received the *In-party cue* to view (normatively positive) reductions in unemployment and inflation favorably. Specifically, those respondents who were told that an out-party leader oversaw the reduction in the unemployment rate viewed the change 12.5 percentage points less favorably than those who received the *In-party cue*. Similarly, respondents who received the *Out-party cue* viewed the reduction in the inflation rate 9.1 percentage points less favorably than those who received the *In-party cue*. The magnitude of these effects is sizeable.

Table 2: *Impact of Treatment on Economic Evaluations*

	Unemployment	Inflation
Out-party cue	-0.125***	-0.091***
	(0.020)	(0.021)
Constant	0.775***	0.720***
	(0.015)	(0.015)
Observations	861	861
R-squared	0.043	0.022

Standard errors in parentheses.
***$p<0.01$, **$p<0.05$, *$p<0.1$, two-tailed.
All variables have been rescaled 0-1 for ease of interpretation.

We replicated the experiment on Lucid in 2023. Table 3 presents the results. Once

tify these suspicious responses and only present results among non-suspicious respondents in both Figure 1 and Table 2 (*n*=861). Consistent with previous research, we find that suspicious respondents attenuated treatment effects. Nevertheless, we still find impressive effects when including suspicious respondents in our analysis: among the full sample (*n*=1,425), receiving the *Out-party cue* causes respondents to view the change in unemployment 9.6 percentage points more negatively and the change in inflation 6.0 percentage points more negatively than those respondents who received the *In-party cue*. For more information, please see SI 2.

again, we see that respondents are less likely to pick the option that things are getting better when we present an uncongenial cue versus a scenario where we present a congenial cue.

Table 3: Impact of Treatment on Economic Evaluations (Lucid)

	Dependent variable:	
	Unemployment	Inflation
	(1)	(2)
Out-party cue	−0.072**	−0.102***
	(0.030)	(0.030)
Constant	0.741***	0.771***
	(0.021)	(0.021)
Observations	529	529
R^2	0.011	0.021

Note:	*p<0.1; **p<0.05; ***p<0.01
	Standard errors in parentheses.
	All variables have been rescaled 0-1 for ease of interpretation.

Our results suggest that a significant portion of partisans' disagreement about the "acceptance of basic political facts, such as the state of the economy" (Berinsky 2017, 211) might be better explained as a biased interpretation of response categories rather than genuine differences in knowledge. Even when partisans are presented with the *exact same* factual information, they classify it differently based upon their preexisting biases and political context (see also Gaines et al. 2007). The fact that vague response options are common on the ANES—perhaps the most commonly used source of public opinion data in the discipline—helps contribute to the (mistaken) belief that differences in what Democrats and Republicans are large enough to warrant serious concerns about democratic accountability.

References

Ahler, Douglas J., Carolyn E. Roush and Gaurav Sood. 2020. "The Micro-Task Market for Lemons: Data Quality on Amazon's Mechanical Turk." *Political Science Research & Methods* Revise and resubmit.

Amazon Mechanical Turk. 2019. "MTurk Worker Quality and Identity." Available at https://blog.mturk.com/mturk-worker-identity-and-task-quality-d3be46d83d0d.

American National Election Studies. N.d. "American National Election Studies." Available at http://electionstudies.org/.

Bai, Hui. 2018. "Evidence that a Large Amount of Low Quality Responses on MTurk Can be Detected with Repeated GPS Coordinates." Available at https://www.maxhuibai.com/blog/evidence-that-responses-from-repeating-gps-are-random.

Bartels, Larry M. 2002. "Beyond the Running Tally: Partisan Bias in Political Perceptions." *Political Behavior* 24(2):1061–1078.

Berinsky, Adam J. 2017. "Telling the Truth About Believing the Lies? Evidence for the Limited Prevalence of Expressive Survey Responding." *Journal of Politics* 80(1):211–224.

Beyth-Marom, Ruth. 1982. "How Probable is Probable? A Numerical Translation of Verbal Probability Expressions." *Journal of Forecasting* 1(3):257–269.

Bisgaard, Martin. 2015. "Bias Will Find a Way: Economic Perceptions, Attributions of Blame, and Partisan-Motivated Reasoning During Crisis." *The Journal of Politics* 77(3):849–860.

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* 10:519–578.

Dreyfuss, Emily. 2018. "A Bot Panic Hits Amazon's Mechanical Turk." *Wired* 17 August. Available at https://www.wired.com/story/amazon-mechanical-turk-bot-panic/.

Gaines, Brian J., James H. Kuklinski, Paul J. Quirk, Buddy Peyton and Jay Verkuilen. 2007. "Same Facts, Different Interpretations: Partisan Motivation and Opinion on Iraq." *Journal of Politics* 69(4):957–974.

