Merely Asking
A Replication of Ahler and Sood (2018)

January 16, 2026

Abstract

Since the publication of Ahler and Sood (2018), a growing literature finds that correcting partisan citizens' misperceptions about the out-party reduces affective polarization. However, Ahler and Sood's original experiment also shows that "merely asking" respondents to report their perceptions (without any correction of misperceptions) leads to a statistically significant reduction in affective polarization. This result has been largely ignored in subsequent research, but has important implications for both theory and practice. In this preregistered report, we propose to replicate Ahler and Sood's important result using a well-powered, pre-registered survey experiment.

Introduction

An innovative and growing body of work has emphasized the importance of affective polarization for understanding politics in the United States and abroad. In the U.S., this work typically defines partisan affective polarization as the difference between one's feelings toward their own party relative to their feelings toward their outparty (see e.g., Druckman and Levy 2022). As Ahler and Sood (2018) famously show, affective polarization can be reduced by correcting citizens' misperceptions about the demographic make-up of the out-party. Subsequent work builds on their finding. Lees and Cikara (2020) and Marsh (2023) each correct misperceptions about partisan beliefs. Ruggeri et al. (2021) examine the effect of corrections in different countries. Broockman, Kalla, and Westwood (2023) use corrections to experimentally manipulate affective polarization. And scholars and practitioners have taken these results seriously. Five of the 25 treatments in Voelkel et al.'s (2024) "megastudy" of interventions to reduce affective polarization focus "a great deal" on correcting misperceptions of the out-party.

While the work following Ahler and Sood focuses almost entirely on *correcting* misperceptions, Ahler and Sood are careful to include a condition in their experiment that *merely asks* respondents to report their perceptions of the out-party, without a correction. Although not widely appreciated in subsequent research, *Ahler and Sood find that merely asking respondents about the demographic make-up of the out-party—without any correction of misperceptions—also reduces affective polarization.* Indeed, they do not find a statistically significant difference between (A) the effect of asking perceptions and (B) the effect of asking perceptions and then correcting misperceptions. We argue that this result merits further attention because, if replicable, it offers an alternative intervention for reducing affective polarization and it opens up new opportunities for theoretical advancement for understanding the antecedents of affective polarization.

Despite the centrality of Ahler and Sood (2018) to the literature on affective polarization, its primary experiment has yet to receive a complete, well-powered replication.[1] Instead, it has received only partial replications (e.g., Broockman, Kalla, and Westwood 2023) or extensions (e.g., Lees and Cikara 2020; Ruggeri et al. 2021). In each case, subsequent work has focused on the effects of corrections. The results of these extensions have been mixed, leading Dias et al. (2024, p. 2) to emphasize the need to "explore other interventions that could be plausibly implemented at scale." By ignoring the effect of "merely asking," academics and practitioners alike may be missing a simple and effective method to reduce affective polarization. Yet it is hard to blame anyone for focusing on correction effects, given the large amount of subsequent attention to correction as an intervention (e.g., Voelkel et al. 2024). In contrast, as we explain below, it would risky to rely on an intervention of merely asking without replication of the original experiment. We therefore propose to replicate the experimental design from Ahler and Sood (2018) using (1) a larger sample size, (2) a more powerful experimental design, and (3) an improved measure of affective polarization. As an added benefit, the proposed replication allows us to validate the increasingly popular pre-post design (Clifford, Sheagley, and Piston 2021; Jordan, Ollerenshaw, and Trexler 2025) for the study of affective polarization.

Isager et al. (2023) argue that we ought to prioritize replicating claims with (1) the highest value and (2) the highest uncertainty. Concretely, Isager, van 't Veer, and Lakens (2021) argue that we can roughly approximate "value" with citations and "uncertainty" with sample size. Ahler and Sood (2018) is an extremely well-cited paper and their tests use relatively small sample sizes.[2] Using this framework focusing on value and uncertainty, we

[1] Ahler and Sood (2018) also provide a variety of observational analyses. We focus specifically on the experimental design and results reported in Figure 4A on p. 997.

[2] Importantly, though, Rainey (2025*b*) notes that citations and sample size are only rough heuristics, which can be misleading in the absence of more detailed arguments about value and uncertainty. Further, Feldman (2025) notes several other tangible benefits of close replication not contained in the concept of "replication value" (Isager et al. 2023), all of which apply to our study.

argue below that Ahler and Sood's results, especially with respect to *merely asking*, should be a high-priority for replication. First, Ahler and Sood (2018) has received exceptional amount of scholarly and popular attention. Second, a general interest is emerging in exploring more interventions that influence affective polarization. Third, merely asking has practical benefits for survey experiments and a compelling theoretical rationale. Fourth, the evidence for the effect of merely asking on affective polarization remains weak.

The Legacy of Ahler and Sood (2018)

Ahler and Sood (2018) has received a great deal of attention from both the popular media and academics. It has been cited in over 100 stories from over 60 different news outlets, leading to an Altmetric Attention Score greater than every other article published in the *Journal of Politics* and placing it in the top 5% of all research from any discipline in the Altmetric database (Altmetric 2018). It has been cited by 742 academic papers in the Google Scholar database, as of December 31, 2025. In their review of the affective polarization literature, Iyengar et al. (2019, 139) write, "scholars have shown that correcting misperceptions about party supporters reduces animus toward the other side (Ahler and Sood 2018)." This example not only demonstrates a typical reference, but also highlights the central role Ahler and Sood play within the affective polarization literature. Iyengar et al. (2019) argue that correcting misperceptions is one of only two methods shown to reduce affective polarization (as of its 2019 publication). And they cite only Ahler and Sood for evidence of this causal pathway.

Subsequent work has largely, though not entirely, found evidence in support of the claim that correcting misperceptions tends to reduce affective polarization. For instance, Broockman, Kalla, and Westwood (2023) offer a well-powered, partial replication of Ahler and Sood's experiment. In their replication, they find that subjects who receive corrections for their out-party misperceptions, relative to those in a pure control group, tend to have (1)

warmer affect toward the out-party, (2) cooler affect toward the in-party, and hence (3) lesser differences between in-party and out-party affect. Their design cannot estimate the effect of merely asking because it lacks a treatment group that asks respondents about misperceptions without correcting them.

More recently, Voelkel et al. (2024) field a survey experiment to over 30,000 respondents with 25 distinct treatments designed to reduce affective polarization. The 25 treatments were chosen by a group of experts to be the "most promising" among 253 submissions. Of the 25 selected, five of the treatments relied a "great deal" on correcting misperceptions, though none correct the demographic misperceptions that are the focus of Ahler and Sood (2018).[3] Nonetheless, these five treatments each reduced affective polarization relative to a pure control.

Demand for Alternative Interventions

Though a variety of studies have found apparent correction effects, Dias et al. (2024) argue that alternative interventions are needed to meaningfully reduce affective polarization. By their count, only 65% of prior correction studies yield estimates indicating that corrections reduce affective polarization by a statistically significant margin. They also field two experiments examining correction effects on affective polarization.[4] In each experiment, they randomly vary the magnitude of corrections they provide, with some subjects receiving corrections close to their baseline perceptions and others receiving corrections further from their baseline perceptions. They show that larger corrections do not tend to produce statistically significant differences in either in-party affect or out-party affect. Based on these results, (Dias et al. 2024, p. 6) conclude:

> "the present research suggests that correcting out-party misperceptions is not an

[3]And, again, none of those treatments merely ask perceptions without also correcting misperceptions.
[4]They report a total of three experiments, but only two include affective polarization as an outcome.

especially promising intervention. For these reasons, we encourage researchers to determine when corrections work; to stress-test other promising, individual-level interventions; and to explore new, scaleable interventions."

