


ARTICLE

Evidence Can Change Partisan Minds but Less So in Hostile Contexts

Jin Woo Kim 

School of Communication, Kookmin University, Seoul, South Korea
Email: jinwoo01@gmail.com

(Received 19 July 2023; revised 11 July 2024; accepted 5 November 2024)

Abstract

A large body of literature indicates that partisan-motivated reasoning drives resistance to political persuasion. But recent scholarship has challenged this view, suggesting that people don't always resist uncongenial information, and even when they do, it is not clear why. In this article, I present two survey experiments that examine when and why partisans selectively dismiss uncongenial information. The findings show that, in the absence of affective triggers, partisans were persuaded by both congenial and uncongenial information. But when randomly induced to feel adversarial, they became more dismissive of uncongenial information and ultimately disagreed more, not less, after considering the same information. These results (1) identify a crucial condition that provokes resistance to political persuasion; (2) demonstrate partisan-motivated reasoning more clearly than previous studies; and (3) underscore the importance of the quality of elite-level political discourse in determining the quality of citizen-level opinion formation.

Keywords: partisan motivated reasoning; resistance to persuasion; hostile context; affective polarization; survey experiment

Donald Trump famously said that he could shoot someone on Fifth Avenue and not lose any voters. Many observers of American politics have commented that this is 'one of the truest things he's ever said' (Remnick 2017). No amount of evidence against Trump will change the minds of his loyalists. While few may *literally* believe these claims, there is wider acceptance of what is implied in them: partisan minds are stubborn things – and facts cannot change them. These views are frequently expressed in both academic and popular discourse (for example, Klein 2020).

A prominent explanation for such biased resistance to information is partisan-motivated reasoning (Kunda 1990; Taber and Lodge 2006). According to this account, political opinions are 'hot cognitions' charged with affects, which automatically beget the motivation to defend. The impetus to uphold one's partisan viewpoints, known as directional motivation, results in the dismissal of counter-attitudinal information and uncritical acceptance of pro-attitudinal information. This propensity is thought to foster political misperceptions (Flynn et al. 2017), incite resistance to corrective information (Nyhan and Reifler 2010), exacerbate political polarization (Bail et al. 2018), and, consequently, undermine democratic accountability (Arceneaux and Vander Wielen 2017; Little et al. 2020). Partisan-motivated reasoning is particularly concerning, as it is often thought to be pervasive or even inevitable. As Lodge and Taber (2000, 184) put it, 'all political objects are affectively charged' and therefore 'most, if not all,

citizens will be biased reasoners, finding it nearly impossible to evaluate *any* new information in an evenhanded way' (emphasis added).

Nevertheless, prior work as a whole does not seem to justify the claims of pervasive or inevitable partisan-motivated reasoning (Coppock 2023; Druckman and McGrath 2019; Little 2021; Tappin et al. 2020). First, while frequently cited studies (for example, Bail et al. 2018; Hart and Nisbet 2012; Lord et al. 1979; Nyhan and Reifler 2010; Taber et al. 2009; Taber and Lodge 2006) found that exposure to discordant information was ineffective or even counterproductive in persuading people with standing viewpoints, recent experimental studies have demonstrated that partisans update toward new information, even when it contradicts their opinions (Coppock 2023; Guess and Coppock 2020; Tappin et al. 2023; Wood and Porter 2019). Why do citizens sometimes appear to be closed-minded, but not always? Though a growing number of studies grapple with this question (for example, Bolsen et al. 2014; Christensen and Moynihan 2024; Groenendyk and Krupnikov 2021; Redlawsk et al. 2010), this puzzle has not been fully resolved.

Moreover, some scholars have argued that even when partisans selectively resist or dismiss uncongenial information, it is not evident whether partisan-motivated reasoning is the *cause* of such reactions (see Baron and Jost 2019; Coppock 2023; Druckman and McGrath 2019; Gerber and Green 1999; Graham and Singh 2024; Little 2021; Tappin et al. 2020). This complexity arises primarily because most previous work did not randomize the key independent variable – directional motivation. Therefore, critics argue, people might reject or dismiss counter information not because they *want* to maintain their views but because they find the information irrelevant to their core values, untrustworthy given their knowledge of the issue, or both. As such, the observation that partisans dismiss incongruent information is never sufficient evidence of partisan-motivated reasoning.

In this article, I address these common limitations in the study of motivated reasoning, based on two well-powered survey experiments (Total $N = 6020$) that investigate how partisans update their opinions after considering new information that either supports or opposes the Affordable Care Act (ACA). These experiments are designed to clarify (1) when partisans selectively resist uncongenial information and (2) whether directional goals – instead of confounding factors – are the cause of resistance to persuasion. Specifically, I test the following theoretical expectations derived from the motivated reasoning theory.

First, I examine whether information about a polarizing issue by itself breeds partisan resistance (Lodge and Taber 2000). Second, I probe whether people become more prone to biased resistance when presented with ambiguous information (see Kunda 1990, 482–83; Leeper and Slothuus 2014, 145; Taber and Lodge 2006, 757) by manipulating the strength of the presented evidence. Third, I consider multiple outcome variables, examining whether partisans maintain or intensify their initial *attitudes* toward the ACA or Obama, even when they adjust their factual *beliefs* about the law to reflect new information (Bisgaard 2019; Graham and Singh 2024; Nyhan et al. 2020). Fourth, using a four-wave panel design, I explore whether people remember (and forget) information in biased ways, such that the persuasive effects of incongruent information dissipate over time, while those of congruent information persist (Bartels, 2014; Zimmermann, 2020). Finally, and most importantly, I test the hypothesis that hostile political contexts amplify partisan resistance to information (Bolsen et al. 2014; Groenendyk and Krupnikov 2021; MacKuen et al. 2010), by randomly altering partisan sentiments before presenting information treatments that are either civil or uncivil.

Collectively, these approaches permit a comprehensive search for biased resistance to persuasion, which may emerge only under specific conditions and thus go unnoticed in a standard survey experiment that (1) does not systematically vary the quality of evidence, (2) focuses solely on changes in factual beliefs but not in attitudes, (3) considers only the immediate effect, or (4) randomizes either exposure to information or political contexts, *but not both*. Many previous studies address some of these limitations (for example, Bartels 2014; Bolsen et al. 2014; Coppock 2023; Graham and Singh 2024; Groenendyk and Krupnikov 2021; Klar 2014). But none,

to my knowledge, addresses them all. By investigating whether partisans' selective tendencies to resist uncongenial information differ across variations in treatment, outcome, timing of outcome measurement, and hostile vs. non-hostile contexts, I assess the extent to which evidence of partisans' openness to persuasion (for example, Guess and Coppock 2020; Wood and Porter 2019) generalizes across these dimensions. This way, I systematically explore various possible explanations for the discrepant findings in the literature.

Across the two studies, I find that, *when unprovoked*, Democrats and Republicans updated their opinions in the same direction – and often reduced their disagreement. This finding was generally consistent across different evidence strengths (weak versus strong) and different outcome variables (beliefs versus attitudes). Furthermore, I find no evidence of asymmetrical decay in these persuasive effects; partisans forgot congenial information just as easily as uncongenial information. However, when partisans were randomly induced to feel adversarial toward the opposing party before receiving the information treatment, or when the presented treatment message included derogatory remarks, they became less receptive to uncongenial information and more receptive to congenial information. As a result, Democrats and Republicans disagreed more – not less – after considering the same information. I use a simulation approach to show that the opinion changes observed in non-hostile contexts corresponded almost perfectly to how *unbiased* Bayesians are expected to update their views, but the same was not true in hostile contexts.