Keith, Bruce E., David B. Magleby, Candice J. Nelson, Elizabeth Orr, Mark C. Westlye and Raymond E. Wolfinger. 1992. *The Myth of the Independent Voter*. Berkeley: University of California Press.

Kennedy, Ryan, Scott Clifford, Tyler Burleigh, Philip D. Waggoner, Ryan Jewell and Nicholas Winter. 2020. "The Shape of and Solutions to the MTurk Quality Crisis." *Political Science Research & Methods* forthcoming.

Lopez, Jesse and D. Sunshine Hillygus. 2018. Why So Serious? Survey Trolls and Misinformation. In *Annual Meeting of the Midwest Political Science Association*. Chicago: .

Parker-Stephen, Evan. 2013. "Tides of Disagreement: How Reality Facilitates (and Inhibits) Partisan Public Opinion." *Journal of Politics* 75(4):1077–1088.

Roush, Carrie and Gaurav sood. 2023. "A Gap in Our Understanding? Reconsidering the Evidence for Partisan Knowledge Gaps." *Quarterly Journal of Political Science* .

Ryan, Timothy J. 2018. "Data Contamination on MTurk." Available at http://timryan.web.unc.edu/2018/08/12/data-contamination-on-mturk/.

Sood, Gaurav and Andrew Guess. 2017. "By the Numbers: Toward More Precise Numerical Summaries." Working paper.

SI 1   Question-Wording from June 2020 MTurk Experiment

Switching gears, we'd like to understand how you think various measures of the economy performed a few years ago, *(when Barack Obama was president | when Republicans were in control of both Houses of Congress)*.

During 2016, *(when Barack Obama was president | when Republicans were in control of both Houses of Congress)*, unemployment decreased from 5.0% to 4.8%, a change of 0.2 percentage points. How would you interpret this change? Would you say that unemployment got better, stayed about the same, or got worse?

- Got better

- Stayed about the same

- Got worse

In 2016, inflation also decreased from 2.1% to 1.9%, a change of 0.2 percentage points. How would you interpret this change? Would you say that inflation got better, stayed about the same, or got worse?

- Got better

- Stayed about the same

- Got worse

SI 2 MTurk Data Quality and Attenuation of Treatment Effects

As mentioned previously, we followed the advice of Ahler, Roush and Sood (2020) and Kennedy et al. (2020) to identify low-quality responses on Amazon's Mechanical Turk (MTurk). To do so, we first used a Qualtrics plugin to record the IP addresses from which respondents were taking the survey. We further collected IP-level metadata and flagged any respondent who took the survey from outside the United States (as the survey was limited to American adults), completed the survey more than once, or completed the survey from a blacklisted address as suspicious/of potential low-quality. In order to identify respondents who may provide humorous or insincere responses to survey questions, we also asked respondents a series of low-incidence screener questions. These questions ask about rare afflictions and behaviors, such as whether the respondent was a member of a gang, whether the respondent used a prosthetic, etc. Following Ahler, Roush and Sood (2020) and Lopez and Hillygus (2018), we classified any respondent as suspicious/a potential provider of low-quality data if they answered in the affirmative to two or more of these questions. In all, we found that 38% of our data is of questionable quality.

To determine whether low-quality responses attenuate treatment effects, we estimated four regression models in SI 2.1. For context, the first two models provide the results among the entire sample (including low-quality responses) using the unemployment and inflation dependent variables, respectively. The last two models include an indicator for whether the respondent was flagged as providing a *Low-quality response* and an interaction term comprised of the *Low-quality response* indicator and assignment to the *Out-party cue* condition. As we can see, flagged respondents attenuate treatment effects for both dependent variables, but impressive effects remain: even when including low-quality respondents in our data, respondents are 9.6 percentage points less likely to view the change in unemployment as having "gotten better" and 6.0 percentage points less likely to view the change in inflation similarly under out-party

leadership.

Table SI 2.1: *Impact of Low-Quality Responses on Treatment Effects*

	(1) Unemployment	(2) Inflation	(3) Unemployment	(4) Inflation
Out-party cue	-0.096***	-0.060***	-0.125***	-0.091***
	(0.015)	(0.016)	(0.020)	(0.021)
Low-quality response			0.045**	-0.022
			(0.022)	(0.024)
Out-party cue * low-quality response			0.076**	0.079**
			(0.031)	(0.033)
Constant	0.793***	0.711***	0.775***	0.720***
	(0.011)	(0.012)	(0.014)	(0.015)
Observations	1,425	1,425	1,425	1,425
R-squared	0.027	0.010	0.050	0.014