In addition to these concerns, at least two logistical challenges are likely to impede the effectiveness of corrections in some contexts.

1. First, corrections are likely to be less effective when implemented by someone who has yet to develop a credible reputation among the target audience. Since corrections are intended to change beliefs, we should expect their effectiveness hinges on a recipient's assessment of a source's credibility. Source credibility has long played a central role in theories of information processing and persuasion (Petty and Cacioppo 1984; Traberg and van der Linden 2022). Indeed, Ecker et al. (2022) conduct a review of the misinformation correction literature, concluding, "even though credibility matters less for correction sources compared with misinformation sources, corrections are ideally delivered by or associated with high-credibility sources. There is also emerging evidence that corrections are more impactful when they come from a socially connected source (for example, a connection on social media) rather than a stranger."
2. Second, Druckman (2023) demonstrates that corrections are less effective in competitive information environments. That is, corrective messages have limited impact when external actors provide alternative interpretations or contradictory information. Given the conflictual nature of partisan politics, many relevant applications will occur in such contexts.

These challenges highlight the need for interventions that do not require the acceptance and integration of new, surprising information.

The Promise of Merely Asking

Given the need for alternative interventions, we argue that *merely asking* merits further attention. It has a compelling theoretical rationale, and, if the effect is large, important practical benefits.

Theoretical Rationale

Why might merely asking reduce affective polarization? Ahler and Sood (2018, p. 975) theorize:

> [asking about the demographic composition of the outparty] very likely changes people's beliefs about the parties' composition. As we contend above, people's "mental pictures" of parties reflect party-stereotypical groups. These pictures implicitly map to inflated beliefs about shares of party-stereotypical groups. However, being forced to offer explicit, precise numerical estimates of the shares of different groups in the party disrupts the tendency to think about parties in terms of prototypes. It forces people to take stock of some of the large numbers that the pictures in their heads map to and to revise their beliefs to lower, more plausible numbers. Experimental data support such a mechanism: more "considered" responses tend to be less inflated than "automatic" responses (Ahler and Sood 2017).

Ahler and Sood's explanation aligns with research on information processing and social identity. Many prominent models of information processing and judgment rely on dual-process models (e.g., Chaiken 1980; Kahneman 2011; Petty and Cacioppo 1986). These models suggest people rely on one of two modes of processing when evaluating information or forming judgments. The first mode is fast, often automatic or unconscious. The second

mode is slower, and more conscious and deliberate.[5] People typically default to the first mode unless they are given reason to expend more effort. As Ahler and Sood note, if people are asked to stop and think, they tend to rely less heavily on automatic judgments and mental shortcuts. Moreover, people often abandon automatic responses when their accuracy or applicability is called into question (Maheswaran and Chaiken 1991). In this case, merely asking about the composition of the outparty may draw the respondent's attention to their uncertainty about these questions.[6] And thus, this uncertainty may reduce the extremity of antipathy toward the outparty by disrupting the automatic affective response.[7]

Social identity theory offers a complimentary perspective. Drawing from this perspective, Green, Palmquist, and Schickler (2002, p. 8) posit that an individual's partisan identity hinges on their perception of the social groups that support each party:

> "Social identification involves comparing a judgment about oneself with one's perception of a social group. As people reflect on whether they are Democrats or Republicans (or neither), they call to mind some mental image, or stereotype, of what these sorts of people are like and square these images with their own self-conceptions. In effect, people ask themselves two questions: What kinds of social groups come to mind as I think about Democrats, Republicans, and Independents? Which assemblage of groups (if any) best describes me?"

This perspective suggests that corrections reduce affective polarization by altering these perceptions. But merely asking also has the potential to do so by priming some social groups and thus reducing the relative salience of others. When one asks themselves how they feel about the outparty, the social groups they were just asked about are likely to come more

[5] Rather than discrete modes of judgment, these may best be thought of as poles of a continuum (Evans and Stanovich 2013).

[6] In addition, even while overestimating the size of these groups in the outparty, people tend to estimate that each group constitutes only a minority of the party (Ahler and Sood 2018, 967). Therefore, merely asking sends the implicit signal that the majority of the outparty are *not* group members.

[7] Indeed, attitude extremity and certainty have a strong, positive association (Krosnick et al. 1993).

readily to mind (Zaller and Feldman 1992). And, as we discuss above, having been made aware of their own ignorance about these groups, the subsequent affective judgment should be more positive.

Practical Benefits

Many of the same mechanisms that shape responses to corrections are likely to operate in response to merely asking. And if the effect of merely asking is large, it offers a relatively affordable alternative intervention to corrections for the purposes of manipulating affective polarization. We mean this in the following concrete, narrow sense: it takes less survey time to merely ask than to *ask and correct*. For instance, the Cooperative Election Study (CES) uses a point system that translates to one point per second of survey time. According to the 2023 CES definitions, the items asking for respondent perceptions follow a "grid" format, which costs 10 points plus 5 points per item. With four items, this grid costs 30 points. The correction treatment includes those same four items along with about 80 words of text to deliver the corrections. The CES charges one point for four words of informational text and thus the correction treatment would cost an additional 20 points. Therefore, the 30-point merely asking treatment is about 40% shorter than the 50-point correction treatment.

Uncertainty about the Effect of Merely Asking

Yet the value of merely asking as an intervention depends on the existence and magnitude of the effect. And this portion of the Ahler and Sood's design has not been directly replicated. In contrast, Ahler and Sood's correction effect has been replicated carefully by existing work—most recently and closely by Broockman, Kalla, and Westwood (2023).

Despite an absence of direct replication, a number of studies have fielded similar designs,

but differ in several crucial ways.[8] In particular, these studies tend to correct different types of misperceptions or examine outcomes other than partisan affective polarization (e.g., Braley et al. 2023; Druckman 2023; Landry et al. 2023). For instance, Lees and Cikara (2020) correct misperceptions about the out-party's reactions to various policies. Their control group is asked about their perceptions of the out-party and two different treatment groups are asked—and then told—the correct answers. But their outcome variable is an assessment of out-party obstructionism on a specific policy rather than general affect toward the out-party. Nonetheless, they find that the groups receiving the correction evaluate the out-party as less obstructive than do individuals whose perceptions were not corrected. But, without a pure control, the study cannot shed light on the effect of merley asking relative to not asking.

Without direct replication, the evidence for the effect of merely asking rests on Ahler and Sood's original results. While Ahler and Sood find a statistically significant effect for merely asking, their experiment appears under-powered and their evidence might overstate the impact of merely asking about the composition of the out-party on affective polarization.

The estimates that Ahler and Sood report are quite large—they find that merely asking about perceptions of the out-party decreases affective polarization by about 4.9 points on the 101-point feeling thermometer scale. This estimate is about 80% as large as their estimated effect of correcting misperceptions. However, they rely on a relatively small sample, so their confidence intervals are rather wide.

Table 1 shows the statistical power and exaggeration ratio (Gelman and Carlin 2014) for various effect sizes that are both substantively plausible and consistent with their data. The statistical power is the chance of rejecting the null hypothesis of no effect for a particular assumed true effect. The exaggeration ratio (or "Type M error") is how many times larger the expected estimate is than the true effect, filtering out estimates that are not statistically

[8]For reviews of relevant studies, see Bursztyn and Yang (2022); Hartman et al. (2022); Moore-Berg and Hameiri (2024).

significant.