These findings make three important contributions. First, they suggest that biased resistance to persuasion may be triggered not by incongruent political *information* itself but by the hostile and contentious *contexts* in which the information is encountered. This offers a reconciling explanation for a critical discrepancy in the literature (for example, Bail et al. 2018; Guess and Coppock 2020; Taber and Lodge 2006; Wood and Porter 2019) by demonstrating that two things can be true at the same time: (1) partisans are open to persuasion, *and* (2) they routinely engage in motivated reasoning within today's information environment, which is marked by intense partisan hostility and toxic rhetoric (Bail 2022; Berry and Sobieraj 2013; Jamieson and Cappella 2008; Kim et al. 2021; Levendusky 2013).

Second, the randomization of partisan antipathy allows me to separate the effect of 'hot cognition' on partisans' divergent responses to new information from the effects of prior values or knowledge that led people to divergent opinions in the first place. By demonstrating that otherwise similar partisans remain open to persuasion when partisan animosity is not provoked, but less so when affective polarization is intensified, I provide unusually clear evidence of partisan-motivated reasoning (see Little 2021).

Finally, from a normative standpoint, the results underscore the importance of elite-level political discourse in shaping the quality of citizen-level opinion formation. While partisan-motivated reasoning is often seen as a sign of citizens' incapacity for rational thought, my findings suggest that its primary cause may stem from political elites' failure (or refusal) to foster a constructive political information environment (Berry and Sobieraj 2013; Jamieson and Cappella 2008; Levendusky 2013).

Before proceeding, I define key terms and specify the scope of this study, given that there are multiple ways in which people can engage in motivated reasoning (Little et al. 2020; Taber and Lodge 2006). I define 'partisan resistance' as the selective tendency to reject uncongenial information while accepting congenial information during opinion updating; and 'partisan-motivated reasoning' as the causal process in which directional motives drive partisan resistance to persuasion. The focus of this study is exclusively on whether and how people change their minds *after* being exposed to *valenced information* – defined as arguments and evidence that have positive or negative implications for the performance of public officials or their policies. Therefore, the claims made in this article do not necessarily extend to other forms of partisan-motivated reasoning, such as the tendencies to avoid uncongenial information altogether (Kim and Kim 2021; Peterson and Iyengar 2021) or rely on party cues in lieu of policy details (Bakker et al. 2020; Bolsen et al. 2014; Bullock 2011).

Potential Causes of Partisan Resistance to Persuasion

Generally, previous studies on partisan-motivated reasoning expose participants to new information and measure their interpretations of the information or the resulting attitude/belief changes. In one of the best-known studies, Taber and Lodge (2006, 756) found that people ‘overly accommodate supportive evidence while dismissing out-of-hand evidence that challenges their prior attitudes’, which led them to strengthen their attitudes after considering a balanced set of pro and con arguments (see also Lord et al. 1979). More recently, Bail et al. (2018) found that partisans randomly assigned to receive uncongenial information on Twitter reinforced their partisan viewpoints (see also Bolsen and Druckman 2018; Hart and Nisbet 2012; Nyhan and Reifler 2010). Collectively, these studies suggest that ‘once attitudes have become crystallized, persuasion is difficult’ (Taber and Lodge 2006, 757) and inviting those with opposing views to consider the same information can *increase* their disagreement (see also Bartels 2002; Jerit and Barabas 2012).

These findings of *partisan resistance* to persuasion are often taken as direct evidence of *partisan-motivated reasoning*. To demonstrate motivated reasoning, however, it is necessary to show not only that partisans resist new information in biased ways, but also that these biases are caused by a desire to reach a particular conclusion (Druckman and McGrath 2019; Little 2021; Tappin et al. 2020). Most previous studies are not equipped to verify this causal relationship because they did not randomize the key independent variable, the motivation to defend a particular position, which makes it difficult to determine the cause of biased resistance.

Most notably, resistance to persuasion may stem from value preferences rather than directional goals (see Gerber and Green 1999, 206). For example, consider the following passage included in one of Taber and Lodge’s (2006) stimuli:

The Bill of Rights guarantees the right of all citizens to bear arms. . . . A national council reported in 1991 that handgun accidents killed less than 15 children under the age of 6. This number is minuscule when compared to the total number of accidental deaths of young children.

A gun-control supporter may reject this argument, not because they are motivated to support their party’s position but because they value the lives of young children more than the Second Amendment. This sort of confounding is particularly likely to emerge when the content of manipulated information is not closely matched across different positions – for example, in studies where pro-Republican information appeals to a value that Democrats may find irrelevant and vice versa (for example, Edwards and Smith 1996; Taber and Lodge 2006). In such cases, even Democrats and Republicans who lack directional motives can plausibly exhibit differential reactions to the same information.

Another possibility is that prior knowledge or beliefs, not directional motives, trigger selective scepticism and resistance, where people scrutinize or reject uncongenial information because they find it untrustworthy or inaccurate based on what (they think) they already know (see Druckman and McGrath 2019). Lord et al. (1979) provide a fitting analogy: an oncologist may dismiss news about herbal cures for cancer, not because they want herbal cures to be ineffective but because their oncological knowledge leads them to find the news untrustworthy. Likewise, observations that partisans selectively scrutinize or dismiss information that runs counter to their political positions (see Guay and Johnston 2020; Kahan et al. 2017; Lord et al. 1979; Taber and Lodge 2006) could reflect people evaluating the validity of new information in light of prior knowledge and discarding seemingly questionable information.

Distinguishing between these alternative explanations requires a research design that can separate the effect of partisans’ defensive goals from the effects of the various factors that shaped their opinions in the first place. I implement two strategies. First, I examine the effects of a set of

ACA-related information, which varies in direction and strength of evidence but focuses on a single sub-topic – healthcare costs – that Democrats and Republicans have reported as equally important (Druckman et al. 2012). This design minimizes the possibility that partisan differences in values lead Democrats and Republicans to update their opinions in divergent ways, thereby reducing ambiguity in detecting partisan bias.

Second, in one of my experiments, I randomly intensify partisan sentiments before presenting information treatments that used either civil or uncivil language. Although prior research does not provide clear guidance on how to manipulate the directional goal *per se* (see Bolsen et al. 2014, 244; Tappin et al. 2020, 85), randomizing affects toward political parties seems to be the closest alternative (see Bolsen et al. 2014; Klar 2014; Lavine et al. 2012), given that directional motives are theorized to stem from affective judgements (Lodge and Taber 2000, 183–184). This approach enables me to assess whether partisans, with *identical* prior knowledge and values, respond to new information differently when partisan animosity is heightened or not – a test that is not susceptible to the alternative explanations considered here.

Theoretical Conditions of Partisan Resistance

As noted earlier, a growing number of experimental studies have shown that people do update their opinions toward contrary information (Coppock et al. 2023; Guess and Coppock 2020; Tappin et al. 2023; Wood and Porter 2019), in contrast to other studies finding partisan resistance (for example, Bail et al. 2018; Hart and Nisbet 2012; Lord et al. 1979; Nyhan and Reifler 2010; Taber and Lodge 2006). These studies differ in many aspects – such as research design, issue area, information content, outcome variable, study duration, and setting – making it impossible to explain the discrepancies with a single answer. To address this issue, I derive multiple hypotheses specifying several conditions under which partisans may or may not revise their opinions toward new information from the motivated reasoning literature. By examining these hypotheses together, I explore why some studies, but not others, found partisan resistance. I test them in two survey experiments that measure the persuasive effects of information about the ACA.¹

A widely accepted account in the literature is that partisan resistance is most likely when politically contested issues – such as the ACA – are involved, as people are likely to be emotionally invested in their views on such issues (for example, see Bolsen and Palm 2019, 9; Leeper and Slothuus 2014, 143; Druckman 2012, 204). Therefore:

H1a: Pro-ACA information will increase support for the ACA among Democrats, whereas it will either decrease or fail to increase support for the ACA among Republicans.

H1b: Anti-ACA information will decrease support for the ACA among Republicans, whereas it will either increase or fail to decrease support for the ACA among Democrats.

H1c: Both information treatments will increase the partisan gap in ACA support.

However, many previous studies demonstrating people's openness to opposing information also focused on controversial topics (for example, Coppock 2023; Tappin et al. 2023; Wood and Porter 2019), suggesting that the mere presence of a polarizing issue may not be enough to trigger partisan resistance. What other factors could be involved?

One possibility is that people resist uncongenial information *only* to the extent that they can construct a counter-argument against it, thereby maintaining 'the illusion of objectivity' (Pyszczynski and Greenberg 1987) – a caveat noted by motivated reasoning scholars (see Kunda 1990, 482–483; Redlawsk et al. 2010; Taber and Lodge 2006, 757). That is, partisans may be able to

¹See Appendix C.3 for background information about the ACA as a political issue.

selectively reject uncongenial arguments (while uncritically accepting congenial arguments) at will when the presented evidence is relatively weak. However, ‘even those committed to their positions can be persuaded by strong and credible counter-evidence’ (Lodge and Taber 2000, 188). According to this explanation, variations in the strength of the evidence presented to subjects across different studies may have contributed to the conflicting results. I systematically test this possibility using an experiment that varies evidence strength. Specifically, I consider the following hypothesis:

H2: The patterns described in H1 will be more pronounced when the information about the ACA provides weak evidence for its position.

Another possibility is that, even when partisans update their *factual beliefs* to reflect new information, they may still resist changing their *attitudes* toward a policy or a public official responsible for the policy (for example, see Nyhan et al. 2020; Porter et al. 2019).² This resistance may occur through counter-arguing, where partisans search their memories for other beliefs to justify and maintain their attitudes (for example, see Kunda 1990, 483) or through biased attribution of blame or credit (for example, Bisgaard 2019; Graham and Singh 2024). For example, an Obama supporter exposed to information that healthcare costs increased rapidly under Obamacare may accept the fact but still maintain or even strengthen their positive attitudes toward the ACA and Obama by thinking, ‘Obamacare expanded healthcare access despite the cost increase,’ or ‘Healthcare costs increased because Obamacare opponents prevented Obama from including the public option, so it’s not his fault’. Therefore, I propose:

H3: The patterns described in H1 will be more pronounced for attitudes toward the ACA and Obama than for beliefs about the ACA.

Furthermore, the broader literature on motivated reasoning (beyond the political domain) demonstrates that people tend to selectively forget ego-threatening information – a tendency that helps people maintain self-serving views of themselves, even in the face of negative feedback (Bénabou and Tirole 2002; Mischel et al. 1976; Zimmermann 2020). Applied to political communication, this suggests that even when partisans initially appear to be persuaded by both pro-attitudinal and counter-attitudinal information, these effects may decay asymmetrically over time, ultimately resulting in persistently polarized perceptions and attitudes on public issues (for example, Bartels 2002; Jerit and Barabas 2012). In line with this explanation, Bartels (2014) found that while exposure to an Obama campaign ad immediately increased both Obama and Romney supporters’ intention to vote for Obama, only the effects on the latter group eroded over time. Therefore, I test:

H4: The effects of congenial information will persist, whereas the effects of uncongenial information will decay over time.

Finally, contextual variations may prompt the same citizen to adopt different reasoning strategies. Previous studies have shown that providing incentives for correct answers or requiring justification – contexts that can prime accuracy goals – reduces people’s tendencies to selectively scrutinize information (Christensen and Moynihan 2024; Khanna and Sood 2018) or to engage in expressive responding in surveys (Bullock et al. 2015; Prior et al. 2015). These results (see also Bolsen et al. 2014; De Dreu and Van Knippenberg 2005; Green et al. 2000) imply that people may or may not engage in biased reasoning depending on various contextual factors, rather than being inherently open-minded or closed-minded.

²An attitude is defined as the weighted average of multiple evaluative beliefs (Fishbein, 1963).

In this study, I focus particularly on whether people become more likely to engage in partisan resistance when new information is presented in a context that triggers strong feelings of anger or hostility toward the opposing party, thereby activating partisan motives. In line with this possibility, several studies have found that priming anger, expectation of conflict, partisan loyalty, or disdain for the out-party (compared to the in-party) can lead people to rely more heavily on party cues instead of policy details (Bolsen et al. 2014; Lavine et al. 2012), express stronger support for their own party's policy (Klar 2014), engage in biased information search (MacKuen et al. 2010), exhibit stronger biases in argument evaluation tasks, and become more polarized (Groenendyk and Krupnikov 2021).

Indeed, some of the most influential findings of partisan resistance are consistent with this explanation. For example, Taber and Lodge (2006, 756) noted that they deliberately chose contentious political arguments that could 'generate strong affective responses' because 'cold' arguments are unlikely to arouse partisan motivation. This, conversely, suggests that partisans may have been more open to changing their minds had the stimuli been less evocative. Additionally, in Bail et al.'s (2018) experiment showing polarization among those exposed to opposing views on Twitter, participants may have been exposed to not just opposing information but also to the toxic political discourse common on social media (Kim et al. 2021). This leaves open the possibility that, absent the incivility, exposure to opposing information could have had the opposite (that is, depolarizing) effect.

However, it is crucial to note that prior studies have not conclusively determined whether partisan resistance depends on hostile contexts – at least not in the precise sense that the *persuasive effects* of congenial versus uncongenial information vary across hostile versus non-hostile contexts. To determine this, one should randomize both (1) exposure to valenced information and (2) the hostility of the context in which the information is presented. This approach will allow one to examine whether those exposed to new information develop more polarized opinions compared to those who receive no information *and* how a hostile context amplifies these differences. Most of the political persuasion studies discussed above (for example, Bail et al. 2018; Guess and Coppock 2020; Nyhan and Reifler 2010) randomized exposure to information but did not manipulate the informational context. Conversely, other studies that varied contexts randomly (for example, Groenendyk and Krupnikov 2021; Klar 2014; Lavine et al. 2012) did not measure the persuasive effects of information using a randomized controlled trial design. To complement these studies, I use a design that first randomly manipulates partisan sentiments and then randomizes exposure to information treatments, which use either civil or uncivil language.

This investigation directly addresses one of the central theoretical arguments of partisan-motivated reasoning: that biased reactions to information are driven by people's affective involvement with the political world (Lodge and Taber 2000; see also Bolsen and Palm 2019; Leeper 2014). Furthermore, this approach tackles a common limitation in much of experimental research on political persuasion: the sterile and artificial experimental settings, which are largely divorced from the contemporary political information environment, fraught with toxicity, hostility, outrage, stereotyping, trolling, mockery, and even outright hate speech (Bail 2022; Berry and Sobieraj 2013; Jamieson and Cappella 2008; Kim et al. 2021; Levendusky 2013). The result of encountering the same attitudinally inconsistent information could be markedly different depending on whether it was delivered in such contentious contexts or not. By examining how people respond to new information 'in a vacuum' as well as in conditions that mimic the messier aspects of the real world, this research seeks to measure both the potential effects that political information could have in theory and the actual effects it is likely to have in practice. In short, I examine:

H5: The patterns described in H1 will be more pronounced when partisans are induced to feel hostile toward the opposing party or when the persuasive message incorporates uncivil language.

I test the hypotheses in two experiments presented in the following sections. Study 1 considers H1 to H4. Study 2 considers H5.

Study 1

In Study 1, I assigned respondents to read a message about the ACA or a placebo message to estimate the treatment effects on (1) attitudes toward the ACA and (2) beliefs about the ACA's impact on health costs. The treatment effects are defined as the mean differences between the control and treatment groups.