Table 1: Statistical power and exaggeration ratio for plausible effect sizes, based on Ahler and Sood.

Description	Effect Size	Power (Two-Tailed)	Power (One-Tailed)	Exaggeration Ratio
Ahler and Sood's estimate for merely asking	4.9 pts	78%	87%	1.02
Half of Ahler and Sood's estimate for correcting misperceptions	3.2 pts	43%	56%	1.19
Lower bound of Ahler and Sood's 90% CI for merely asking	2.0 pts	20%	30%	2.27

For Ahler and Sood's estimate of the effect of merely asking, their experiment has an adequate power of 87% for a one-tailed test. However, this estimate is surprisingly large—perhaps implausibly so—thus it is important to consider the statistical power for smaller effect sizes. First, we consider an effect of 3.2 points, or half as large as Ahler and Sood's estimate for the effect of correcting misperceptions. This effect is still substantively large, but plausible. In this case, the statistical power drops to 56%, which falls below the typical minimal threshold of 80%. Second, we consider an effect of 2.0 points, which is the lower-bound of Ahler and Sood's 90% confidence interval for the effect of merely asking and about 30% of Ahler and Sood's point estimate correcting misperceptions. Below, we argue 2.0 is the smallest effect of substantive interest. In this scenario, the statistical power shrinks to about 30% for a one-tailed test. The exaggeration ratio is 2.27, which means that statistically significant results tend to overestimate the true effect by about 127%.

Further, in their replication of Ahler and Sood, Broockman, Kalla, and Westwood (2023) find an effect of only 3.9 points on the 101-point scale for correcting these misperceptions. This is a smaller effect than Ahler and Sood find for merely asking. This underscores the importance of carefully reproducing Ahler and Sood's results. From the "replication value" perspective (Isager et al. 2023), these claims are both (1) substantively and theoretically valuable and (2) quite uncertain.

Research Design

We propose a design that largely mimics Ahler and Sood's experiment summarized in their Figure 4A and discussed on pp. 974–978. Our Figure 1 summarizes this design, which includes three conditions: *Tell*, *Ask*, and *Control*. In the *Tell* condition, respondents are asked their perceptions of the out-party and then have any misperceptions corrected; in the *Ask* condition, respondents are merely asked their perceptions of the out-party (without any correction); in the *Control* condition, respondents are not asked their perceptions nor do they receive any correction.

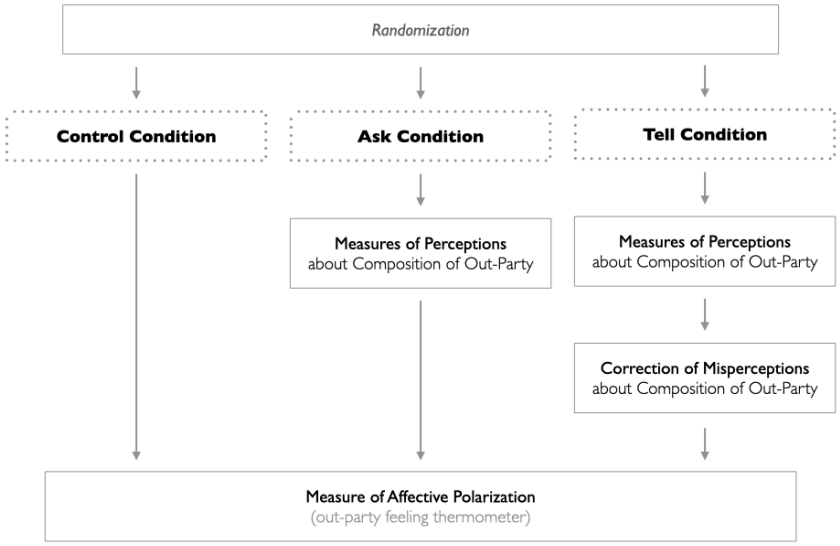


Figure 1: This figure shows Ahler and Sood's (2018) original experimental design.

We make three important changes to their original design: (1) we substantially increase

the sample size of the study, (2) we implement a pre-post and quasi-pre-post design to increase the precision of our estimates, and (3) we use a new measure of affective polarization (in addition to Ahler and Sood's original measure).

This design allows us to test the *Correction Hypothesis* and the *Merely Asking Hypothesis*:

> **The Correction Hypothesis**: The *Tell* treatment decreases affective polarization (relative to the *Control*).
>
> **The *Merely Asking* Hypothesis**: The *Ask* treatment decreases affective polarization (relative to the *Control*).

A replication of the *Correction Hypothesis* is justified by its importance in the affective polarization literature and the doubts raised by Dias et al. (2024). However, these are not our primary reasons for including the *Tell* condition in our experiment. We include the better-studied *Tell* condition in our experiment so that we can compare our results for the less-studied *Ask* condition. Indeed, if we are unable to replicate Ahler and Sood's results for *merely asking*, it will be helpful to know whether (or not) we are able to replicate their results for *correcting* misperceptions.[9] Further, including both conditions will allow us to better inform future work by bounding their relative efficiency.

Experimental Conditions

Following Ahler and Sood (2018), we assign respondents to one of three conditions: *Tell*, *Ask*, and control.

Tell Condition: In the *Tell* condition, we ask respondents to provide their perceptions of the demographic composition of the out-party. We ask Republican respondents to

[9]There are several reasons why we might not be able to replicate Ahler and Sood's original finding in support of the *Correction Hypothesis*. Most importantly, we make two methodological adjustments: a pre-post design and an alternative measure of affective polarization (although we include their original measure as well). While we carefully justify these changes, we must make sure that we can replicate their relatively well-studied hypothesis with these two changes.

report the percentage of Democrats they think are Black; atheist or agnostic; union members; and gay, lesbian, or bisexual. We ask Democratic respondents to report the percentage of Republicans they think are evangelical Christian; earn more than $250,000 per year; live in the South; and are age 65 years or older. We follow up by providing respondents with the correct percentages. We provide the correct response and the percentage difference between their response and the actual percentage.[10] We ask respondents how surprised they are by the correct answers.

Ask Condition: In the *Ask* condition, we ask respondents to provide their perceptions of the demographic composition of the out-party, just as with the *Tell* condition. However, respondents in the *Ask* condition do not receive the correct percentages.

Control Condition: In the control condition, we neither ask respondents their perceptions of the out-party nor do we provide the correct percentages.

Outcomes

We focus on three outcome measures: the out-party feeling thermometer (following Ahler and Sood 2018), the difference between the in-party and out-party feeling thermometer (following Brookman, Kalla, and Westwood 2023), and the recently developed Affective Polarization Scale (Campos and Federico 2025).

As our primary outcome, we follow Ahler and Sood (2018) and use the out-party feeling thermometer. More recently, Druckman et al. (2022) also use the out-party feeling thermometer.

As a secondary outcome, we follow other researchers such as Broockman, Kalla, and Westwood (2023), and use the difference between the in-party and out-party feeling thermometers to measure affective polarization. Broockman, Kalla, and Westwood's data suggest that the difference measure might yield more statistical power compared to the out-party

[10] More information about how we estimate the correct responses can be found in Appendix B.

feeling thermometer.