Procedure

Study 1 was conducted in four waves. The baseline survey measured pre-treatment covariates, including the baseline values of the dependent variables. Survey responses were gathered from 2,029 participants recruited via Amazon's Mechanical Turk (MTurk) in May 2016. After a week, 1,800 of the respondents from the baseline survey were invited to participate in the main experiment. Those who failed a screener question and non-US citizens were not invited; 1,514 of the invited respondents completed the second wave.

The experiment was administered at Wave 2. Before the treatment, participants responded to an attention screener question (Berinsky et al. 2014).³ Participants were then provided with the message to which they were randomly assigned and asked to read it carefully. Next, respondents answered questions measuring their post-treatment opinions about the ACA.

Those who took part in the main experiment were invited to complete two follow-up surveys. Wave 3 ($n = 899$) was conducted about 80 days after Wave 2 (between late July and early August). Wave 4 ($n = 850$) was administered about 160 days after Wave 2 (late October to early November).

Sample

Participants were 18 per cent ages 18–25, 62 per cent ages 25–44, 51 per cent female, and 76 per cent white. About half (53 per cent) held a four-year college degree. Politically, 60 per cent were Democrats and 27 per cent were Republicans (including leaners). Appendix A.4.4 compares the Study 1 sample with the benchmark 2016 American National Election Studies (ANES) sample and provides additional analyses suggesting that the potential impacts of the sample characteristics on treatment effect estimates may have been modest (see also Coppock 2019).

Experimental Conditions

Participants were randomly assigned to one of five conditions: placebo, strong con, weak con, weak pro, and strong pro. Those assigned to the non-placebo conditions received an article about the ACA, with its stance either in support or opposition to the ACA and the strength of evidence either strong or weak.

Message Contents

When crafting the stimuli, I ensured that the focus was on a sub-topic that *both* Democrats and Republicans would consider important. Based on prior research (Druckman et al. 2012), I expected that people would care about health costs regardless of their partisanship. I test this

³Those who failed to answer this question correctly were not excluded from the main analyses in both Studies 1 and 2. The results remained robust when they were excluded (see Tables S1 and S18).

assumption in Fig. S3 in Appendix A.4.1, which shows that partisans were in near-complete agreement about the importance of health costs as an evaluative criterion.⁴

The two anti-ACA messages claimed that the ACA is driving up health cost growth, and the two pro-ACA messages claimed the opposite. The anti-ACA messages then presented four pieces of evidence indicating the ACA was accelerating health cost growth, which were manipulated to be either strong or weak depending on the assigned condition. Likewise, the pro-ACA messages presented four facts suggesting that the ACA was slowing health cost growth. To realistically vary the strength of the evidence, I based the ‘strong’ versions on real research findings that described (a) noticeable differences (b) in relevant healthcare data (c) at the national level. I then created weaker counterparts by manipulating some of these dimensions – (a) smaller differences, (b) arguably irrelevant statistics or (c) local-level evidence – as permitted by the available information. Appendix A.1 provides the full texts of the treatment messages.

Measurement of Outcome Variables

The primary outcome variables are participants’ attitudes toward the ACA and their beliefs about the ACA’s effect on health costs. Following Taber and Lodge (2006, 757), I used multiple items to measure these variables to avoid the possibility that the upper and lower bounds of a single-item scale prevent the detection of polarization.

For the measurement of attitude toward the ACA, participants were asked whether they ‘support or oppose the health care law sometimes referred to as Obamacare’ followed by four more items also probing their attitudinal support for the law (Cronbach’s $\alpha = 0.96$). The belief that the ACA saves health costs was measured by asking participants how the law is changing the health care system ‘in terms of healthcare costs’, followed by two additional items (Cronbach’s $\alpha = 0.77$). Study 1 also considers the downstream effects on people’s evaluations of Obama on (1) his handling of healthcare and (2) his job as President. These variables were rescaled to vary between 0 and 1, where a high value indicates a favourable opinion on the ACA or Obama. More information about the measurement of the outcomes and other covariates in Study 1 is provided in Appendix A.2.

Results

Figure 1 presents the main results of Study 1.⁵ The upper panels plot the OLS estimates of the marginal effects of each information treatment – that is, the mean difference from the control group among each partisan group (see Appendix A.3.1 for more details on the regression). These panels show that people updated their overall attitudes toward the ACA (Panel A) and their beliefs about its effect on cost growth (Panel B) in accordance with the assigned information, regardless of whether this information aligned or conflicted with their partisan identities. Among the sixteen estimates, all eight estimates of the effects of uncongenial information (for example, anti-ACA information presented to Democrats) were statistically significant, while five of the eight estimates of the congenial information effects were significant. Collectively, these results offer no support for H1a and H1b, which predicted that uncongenial information would either have no effect or cause a ‘backfire’ effect (for example, a positive effect of anti-ACA information).

Comparing the circles (Democrats) and triangles (Republicans) in Panels (A) and (B), we see that each information treatment led Democrats and Republicans to update their viewpoints in similar ways. The lower panels of Fig. 1 show the estimates of partisan differences in the treatment

⁴These results are from the baseline Wave 1 survey.

⁵Appendices A.4.5 and B.5.6 report randomization checks that show the experimental groups in Studies 1 and 2 were well-balanced on pre-treatment covariates. Appendices A.4.6 and B.5.7 present unadjusted outcome means and difference-in-means estimates for both studies. These are similar to, but less precise than, the covariate-adjusted estimates reported here.

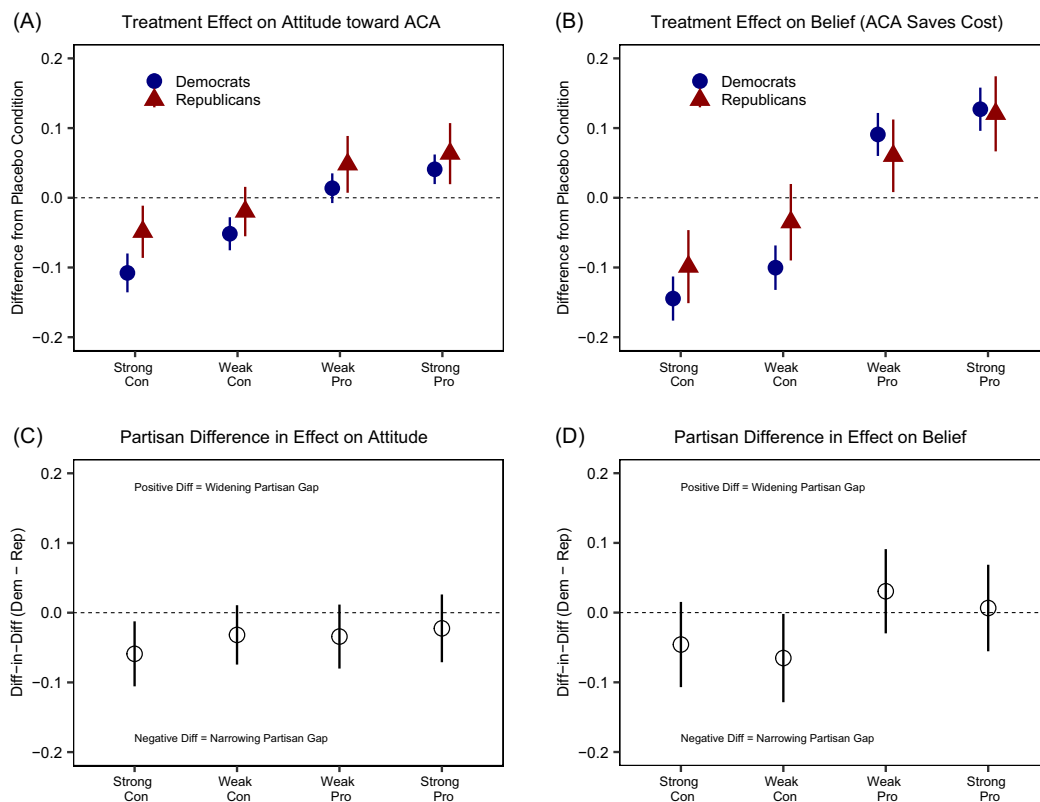


Figure 1. Effects of New Information on ACA Attitude and Belief (Study 1).