As another secondary outcome, we use the nine-item Affective Polarization Scale that Campos and Federico (2025) develop using the concept of affective polarization or "political sectarianism" by Finkel et al. (2020). Finkel et al.'s conception of affective polarization has three components: othering, aversion, and moralization. See Landry et al. (2024) for a very similar measure.

Table 2 provides the nine-items.

Table 2: Final 9-Item Affective Polarization Scale

Othering	• I feel as though [in-party] are very different from [out-party]. • [Out-party] live in a different world from us [in-party]. • [Out-party] act in ways that us [in-party] could never understand.
Aversion	• As a [in-party], I would not want to be friends with someone who was a [out-party]. • If I found out a friend of mine was a [out-party], I would want to stop spending time with them. • *Although I do not agree with their political views, there are people I like who are [out-party].
Moralization	• My identity as a [in-party] is connected to my core moral beliefs. • My identity as a [in-party] reflects my beliefs about the difference between right and wrong. • My identity as a [in-party] is rooted in moral principles.

Note: (*) denotes reverse coded items.

Pre-Treatment Measures

Following Clifford, Sheagley, and Piston (2021), we include pre-treatment measures of affective polarization that are closely related to the outcomes. In the survey, we separate

these pre-treatment measures from the experimental treatments and the outcomes by several demographic questions and an unrelated survey experiment vignette.[11] In the analysis, we adjust for these pre-treatment measures (Lin 2013).[12]

We ask identical versions of the feeling thermometer questions pre- and post-treatment, which Clifford, Sheagley, and Piston (2021) refer to as a pre-post design. Importantly, both Clifford, Sheagley, and Piston (2021) and Jordan, Ollerenshaw, and Trexler (2025) show that the attenuation in the treatment effect compared to a post-only design is relatively small.[13]

We also use the three items in Table 3 to create an alternative scale that serves as a quasi-pre-treatment measure of the nine-item scale. As Campos and Federico (2025) show, these three alternative items are closely related, both conceptually and empirically, to the items in their final nine-item scale.[14] Also, the alternative scale uses *different* items, reducing concerns that respondents will simply remember and repeat their answers to the pre-treatment items. However, Clifford, Sheagley, and Piston (2021) and Jordan, Ollerenshaw, and Trexler (2025) show that this should not be a major concern.

[11]The unrelated survey instrument is reported in Appendix C.

[12]Jordan, Ollerenshaw, and Trexler (2025) show that the pre-post design attenuates treatment effects by about 20% compared to a post-only design. But controlling for these pretreatment measures shrinks standard errors by about 50%. In general, the standard errors shrink by about $[1-\sqrt{1-r^2}]\%$, where r is the correlation between the pre- and post-treatment measures (Bloom 1995; Rainey 2025*a*). For the pre-post design to increase statistical power, the (percent) shrinkage in the standard errors must be greater than the (percent) attenuation in the treatment effect. Setting $[1-\sqrt{1-r^2}]\% > 20\%$ and solving for r, we need $|r| > 0.6$. The *smallest* correlation that Jordan, Ollerenshaw, and Trexler (2025) observe across their $6 \times 3 = 18$ designs and samples is about 0.73 and the median is about 0.86. In our pilot data, the correlation between pre- and post-treatment measures is usually above 0.8, which would generate a 40% shrinkage in the standard error. For concreteness, a 20% attenuation in the treatment effect and a 40% shrinkage in the treatment effect would increase the ratio of the treatment effect to the standard error by $(100\% - 20\%)/(100\% - 40\%) = 1.33$, which is equivalent to a 77% increase in sample size. If the ratio for the post-only design were 2.5 (80% power), then the ratio for the pre-post design would increase to 3.3 (95% power). This would decrease the rate of Type II errors by about 75%.

[13]In a pilot study with 250 respondents, we find that the pre-treatment out-party feeling thermometer has a correlation of 0.87 with the post-treatment out-party feeling thermometer (pooling across the three conditions). The pre-treatment difference in feeling thermometers has a correlation of 0.91 with the post-treatment difference in feeling thermometers. Appendix D provides the details. This correlation suggests that the pre-post strategy of Clifford, Sheagley, and Piston will shrink the standard error by about $\left(1-\sqrt{1-0.9^2}\right) \times 100\% = 57\%$.

[14]In the pilot study, the pre-treatment alternative scale has a correlation of 0.78 with the post-treatment scale. Again, see Appendix D for details. Adjusting for a pre-treatment measure with a correlation of 0.8 with the outcome will shrink the standard error by about $\left(1-\sqrt{1-0.8^2}\right) \times 100\% = 40\%$.

Table 3: Alternative Three-Item Affective Polarization Scale

Othering	Republicans/Democrats seem out of place when they are around us [Democrats/Republicans].
Aversion	Identifying as a [Republican/Democrat] rather than a [Democrat/Republican] makes someone a bad person.
Moralization	As a [Democrat/Republican], my feelings about politics are connected to my core moral beliefs.

Figure 2 summarizes the design.

Sample

We will field the experiment to 3,000 US respondents through CloudResearch Connect, paying each respondent $2 to complete the survey.[15] While CloudResearch Connect does not provide representative samples of the United States population, an ongoing literature consistently demonstrates that "results from convenience samples provide estimates of causal effects comparable to those found on population-based samples" (Mullinix et al. 2015, p. 122) and recent work suggests that CloudResearch Connect respondents are especially useful for experiments with a high requirement for attention (Stagnaro et al. 2024).[16]

Estimation and Testing

To estimate the effect of the treatment(s) using the pre-post and quasi-pre-post designs, we estimate the average treatment effect (ATE) with ordinary least squares and estimate the standard errors using the HC2 Huber-White sandwich estimator (MacKinnon and White

[15]Following Ahler and Sood (2018), each pure independent in the sample will be assigned at random to receive either the Republican or Democratic items. In an exploratory analysis, we will also provide the estimates of the treatment effects after dropping pure independents.

[16]Many crowdsourced survey experiments appear in recent issues of top journals. See examples in the *APSR* (Hansen 2023; Graham 2023*b*; Graham and Singh 2023; Zárate, Quezada-Llanes, and Armenta 2023; Kreft and Agerberg 2023; Pérez, Robertson, and Vicũna 2023), *AJPS* (Kim 2023; Virgin 2023; Kobayashi et al. 2023; Powers and Altman 2023; Singh and Tir 2023; Graham 2023*a*), and *JOP* (Andrews, Delton, and Kline 2023; Renshon, Yarhi-Milo, and Kertzer 2023; Abrajano, Elmendorf, and Quinn 2023; Chow and Han 2023; Velez, Porter, and Wood 2023; Elder and Hayes 2023).