Note: This figure displays OLS estimates with 95 per cent CIs. Each model controlled for pretreatment values of the outcome variables and partisanship for statistical precision. In Appendix A.3, Table S1 reports the complete regression estimates with additional robustness checks, and Table S2 reports the marginal effect estimates in tabular format. See Table S10 for distributions of partisanship and demographics, broken down by condition.

effects (Democrats minus Republicans). These estimates are constructed so that a positive number indicates that a treatment increased the partisan gap in opinions (divergence), whereas a negative number signals the opposite effect (convergence). Among the eight estimates of partisan differences, seven were negative, with two reaching statistical significance ($p < 0.05$). In both cases of significant heterogeneous effects, anti-ACA information exerted a stronger influence on *Democrats*, thereby reducing the partisan gap in attitude by 6 percentage points and the belief gap by 7 points. The average of the four estimates in Panel (C) was -0.04 ($p < 0.05$), suggesting that the four information treatments, on average, decreased partisan disagreement by about 4 percentage points. The average of the estimates in Panel (D) was also negative (-0.02), though statistically insignificant. In short, the hypothesis that exposure to new information treatments will increase partisan disagreement (H1c) was rejected.

The figure also demonstrates the nuanced role of evidence strength in opinion updating. Specifically, while evidence strength affected the overall magnitude of the treatment effects, leading to larger opinion shifts with stronger evidence, it did not increase or decrease partisan biases. The partisan differences in the treatment effects in Panels (C) and (D) were roughly equivalent between the strong and weak versions of the information, with none of the comparisons reaching statistical significance. These results contradict H2, which posited an increase in biased resistance when people are presented with weaker evidence.

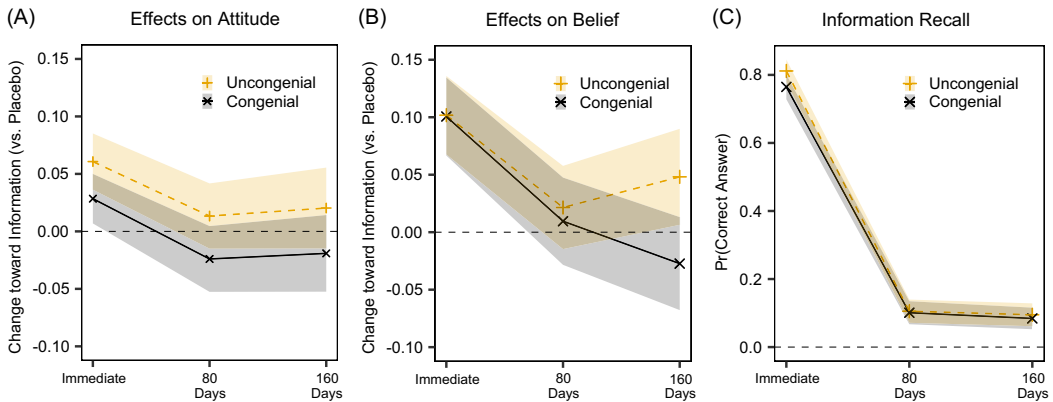


Figure 2. Decay and Persistence of Congenial vs. Uncongenial Information effects.

Note: This figure displays OLS estimates with 95 percent CIs. See Appendix A.3.3 for more details. See Table S10 for distributions of partisanship and demographics, broken down by condition.

Furthermore, the figure shows that both beliefs about a specific aspect (cost) of the ACA and overall attitudes toward the law changed in parallel or in slightly convergent ways. This is inconsistent with the expectation that new information would lead to increased polarization in attitudes, even if it does not result in polarization of factual beliefs (H3). I also examined the downstream effects on people's approval of Obama's handling of healthcare and his overall job performance to test the possibility that the treatments led partisans to disagree more about Obama (see Fig. A.4.1 in Appendix A). However, the treatment effects on approval of Obama's performance in healthcare aligned closely with the findings in Fig. 1, while *neither* congenial nor uncongenial information significantly affected people's overall evaluations of Obama. In sum, there was no evidence to suggest that new information widened partisan divides in *attitudes*, which disconfirms H3.

So far, I have found little evidence of partisan bias in opinion changes measured *immediately* after the information treatments. But could partisan bias have revealed itself in the longer term (Bartels, 2014)? In Fig. 2, I examine whether the effects of congenial information were more persistent. Panels (A) and (B) plot the overall effectiveness of congenial versus uncongenial information treatments, measured at different points in time – immediately after the treatment, roughly 80 days later, and 160 days later.⁶ As these panels illustrate, the estimates of the long-term effects are much smaller than the estimates of the immediate effects and are mostly insignificant. More importantly, I find no evidence of asymmetrical decay. If anything, the incongruent information effects appear to have lasted longer. Thus, partisans appear to forget new information that is favourable to their party just as easily as information unfavourable to their side – a point also reflected in Panel (C), which assesses participants' recall of the statistical facts included in the information treatments (for example, 'Premiums in the ACA Marketplaces rose by 10.1 per cent on average nationwide in 2016'). In short, I find no support for H4.

In summary, Study 1 rejected every hypothesis that posited partisans accept and retain information in biased ways. There are at least two reasonable reservations about these results. The first is whether the findings could generalize to other potentially more polarizing and consequential issues. The second concern is that the time gap between the main experiment and the follow-ups may have been too long, making the test of H4 too conservative. To address these concerns, Appendix D presents an additional original experiment testing H1, H2, and H4 by measuring the immediate and long-term effects of pro-Democratic information about the major

⁶For comparability, the immediate effects in Fig. 2 were drawn from those who responded to Wave 3 (80 days) as well.

parties' ability to manage the economy on partisans' preferences for the Democratic Party on this issue. The experiment shows that exposure to new information led partisans to change their views in parallel or convergent ways regardless of evidence strength, consistent with Study 1's findings on H1 and H2.⁷ Furthermore, even though the economy experiment conducted its follow-up survey 10 days after the treatment – as opposed to 80 to 160 days in Study 1 – it still did not find support for H4, the hypothesis that congenial information effects would be more persistent. This confirms the findings of Study 1 as well.⁸

Study 2

I designed the second study to examine whether the key finding of Study 1 holds even under theoretical conditions that are particularly favourable to instigating 'hot cognition' (Taber and Lodge 2006), or whether people become more closed-minded under such contexts. I took two approaches to create an adversarial environment. First, I used ambivalent versus univalent partisanship priming techniques (Bolsen et al. 2014; Klar 2014; Lavine et al. 2012) to experimentally induce variations in affective polarization before treating participants with new information. Second, I varied the tone of the treatment messages to be either civil or uncivil. I explored the extent to which partisans exhibited biased responses to congruent versus incongruent information when situated in a contentious environment.

Procedure

A total of 4,506 participants, recruited via MTurk, completed the survey in August 2018.⁹ Study 2 was conducted in one wave. Participants first answered questions measuring various pre-treatment covariates, including prior ACA attitudes, partisanship, and demographic traits, followed by a screener question (Berinsky et al. 2014). They then engaged in one of two partisanship priming tasks that were designed to randomly vary feelings towards the major parties.¹⁰ Subsequently, participants read either a placebo message or one of four messages related to the ACA. Next, they answered survey questions measuring the outcome variables. Finally, they gave feeling thermometer ratings for the parties.