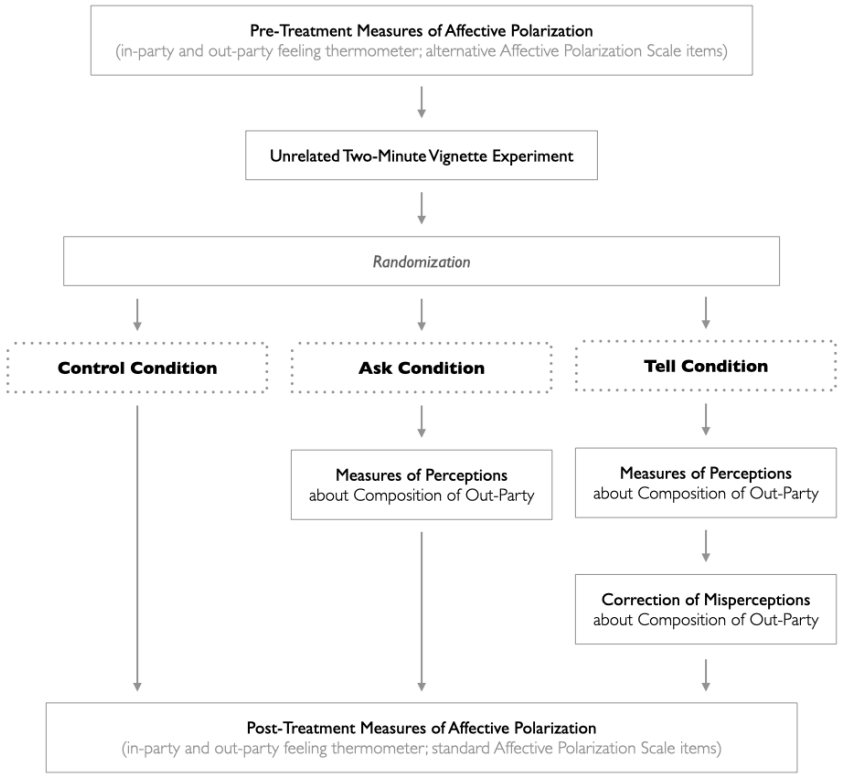


Figure 2: This figure shows our experimental design.

1985). Table 4 shows the control variables we use for each outcome. In each case, we center the control variables at their sample averages and interact each control variable with the treatment (Lin 2013). We use 90% Wald confidence intervals to evaluate our directional hypotheses (McCaskey and Rainey 2015).

Power Analysis

We demonstrate adequate power for our experiment in two ways. First, we compare our sample size to previous research. Second, we use pilot data to conservatively estimate the minimal detectable effect with 80% power, and compare this estimate to the smallest effect

Table 4: This table summarizes the outcome measures, pre-treatment measures that we adjust for in our regression models, and the design. For the measures based on feeling thermometers, we adjust for respondents' answer to an identical question asked pre-treatment as well as the sum of the three alternative scale items. For the Affective Polarization Scale, we use the sum of the nine-item scale and adjust for the difference in feeling thermometer scores as well as the sum of the alternative scale items, giving us a *quasi*-pre-post design.

Outcome	**Control Variables (Measured Pre-Treatment)**	**Design**
Out-Party Feeling Thermometer	(i) Out-Party Feeling Thermometer and (ii) Alternative, Three-Item Affective Polarization Scale	Pre-Post
Difference in Feeling Thermometers	(i) Difference in Feeling Thermometers and (ii) Alternative, Three-Item Affective Polarization Scale	Pre-Post
Affective Polarization Scale	(i) Difference in Feeling Thermometers and (ii) Alternative, Three-Item Affective Polarization Scale	Quasi-Pre-Post

of substantive interest.

We Obtain More Precise Estimates than Earlier Studies

Ahler and Sood's experiment has about 275 respondents per condition. Broockman, Kalla, and Westwood's (2023) replication has about 500 respondents per condition. We plan to have about 1,000 respondents per condition. This should give us $\left(1 - \sqrt{\frac{550}{2{,}000}}\right) \times 100\% = 48\%$ smaller standard errors than Ahler and Sood and $\left(1 - \sqrt{\frac{1000}{2{,}000}}\right) \times 100\% = 29\%$ smaller standard errors than Broockman, Kalla, and Westwood. However, as we explain above, we include the pre-treatment measures to further reduce the standard errors . And we carefully estimate this reduction with a pilot study, described below.

We Have *At Least* 80% Power

To estimate the minimal detectable effect with 80% power using our pre-post and quasi-pre-post designs described in Table 4, we use Rainey's (2025) rules for pilot data. In the planning stage, we collected pilot data that mimic the full study, except with 250 respondents rather

than 3,000. We perform the planned analysis on the pilot data, and then we use the standard errors from the pilot data to conservatively predict the standard errors in the full study. We report the minimum detectable effects (Bloom 1995; Rainey 2025*a*) for 80% and 95% power.

But what effect sizes are we interested in detecting? We are interested in effects as small as about two points on the out-party feeling thermometer scale. For comparison, Ahler and Sood estimate that the effect of the *Ask* treatment is about 4.9 points on the 101-point scale, with a 90% confidence interval of [2.0, 7.8]. For the *Tell* treatment, they estimate an effect of 6.4 points [3.5, 9.3]. Broockman, Kalla, and Westwood (2023) estimate a comparable effect of the *Tell* treatment of 3.9 points [1.2, 6.7]. Given that the purpose of our study is to replicate Ahler and Sood's findings for their *Ask* treatment, the lower-bound of their 90% confidence interval seems like a natural threshold to target and an effect of 2.0 points on the 101-point scale seems substantively reasonable as well.

Table 5 shows that our design is well-powered to detect effects of about two points on the 101-point out-party feeling thermometer scale. *Conservatively*, we have 80% power to detect effects of about 1.64 points, and we have 95% power to detect effects of about 2.16 points.[17]

While we must avoid mindlessly applying hard and fast, generic rules about the smallest effect of substantive interest, it can be helpful to consider effect sizes rescaled to the (pooled, pilot) sample standard deviation or "Cohen's *d*." Conservatively, we have 80% power to detect rescaled effects of $d = 0.07$ for the out-party feeling thermometer, $d = 0.05$ for the difference between the in- and out-party feeling thermometer, and $d = 0.08$ for the Affective Polarization Scale. Cohen (1988) tentatively suggests $d = 0.2$ as a threshold for "small" effects. Lovakov and Agadullina (2021) derive empirical thresholds for social psychology and suggest $d = 0.15$ as a threshold for "small" effects, which falls at their 25th percentile. Effects of $d = 0.11$ and $d = 0.07$ fall at their 20th and 15th percentiles, respectively. By either of these guidelines, we are well-powered to detect "small" effects, though we must treat these

[17]Removing the conservative bias, we have 95% power to detect an effect of about 1.95 points.

Outcome	Adjusting For	Conservative $\text{MDE}_{80\%}$	Conservative $\text{MDE}_{95\%}$
Out-Party FT	*None*	*2.95*	*3.89*
	Variables in Table 4	1.64	2.16
Difference in FT	*None*	*4.04*	*5.33*
	Variables in Table 4	1.65	2.18
Affective Polarization Scale	*None*	*1.3*	*1.72*
	Variables in Table 4	0.77	1.02

Table 5: This table shows our conservative estimates of the minimum detectable effect (MDE) at 80% and 95% power for our planned 3,000-respondent survey experiment. To compute these MDEs, we use use Rainey's (2025) rules for pilot data with our completed 250-respondent pilot study. Based on prior research, we are interested in effects larger than about 2.0 points on the 101-point feeling thermometer scale. Statistically adjusting for pre-treatment measures of the feeling thermometers and our alternative, three-item variant of Campos and Federico's (2025) Affective Polarization Scale allows us to reliably detect this effect. For comparison, Ahler and Sood's original experiment has a minimum detectable effect of about 4.4 points with 80% power—our design offers a meaningful improvement.

context-free, default thresholds with skepticism.