Sample

Participants were 14 per cent ages 18–25, 64 per cent ages 25–44, 52 per cent female, and 74 per cent white. About half (56 per cent) held a four-year college degree. Politically, 56 per cent were Democrats and 33 per cent were Republicans. See Appendix B.5.5 for more information about the sample characteristics and their potential impact on generalizability.

Experimental Conditions

Participants were randomly assigned to one of ten conditions in a 2 (ambivalent versus univalent partisanship prime conditions) by 5 (placebo, civil con, uncivil con, civil pro, and uncivil pro message conditions) factorial design.

⁷See Appendix D.3.1 for further discussion on generalizability across issues.

⁸Previous research has also shown limited evidence of asymmetrical decay over a week to a month (Coppock, 2023, 118, Coppock et al. 2023, Nyhan et al. 2020). However, H4 warrants further investigation, as there may be a time window in which only congenial information is remembered and remains effective before completely decaying, as shown in Study 1. See Appendix D.3.2 for further discussion.

⁹See Appendix C.2 for a discussion on the potential impact of the time gap between the studies on the conclusions.

¹⁰Pure independents skipped this task.

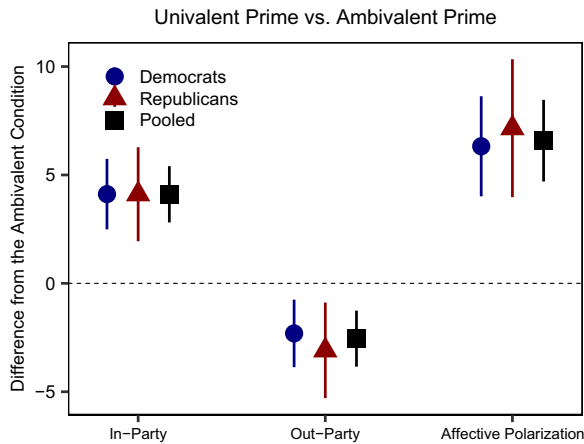


Figure 3. Effect of Priming Treatments on Affective Polarization.

Note: This figure displays OLS estimates with 95 per cent CIs. Table S25 in Appendix B presents the regression models, from which these marginal estimates are calculated.

Partisan Affect Priming Tasks

Following previous studies (Klar 2014; Lavine et al. 2012), I randomly induced partisans to feel either ambivalent or univalent about the major parties by asking those assigned to the ambivalent partisanship condition to write reasons why they are dissatisfied with their party, and those assigned to the univalent condition to write why they prefer their own party and dislike the other party.¹¹ Appendix B.2 provides the priming instrument.

I conducted a manipulation check on the priming tasks using the standard feeling thermometer ratings for in-party vs. out-party (each ranging from 0 to 100). Figure 3 shows that the univalent partisanship priming induced a 6- to 7-degree increase in affective polarization relative to the ambivalent priming. This effect is roughly half of the increase in affective polarization observed over three decades in the US (Iyengar et al. 2012).

Message Contents

The civil versions of the messages were modified from the strong versions in Study 1. The uncivil versions were identical to the civil versions in substance but included several insulting comments aimed at the other side. For example, the uncivil version of the pro-ACA message called the Trump administration ‘irresponsible and despicable’ and Republicans in Washington ‘idiotic’, among others. Figure S10 in Appendix B.5.4 reports manipulation checks on the civility treatment, showing very large differences in perceived civility between the civil and uncivil versions of the messages. Participants in the placebo (that is, control) condition read a message about the debt crisis in Greece. The full texts are available in Appendix B.1.

Measurement of Outcome Variables

As with Study 1, multi-item batteries were used to measure the two primary outcome variables: overall attitude toward the ACA (4 items; Cronbach’s $\alpha = 0.93$) and the belief that the ACA controls health costs (4 items; Cronbach’s $\alpha = 0.88$). The survey items were similar to those used in Study 1 (see Appendix B.3).

¹¹I did not include a third group, who would not participate in either of the priming tasks, as doing so would increase the number of cells to 15 and consequently decrease statistical power. Study 1 had already examined partisans’ reactions to information in such a condition. The primary focus of Study 2 was to maximize the *variation* in partisan sentiments.

Results

The objective of Study 2 is to analyze how partisans' reactions to new information are influenced by contextual variation in hostility. The study measures the differences in updating patterns when partisans are: (1) primed to feel politically ambivalent and then exposed to civil discourse; (2) primed to feel ambivalent and exposed to uncivil discourse; (3) primed to feel politically univalent and then exposed to civil discourse; or (4) primed to feel univalent and exposed to uncivil discourse. The baseline for comparison is the first context (ambivalent and civil). This study examines whether other contexts, referred to as 'contentious' or 'hostile' contexts, facilitate divergent opinion updating, where Democrats and Republicans *increase* their disagreement after learning new facts.

Figure 4 illustrates how participants assigned to different conditions updated their attitudes toward the ACA, based on a regression model identifying treatment effects on each partisan group across different contexts (for more details on the regression, see Appendix B.4). Panel (A) presents the OLS estimates of the marginal effects. The shaded areas represent the non-hostile baseline (that is, civil and ambivalent) and the white areas represent the hostile contexts. Blocks

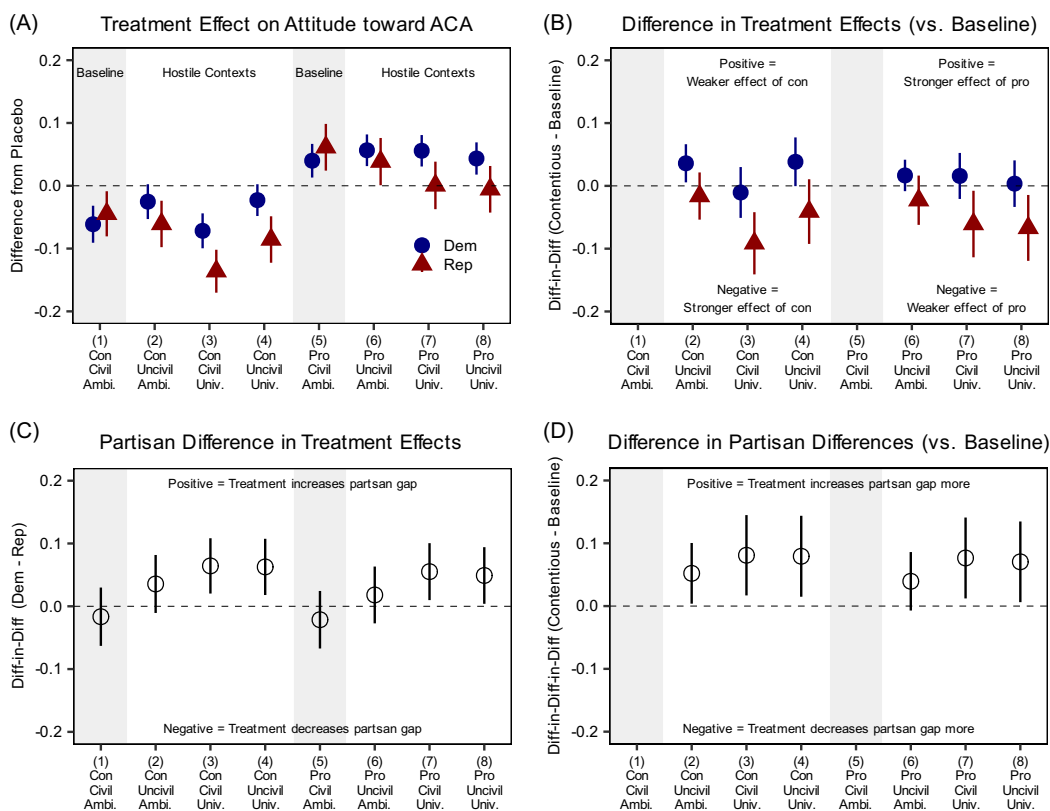


Figure 4. Effects of New Information on Attitude toward the ACA across Contexts (Study 2).