Null Results

In case the treatment effect is not statistically significant, we can use an equivalence test (Rainey 2014) to assess whether we can reject *meaningful* positive effects. Using Rainey's (2025) rules for pilot data, we can conservatively estimate the standard error for the full study as 0.65, which implies a 90% confidence interval with a width of $2 \times 1.64 \times 0.65 = 2.13$. Thus, if the estimated treatment effect is *not* statistically significant, then we will certainly be able to reject effects larger than 2.13 points (and probably much smaller effects), which is about the smallest effect of substantive interest.

Results

The Correction Hypothesis

[Results to be added here.]

The Merely Asking Hypothesis

[Results to be added here.]

Exploratory I: The Difference Between Correcting and *Merely Asking*

Our experiment is not designed to compare the *Ask* and *Tell* conditions—there are substantively important differences between these two conditions that we are not well-powered to detect.[18] However, we can use a 90% confidence interval to bound the comparisons between these conditions (McCaskey and Rainey 2015), which might be *somewhat* informative for future research.

[Results to be added here.]

Exploratory II: The Effects of the Treatments on the Three Components of the Scale

While we do not have specific, theory-driven predictions about the effects of the treatments on the three components of the Affective Polarization Scale, we report estimates of the effects

[18]We have 80% power to detect differences of about 1.64 points on the 101-point scale. For example, if the *Tell* treatment has an effect of three points and the *Ask* treatment has an effect of two points, then the correction has a 50% larger effect than merely asking. This one point increase is substantively meaningful, but we do not have adequate statistical power to detect it reliably.

of both treatments on each of these three outcomes. In each model, we adjust for the sum of the alternative, three-item scale.

Exploratory III: The Effects of the Treatments on Response Latencies

In addition to estimating the effect of the *Ask* and *Tell* on the *responses to* our outcome measures, we explore their effects on the sum of *response times* (in logged milliseconds) to our post-treatment outcome measures. As we noted above, it could be that merely asking respondents to report their perceptions of the out-party disrupts their default fast, automatic, and unconscious affective responses and causes them to respond in a slow, conscious, and deliberate manner. We use the same estimation strategy as the main analysis, adjusting for the sum of response times across the pre-treatment measures. For this analysis, we drop outliers on the post-treatment latencies. On the low end, we define outliers as those who spend less than 200 milliseconds per word of survey text (Zhang et al. 2020). On the high end, we define outliers as those whose response times exceed the third quartile + 2.5 $\times$ the inter-quartile range (see Vankov 2023).

Exploratory IV: Dropping Inattentive Respondents

Ahler & Sood do not drop inattentive respondents and thus our primary analysis will not either. But we will report exploratory analysis that re-estimates the treatment effects after dropping low-attention respondents. To distinguish low-attention respondents from medium- and high-attention respondents, Berinsky et al. (2021) recommend including multiple attention items formatted in a grid. We therefore include a grid of attention items immediately after the pre-treatment measures (see Appendix A for item text). Following their recommendations, we scale these items using a two-parameter item response theory model and report estimates

after dropping the bottom quintile of attention.

Exploratory V: Dropping Independents

Ahler & Sood randomly assign independents to an out-party and in-party. Our primary analysis follows this choice. But we will report an exploratory analysis that re-estimates the treatment effects with independents removed.

Conclusion

[Conclusion to be added here after data collection.]

Ethical Standards

The authors declare the human subjects research in this article will be reviewed and approved by the appropriate human subjects board.

References

Abrajano, Marisa, Christopher S. Elmendorf, and Kevin M. Quinn. 2023. "Measuring Perceived Skin Color: Spillover Effects and Likert-Type Scales." *The Journal of Politics* 85(1): 320–327.

Ahler, Douglas J., and Gaurav Sood. 2017. "Measuring Perceptions of Shares of Salient and Stereotypical Groups." In *Misinformation and Mass Audiences*, ed. Laura Sheble Brian G. Southwell, Emily Thorson. Austin: University of Texas Press , 71–90.

Ahler, Douglas J., and Gaurav Sood. 2018. "The Parties in Our Heads: Misperceptions about Party Composition and Their Consequences." *The Journal of Politics* 80(3): 964–981.

Altmetric. 2018. "The Parties in Our Heads: Misperceptions about Party Composition and Their Consequences." Altmetric.

Andrews, Talbot M., Andrew W. Delton, and Reuben Kline. 2023. "Who Do You Trust? Institutions That Constrain Leaders Help People Prevent Disaster." *The Journal of Politics* 85(1): 64–75.

Berinsky, Adam J., Michele F. Margolis, Michael W. Sances, and Christopher Warshaw. 2021. "Using Screeners to Measure Respondent Attention on Self-Administered Surveys: Which Items and How Many?" *Political Science Research and Methods* 9(2): 430–437.

Bloom, Howard S. 1995. "Minimum Detectable Effects." *Eval. Rev.* 19(5): 547–556.

Braley, Alia, Gabriel S. Lenz, Dhaval Adjodah, Hossein Rahnama, and Alex Pentland. 2023. "Why Voters Who Value Democracy Participate in Democratic Backsliding." *Nature Human Behaviour* 7(8): 1282–1293.

Broockman, David E., Joshua L. Kalla, and Sean J. Westwood. 2023. "Does Affective Polarization Undermine Democratic Norms or Accountability? Maybe Not." *American Journal of Political Science* 67(3): 808–828.

Bursztyn, Leonardo, and David Y. Yang. 2022. "Misperceptions About Others." *Annual Review of Economics* 14(1): 425–452.

Campos, Nicolas, and Christopher Federico. 2025. "A New Measure of Affective Polarization." *American Political Science Review*: Forthcoming.

Chaiken, Shelly. 1980. "Heuristic versus systematic information processing and the use of source versus message cues in persuasion." *Journal of personality and social psychology* 39(5): 752.

Chow, Wilfred M., and Enze Han. 2023. "Descriptive Legitimacy and International Organizations: Evidence from United Nations High Commissioner for Refugees." *The Journal of Politics* 85(2): 357–371.

Clifford, Scott, Geoffrey Sheagley, and Spender Piston. 2021. "Increasing Precision without Altering Treatment Effects: Repeated Measures Designs in Survey Experiments." *American Political Science Review* 115(3): 1048–1065.

Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences.* 2 ed. London, England: Routledge.

Dias, Nicholas C, Laurits F Aarslew, Kristian Vrede Skaaning Frederiksen, Yphtach Lelkes, Lea Pradella, and Sean J Westwood. 2024. "Correcting Misperceptions of Partisan Opponents Is Not Effective at Treating Democratic Ills." *PNAS Nexus* 3(8).

Druckman, James N. 2023. "Correcting Misperceptions of the Other Political Party Does Not Robustly Reduce Support for Undemocratic Practices or Partisan Violence." *Proceedings of the National Academy of Sciences* 120(37): e2308938120.

Druckman, James N., and Jeremy Levy. 2022. "Chapter 18: Affective Polarization in the American Public." In *Handbook on Politics and Public Opinion*, ed. Thomas Rudolph. Northampton, Massachusetts: Edward Elgar Publishing, Inc. , 257–270.

Druckman, James N., Samara Klar, Yanna Krupnikov, Matthew Levendusky, and John Barry Ryan. 2022. "(Mis)estimating Affective Polarization." *The Journal of Politics* 84(2): 1106–1117.