Note: This figure displays OLS estimates with 95 per cent CIs. Panel (A) plots the marginal effects of the negative/positive information on Democrats and Republicans across different contexts. Panel (B) depicts the differences in the treatment effects between the baseline and hostile contexts. Panel (C) plots the difference in the marginal effects between Democrats and Republicans, while Panel (D) presents the differences in the heterogeneous effects between the baseline and hostile contexts. These estimates are derived from a regression model which controls for pre-treatment attitude and partisanship for precision. In Appendix B.4, Table S18 reports the complete regression estimates with additional robustness checks, and Tables S19 to S22 report the estimates plotted in this figure in tabular format. See Table S27 for distributions of partisanship and demographics, broken down by condition.

(1)–(4) display the effects of the anti-ACA information. Block (1) presents the marginal mean differences between the *ambivalent-civil-con* group and the *ambivalent-control* group, and Block (2) the differences between the *ambivalent-uncivil-con* group and the *ambivalent-control* group. Similarly, Blocks (3) and (4) show the differences between the *univalent-civil-con* group and the *univalent-control* group, and between the *univalent-uncivil-con* group and the *univalent-control* group. Blocks (5)–(8), which show the effects of the pro-ACA information, can be interpreted in an analogous way.¹²

Blocks (1) and (5) – in shaded areas – show that, in the non-hostile context, the negative information about the ACA decreased Democrats' favourability towards the ACA by 6 percentage points and Republicans' by 4 percentage points (Block 1) – while the positive information improved Democrats' views on the ACA by 4 points and the Republicans' by 6 points (Block 5). These updating patterns are similar to those found in Study 1, where partisans updated in slightly convergent or parallel ways. But as shown in the white areas of the panel, partisans in the hostile contexts reacted to the (substantively) same information in quite different ways. Relative to the baseline, congenial information effects (for example, the effect of the anti-ACA information for Republicans) generally strengthened, whereas uncongenial information effects weakened. Take the triangle in Block (5) versus the triangle in Block (7), for instance. Republicans in the univalent condition (Block 7) – that is, those primed to think about what they disliked about the Democratic Party (and what they liked about their own party) before the information treatment – remained completely unaffected by the same information that had a 6-percentage-point effect on their counterparts in the ambivalent priming condition (Block 5).

Panel (B) of Fig. 4 presents the differences in the effects of the information treatments *across* the non-hostile versus the hostile contexts, which can be interpreted as the differences between the points in the grey and white areas in Panel (A). If an adversarial context amplifies partisan bias, we would expect these estimates to be consistently positive for Democrats – indicating a weaker effect of the anti-ACA information or a stronger effect of the pro-ACA information in a hostile context – and consistently negative for Republicans. As shown, all but one estimate is in the expected direction, with four of them reaching statistical significance at $p < 0.05$ (and one more just marginally missing the cutoff with $p = 0.053$).

Panel (C) reports the estimated partisan differences in the treatment effects – that is, the differences between the circles and triangles in Panel (A). As with Study 1, these differences are calculated so that a positive number indicates an increased disagreement among partisans. The two estimates corresponding to the non-hostile environment (Blocks 1 and 5) were small and negative (–0.02). By contrast, the six estimates corresponding to the hostile contexts (Blocks 2–4 and 6–8) were all positive, four of which were statistically significant. The average of these six estimates was 0.05 ($SE = 0.01$, $p < 0.005$; see Table S21), suggesting that when the information treatments were delivered in a hostile context, partisans became more polarized by an additional 5 percentage points, on average.

Panel (D) contrasts the estimates plotted in Panel (C) across the hostile and non-hostile contexts. The points represent the differences between the circles in the grey and white areas in Panel (C) or, equivalently, the differences between the circles and triangles in Panel (B). As depicted, all estimates were positive, ranging from 0.04 to 0.08, with an average of 0.07 ($SE = 0.02$, $p < 0.005$; see Table S22). This indicates that, on average, the contentious contexts caused an additional 7-percentage-point divergence between Democrats and Republicans upon exposure to new information, *relative to* their counterparts who received the same information but in the non-adversarial context.

¹²Because I used a fully crossed design, there are two control groups: the *ambivalent-control* and *univalent-control* groups. Note that in Blocks (3), (4), (7), and (8), the treated groups are compared against the latter (*univalent-control*) group. Table S24 compares the *ambivalent-control* and *univalent-control* groups, and shows no evidence that partisanship priming in itself – without exposure to new information – altered people's (expressed) views on the ACA.

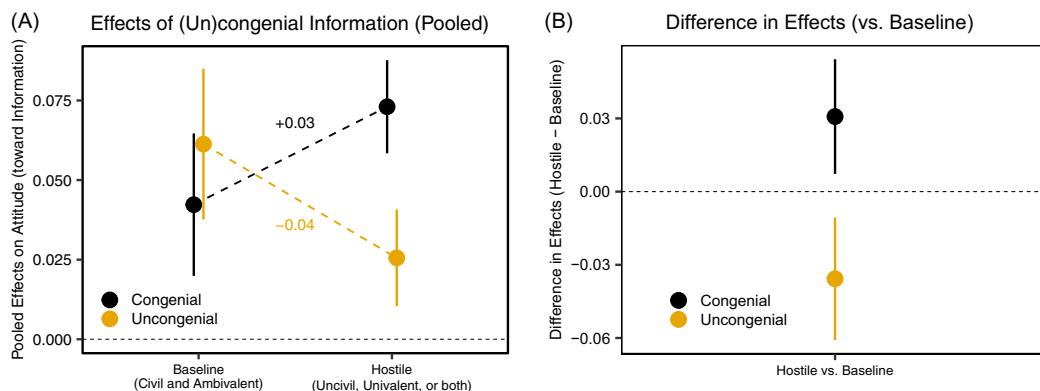


Figure 5. Changes in the Persuasive Effects of Congenial vs. Uncongenial Information across Contexts.

Note: This figure displays OLS estimates with 95 per cent CIs. Panel (A) plots the persuasive effects of congenial information vs. uncongenial information in non-hostile vs. hostile contexts. Panel (B) plots the differences in the persuasive effects across the contexts. These estimates, calculated by pooling the estimates in Fig. 4, are provided in tabular format in Table S23. See Appendix B.4.3 for more information. See S27 for distributions of partisanship and demographics, broken down by condition.

Given that the individual estimates are somewhat noisy, it could help to consider the overall decrease in the effects of *uncongenial* information and the increase in the *congenial* information effects across contexts. To do so, I pooled the estimates from Panels (A) and (B) of Fig. 4 by congeniality and context. In this case, a positive number signifies successful persuasion – that is, an attitudinal shift aligning with the presented information. The pooled estimates are presented in Fig. 5. As shown in Panel (A), congenial information led, on average, to a 4-percentage-point shift in partisans' attitudes toward the ACA in the non-hostile context. This effect size increased to 7 percentage points in the hostile contexts. Conversely, the overall persuasive impact of uncongenial information decreased from 6 to 3 percentage points across these contexts. Panel (B) shows that both of these changes are statistically significant.

Figures S8 and S9 in Appendix B display the estimated treatment effects on the belief that the ACA controls health cost growth, in the same formats as Figs. 4 and 5. I omit a detailed discussion of the estimates, aside from noting that belief updating patterns generally resemble attitude updating patterns.¹³ In conclusion, Study 2 furnishes strong evidence supporting H5, which posited that contextual changes can affect partisans' responses to new information – specifically, that a hostile environment can instigate partisan resistance.