Ecker, Ullrich K. H., Stephan Lewandowsky, John Cook, Philipp Schmid, Lisa K. Fazio, Nadia Brashier, Panayiota Kendeou, Emily K. Vraga, and Michelle A. Amazeen. 2022. "The Psychological Drivers of Misinformation Belief and Its Resistance to Correction." *Nature Reviews Psychology* 1(1): 13–29.

Elder, Elizabeth Mitchell, and Matthew Hayes. 2023. "Signaling Race, Ethnicity, and Gender with Names: Challenges and Recommendations." *The Journal of Politics* 85(2): 764–770.

Evans, Jonathan St BT, and Keith E Stanovich. 2013. "Dual-process theories of higher cognition: Advancing the debate." *Perspectives on psychological science* 8(3): 223–241.

Feldman, Gilad. 2025. "The Value of Replications Goes Beyond Replicability and Is Associated With the Value of the Research It Replicates: Commentary on Isager et al. (2024)." Center for Open Science. https://osf.io/35fh8

Finkel, Eli J., Christopher A. Bail, Mina Cikara, Peter H. Ditto, Shanto Iyengar, Samara Klar, Lilliana Mason, Mary C. McGrath, Brendan Nyhan, David G. Rand, Linda J. Skitka, Joshua A. Tucker, Jay J. Van Bavel, Cynthia S. Wang, and James N. Druckman. 2020. "Political Sectarianism in America." *Science* 370(6516): 533–536.

Gelman, Andrew, and John Carlin. 2014. "Beyond Power Calculations." *Perspectives on Psychological Science* 9(6): 641–651.

Graham, Matthew H. 2023*a*. "Detecting and Deterring Information Search in Online Surveys." *American Journal of Political Science*: Forthcoming.

Graham, Matthew H. 2023*b*. "Measuring Misperceptions?" *American Political Science Review* 117(1): 80–102.

Graham, Matthew H., and Shikhar Singh. 2023. "An Outbreak of Selective Attribution: Partisanship and Blame in the COVID-19 Pandemic." *American Political Science Review*: Forthcoming.

Green, Donald, Bradley Palmquist, and Eric Schickler. 2002. *Partisan Hearts & Minds: Political Parties and the Social Identities of Voters*. New Haven, CT: Yale University Press.

Hansen, Kristina Jessen. 2023. "Greed, Envy, and Admiration: The Distinct Nature of Public Opinion about Redistribution from the Rich." *American Political Science Review* 117(1): 217–234.

Hartman, Rachel, Will Blakey, Jake Womick, Chris Bail, Eli J. Finkel, Hahrie Han, John Sarrouf, Juliana Schroeder, Paschal Sheeran, Jay J. Van Bavel, Robb Willer, and Kurt Gray. 2022. "Interventions to Reduce Partisan Animosity." *Nature Human Behaviour* 6(9): 1194–1205.

Isager, Peder M, Anna E van 't Veer, and Daniël Lakens. 2021. "Replication Value as a Function of Citation Impact and Sample Size." MetaArXiv. https://osf.io/preprints/metaarxiv/knjea_v1

Isager, Peder Mortvedt, Robbie C M van Aert, Štěpán Bahník, Mark J Brandt, K Andrew DeSoto, Roger Giner-Sorolla, Joachim I Krueger, Marco Perugini, Ivan Ropovik, Anna E van 't Veer, Marek Vranka, and Daniël Lakens. 2023. "Deciding What to Replicate: A Decision Model for Replication Study Selection Under Resource and Knowledge Constraints." *Psychololological Methods* 28(2): 438–451.

Iyengar, Shanto, Yphtach Lelkes, Matthew Levendusky, Neil Malhotra, and Sean J. Westwood. 2019. "The Origins and Consequences of Affective Polarization in the United States." *Annual Review of Political Science* 22(1): 129–146.

Jordan, Diana, Trent Ollerenshaw, and Andrew Trexler. 2025. "Repeated Measure Designs are Superior for (Most) Experimental Survey Research Applications." Unpublished working paper. https://atrexler.com/papers/Jordan_Repeated_2025.pdf

Kahneman, Daniel. 2011. *Thinking, fast and slow*. macmillan.

Kim, Eunji. 2023. "Entertaining Beliefs in Economic Mobility." *American Journal of Political Science* 67(1): 39–54.

Kobayashi, Yoshiharu, Menevis Cilizoglu, Tobias Heinrich, and William Christiansen. 2023. "No Entry in a Pandemic: Public Support for Border Closures." *American Journal of Political Science*: Forthcoming.

Kreft, Anne-Kathrin, and Mattias Agerberg. 2023. "Imperfect Victims? Civilian Men, Vulnerability, and Policy Preferences." *American Political Science Review*: Forthcoming.

Krosnick, Jon A, David S Boninger, Yao C Chuang, Matthew K Berent, and Catherine G Carnot. 1993. "Attitude strength: One construct or many related constructs?" *Journal of personality and social psychology* 65(6): 1132.

Landry, Alexander, Eli Finkel, Rick H. Hoyle, James Druckman, and Jay Joseph Van Bavel. 2024. "Partisan Antipathy and the Erosion of Democratic Norms." Center for Open Science. https://dx.doi.org/10.31234/osf.io/ahgy6

Landry, Alexander P., Jonathan W. Schooler, Robb Willer, and Paul Seli. 2023. "Reducing Explicit Blatant Dehumanization by Correcting Exaggerated Meta-Perceptions." *Social Psychological and Personality Science* 14(4): 407–418.

Lees, Jeffrey, and Mina Cikara. 2020. "Inaccurate Group Meta-Perceptions Drive Negative out-Group Attributions in Competitive Contexts." *Nature Human Behaviour* 4(3): 279–286.

Lin, Winston. 2013. "Agnostic Notes on Regression Adjustments to Experimental Data: Reexamining Freedman's Critique." *The Annals of Applied Statistics* 7(1).

Lovakov, Andrey, and Elena R Agadullina. 2021. "Empirically Derived Guidelines for Effect Size Interpretation in Social Psychology." *European Journal of Social Psychology* 51(3): 485–504.

MacKinnon, James G., and Halbert White. 1985. "Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties." *Journal of Econometrics* 29(3): 305–325.

Maheswaran, Durairaj, and Shelly Chaiken. 1991. "Promoting systematic processing in low-motivation settings: effect of incongruent information on processing and judgment." *Journal of personality and social psychology* 61(1): 13.

Marsh, Wayde Z. C. 2023. "The Devil No More? Decreasing Negative Outparty Affect through Asymmetric Partisan Thinking." *Public Opinion Quarterly* 87(1): 170–186.

McCaskey, Kelly, and Carlisle Rainey. 2015. "Substantive Importance and the Veil of Statistical Significance." *Statistics, Politics, and Policy* 6: 77–96.

Moore-Berg, Samantha L., and Boaz Hameiri. 2024. "Improving Intergroup Relations with Meta-Perception Correction Interventions." *Trends in Cognitive Sciences* 28(3): 190–192.

Mullinix, Kevin J., Thomas J. Leeper, James N. Druckman, and Jeremy Freese. 2015. "The Generalizability of Survey Experiments." *Journal of Experimental Political Science* 2(2): 109–138.

Petty, Richard E., and John T. Cacioppo. 1984. "Source Factors and the Elaboration Likelihood Model of Persuasion." *Advances in Consumer Research* 11(1).

Petty, Richard E, and John T Cacioppo. 1986. "Message elaboration versus peripheral cues." In *Communication and persuasion: Central and peripheral routes to attitude change*. Springer , 141–172.

Powers, Kathleen E., and Dan Altman. 2023. "The Psychology of Coercion Failure: How Reactance Explains Resistance to Threats." *American Journal of Political Science* 67(1): 221–238.

Pérez, Efrén, Crystal Robertson, and Bianca Vicũna. 2023. "Prejudiced When Climbing Up or When Falling Down? Why Some People of Color Express Anti-Black Racism." *American Political Science Review* 117(1): 168–183.

Rainey, Carlisle. 2014. "Arguing for a Negligible Effect." *American Journal of Political Science* 58(4): 1083–1091.

Rainey, Carlisle. 2025*a*. "Power Rules: Practical Statistical Power Calculations." Center for Open Science. https://doi.org/10.31219/osf.io/5am9q

Rainey, Carlisle. 2025*b*. "Use and Misuse of a Fast Approximation: Not a Criticism, but a Caution." MetaArXiv. https://osf.io/preprints/metaarxiv/8z45v

Renshon, Jonathan, Keren Yarhi-Milo, and Joshua D. Kertzer. 2023. "Democratic Reputations in Crises and War." *The Journal of Politics* 85(1): 1–18.

Ruggeri, Kai, Bojana Većkalov, Lana Bojanić, Thomas L. Andersen, Sarah Ashcroft-Jones, Nélida Ayacaxli, Paula Barea-Arroyo, Mari Louise Berge, Ludvig D. Bjørndal, Aslı Bursalıoğlu, Vanessa Bühler, Martin Čadek, Melis Çetinçelik, Georgia Clay, Anna Cortijos-Bernabeu, Kaja Damnjanović, Tatianna M. Dugue, Maya Esberg, Celia Esteban-Serna, Ezra N. Felder, Maja Friedemann, Darianna I. Frontera-Villanueva, Patricia Gale, Eduardo Garcia-Garzon, Sandra J. Geiger, Leya George, Allegra Girardello, Aleksandra Gracheva, Anastasia Gracheva, Marquis Guillory, Marlene Hecht, Katharina Herte, Barbora Hubená, William Ingalls, Lea Jakob, Margo Janssens, Hannes Jarke, Ondřej Kácha, Kalina Nikolova Kalinova, Ralitsa Karakasheva, Peggah R. Khorrami, Žan Lep, Samuel Lins, Ingvild S. Lofthus, Salomé Mamede, Silvana Mareva, Mafalda F. Mascarenhas, Lucy McGill, Sara Morales-Izquierdo, Bettina Moltrecht, Tasja S. Mueller, Marzia Musetti, Joakim Nelsson, Thiago Otto, Alessandro F. Paul, Irena Pavlović, Marija B. Petrović, Dora Popović, Gerhard M. Prinz, Josip Razum, Ivaylo Sakelariev, Vivian Samuels, Inés Sanguino, Nicolas Say, Jakob Schuck, Irem Soysal, Anna Louise Todsen, Markus R. Tünte, Milica Vdovic, Jáchym Vintr, Maja Vovko, Marek A. Vranka, Lisa Wagner, Lauren Wilkins, Manou Willems, Elizabeth Wisdom, Aleksandra Yosifova, Sandy Zeng, Mahmoud A. Ahmed, Twinkle Dwarkanath, Mina Cikara, Jeffrey Lees, and Tomas Folke. 2021. "The General Fault in Our Fault Lines." *Nature Human Behaviour* 5(10): 1369–1380.

Singh, Shane P., and Jaroslav Tir. 2023. "Threat-Inducing Violent Events Exacerbate Social Desirability Bias in Survey Responses." *American Journal of Political Science* 67(1): 154–169.

Stagnaro, Michael N, James Druckman, adam berinsky, Antonio A Arechar, Robb Willer, and David G Rand. 2024. "Representativeness Versus Attentiveness: A Comparison Across Nine Online Survey Samples." PsyArXiv. https://osf.io/preprints/psyarxiv/h9j2d

Traberg, Cecilie Steenbuch, and Sander van der Linden. 2022. "Birds of a Feather Are Persuaded Together: Perceived Source Credibility Mediates the Effect of Political Bias on Misinformation Susceptibility." *Personality and Individual Differences* 185: 111269.

Vankov, Ivan I. 2023. "The Hazards of Dealing with Response Time Outliers." *Frontiers in Psychology* 14: 1220281.

Velez, Yamil R., Ethan Porter, and Thomas J. Wood. 2023. "Latino-Targeted Misinformation and the Power of Factual Corrections." *The Journal of Politics* 85(2): 789–794.

Virgin, Sheahan G. 2023. "The Effect of Core Values on Support for Electoral Reform: Evidence from Two Survey Experiments." *American Journal of Political Science* 67(1): 239–256.

Voelkel, Jan G., Michael N. Stagnaro, James Y. Chu, Sophia L. Pink, Joseph S. Mernyk, Chrystal Redekopp, Isaias Ghezae, Matthew Cashman, Dhaval Adjodah, Levi G. Allen, L. Victor Allis, Gina Baleria, Nathan Ballantyne, Jay J. Van Bavel, Hayley Blunden, Alia Braley, Christopher J. Bryan, Jared B. Celniker, Mina Cikara, Margarett V. Clapper, Katherine Clayton, Hanne Collins, Evan DeFilippis, Macrina Dieffenbach, Kimberly C. Doell, Charles Dorison, Mylien Duong, Peter Felsman, Maya Fiorella, David Francis, Michael Franz, Roman A. Gallardo, Sara Gifford, Daniela Goya-Tocchetto, Kurt Gray, Joe Green, Joshua Greene, Mertcan Güngör, Matthew Hall, Cameron A. Hecht, Ali Javeed, John T. Jost, Aaron C. Kay, Nick R. Kay, Brandyn Keating, John Michael Kelly, James R. G. Kirk, Malka Kopell, Nour Kteily, Emily Kubin, Jeffrey Lees, Gabriel Lenz, Matthew Levendusky, Rebecca Littman, Kara Luo, Aaron Lyles, Ben Lyons, Wayde Marsh, James Martherus, Lauren Alpert Maurer, Caroline Mehl, Julia Minson, Molly Moore, Samantha L. Moore-Berg, Michael H. Pasek, Alex Pentland, Curtis Puryear, Hossein Rahnama, Steve Rathje, Jay Rosato, Maytal Saar-Tsechansky, Luiza Almeida Santos, Colleen M. Seifert, Azim Shariff, Otto Simonsson, Shiri Spitz Siddiqi, Daniel F. Stone, Palma Strand, Michael Tomz, David S. Yeager, Erez Yoeli, Jamil Zaki, James N. Druckman, David G. Rand, and Robb Willer. 2024. "Megastudy Testing 25 Treatments to Reduce Antidemocratic Attitudes and Partisan Animosity." *Science* 386(6719): eadh4764.

Zaller, John, and Stanley Feldman. 1992. "A simple theory of the survey response: Answering questions versus revealing preferences." *American journal of political science*: Forthcoming.

Zhang, Chan, Christopher Antoun, H. Yanna Yan, and Frederick G. Conrad. 2020. "Professional Respondents in Opt-in Online Panels: What Do We Really Know?" *Social Science Computer Review* 38(6): 703–719.

Zárate, Marques G., Enrique Quezada-Llanes, and Angel D. Armenta. 2023. "Se Habla Español: Spanish-Language Appeals and Candidate Evaluations in the United States." *American Political Science Review*: Forthcoming.