What do the opinion updating patterns observed in Studies 1 and 2 say about citizens' ability to use information rationally? In Appendix C.1, I compare the observed treatment effects to a benchmark model of rational opinion updating, simulating how *unbiased* Bayesians would update their opinions when faced with incongruent information. The findings indicate that, in the non-contentious environments, partisans updated their opinions in a manner that closely resembles the benchmark. However, in the contentious environments, the observed effects deviated markedly from the benchmark.

¹³Extending H3, which posited different patterns for attitudes and beliefs, one might expect contextual variation in partisan resistance would be more pronounced for attitudes – presuming that a hostile context instigates motivated reasoning mainly by altering people's tendencies to incorporate factual beliefs into attitudes (Bisgaard, 2019), rather than their tendencies to update factual beliefs in the first place. A side-by-side comparison between the results on attitudes and beliefs suggests that the contextual differences in the effects of congenial vs. uncongenial information were mostly similar between the two outcome variables, thus showing no support for this possibility (see Appendix B.5.8). I thank an anonymous reviewer for suggesting this analysis.

Discussion

This article presents two survey experiments designed to explore several conditions under which people resist new information and isolate the role of partisan motives in triggering such resistance. Across the studies, I found that people tend to update their opinions toward new information about the ACA, even when it goes against their party's position. However, people became much more likely to selectively reject uncongenial information when they were randomly assigned to feel hostile toward the opposing side or when they were exposed to uncivil discourse.

These results are important for three reasons. First, they identify an important condition under which partisan resistance to new information becomes more likely. The findings show that it is not necessarily contradictory political *information per se* but, rather, the hostile and contentious political *contexts* in which such information is delivered that lead people to resist political persuasion. Distinguishing between uncongenial information versus uncongenial context as a trigger of resistance can help reconcile at least some of the conflicting observations in the literature on how people react to uncongenial information. Specifically, the stimuli used in some studies documenting resistance may have manipulated both the exposure to information and partisan animosity (for example, Bail et al. 2018; Taber et al. 2009; Taber and Lodge 2006), suggesting the possibility that people may accept contrary information when the latter is not manipulated (for example, Guess and Coppock 2020; Wood and Porter 2019).

On the flip side, this also means that, even though people are persuadable in theory, they may routinely engage in motivated reasoning in practice because a significant part of today's political information environment is filled with partisan hostility and toxicity (for example, Berry and Sobieraj 2013; Kim et al. 2021). The results thus indicate the need to distinguish between the extent to which partisans *can* change their minds under favourable conditions that lack the detrimental features of the current information ecosystem and the extent to which they *do* change their minds in their everyday lives, where such features may be prevalent. This point has an important methodological implication too: typical survey experiments that do not manipulate contexts may capture the former, potentially *overstating* the persuasive effects that counter-information can have on the development of public opinion in the real world.

Second, these results identify partisan motives as the cause of biased resistance to persuasion. As highlighted by several authors, detecting motivated reasoning can be challenging, as those without partisan motives might also selectively dismiss uncongenial information if it seems irrelevant to their values or untrustworthy based on their existing knowledge and beliefs (Druckman and McGrath 2019; Gerber and Green 1999; Little 2021; Tappin et al. 2020). This study mitigates these confounding factors by showing that partisans comparable in all other ways either accept or reject uncongenial information depending on whether affective partisanship is primed.

Finally, the findings suggest that the quality of voter decision-making depends critically on the quality of political discourse at the elite level. Political leaders' inability (or unwillingness) to create a constructive political information environment (Berry and Sobieraj 2013; Jamieson and Cappella 2008; Levendusky 2013) can make otherwise 'rational' citizens incapable of changing their minds in light of arguments and evidence. In this regard, partisan-motivated reasoning may be a self-fulfilling prophecy: the prevailing assumption that voters won't respond to evidence-based persuasion (Remnick 2017; Wehner 2020) leads political elites to rely on acrimonious rhetoric designed to rally loyalists, and that, in turn, leads voters to engage in motivated reasoning even further.

Before concluding, I note several limitations of this study. First, this research used non-probability online samples recruited through MTurk. While this choice was necessary for conducting well-powered experiments, including a panel study, the extent to which these findings can be generalized to the broader American public remains uncertain (but see Appendices A.4.4

and B.5.5 as well as Coppock, 2019). Second, the experiments focused only on the ACA (and the economy) as the target issues. Future research should explore how people respond to new information about other issues, such as immigration, climate change, vaccination, or election fraud charges. It would be especially interesting to examine to what degree informational environments affect people's willingness to change their minds about such issues. Third, previous research suggests alternative methods to manipulate affective partisanship or directional motives (for example, Broockman et al. 2023; Groenendyk and Krupnikov 2021). It would be beneficial to examine whether the findings hold using these alternative approaches. Fourth, while Study 1 and the economy experiment employed a panel design, Study 2 did not. Consequently, I was unable to examine whether the increased partisan gap found in Study 2 persisted over time. A potentially fruitful avenue for future research is to extend H4, the differential decay hypothesis, and explore whether the effects of congenial information delivered in hostile contexts are particularly persistent.

Despite these limitations, this study represents an important step toward sharpening our understanding of whether and when evidence can change partisan minds. The findings suggest that partisan-motivated reasoning may indeed be a 'pervasive force' in contemporary American politics (Bolsen and Palm 2019), given the current climate of constant partisan conflicts and toxic political discourse (for example, see Berry and Sobieraj 2013; Kim et al. 2021). But the findings also indicate that the electorate may potentially 'behave rationally and responsibly given the clarity of the alternatives presented to it and the character of the information available to it' (Key 1966, 7). Curtailing motivated reasoning begins with better political leadership.

Supplementary material. The supplementary material for this article can be found at: <https://doi.org/10.1017/S0007123424000899>

Data availability statement. Replication data for this article can be found in Harvard Dataverse at: <https://doi.org/10.7910/DVN/AWJOYM>.

Acknowledgements. I am deeply grateful to Michael Delli Carpini and Brendan Nyhan for reading multiple drafts of this paper and providing valuable feedback each time. I also thank John Bullock, Joe Cappella, John Carey, Jamie Druckman, D.J. Flynn, Shang E. Ha, Seth Hill, Sun-ha Hong, Yusaku Horiuchi, Devra Moehler, Diana Mutz, Yunjoung Park, June Woong Rhee, four anonymous reviewers at BJPOLS and participants in seminars at the International Communication Association, the Midwestern Political Science Association, the Korean Association of Party Studies, the Program in Quantitative Social Science at Dartmouth College, the Institute for Communication Research at Seoul National University, and the Political Science Department at Sogang University for their helpful comments.

Financial support. This research was supported by the Annenberg School for Communication at the University of Pennsylvania and the Program in Quantitative Social Science at Dartmouth College.

Competing interests. The author declares none.

Ethical standards. The research was conducted in accordance with the protocols approved by the University of Pennsylvania and Dartmouth College.

References

- Arceneaux K and Vander Wielen RJ (2017) *Taming Intuition: How Reflection Minimizes Partisan Reasoning and Promotes Democratic Accountability*. Cambridge, UK: Cambridge University Press.
- Bail C (2022) *Breaking the Social Media Prism: How to Make Our Platforms Less Polarizing*. Princeton, NJ: Princeton University Press.
- Bail CA, Argyle LP, Brown TW, Bumpus JP, Chen H, Hunzaker MF, Lee J, Mann M, Merhout F, and Volfovsky A (2018) Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences* 115(37), 9216–9221.
- Bakker BN, Lelkes Y, and Malka A (2020) Understanding partisan cue receptivity: Tests of predictions from the bounded rationality and expressive utility perspectives. *The Journal of Politics* 82(3), 1061–1077.

- Baron J and Jost JT** (2019) False equivalence: Are liberals and conservatives in the United States equally biased? *Perspectives on Psychological Science* 14(2), 292–303.
- Bartels LM** (2002) Beyond the running tally: Partisan bias in political perceptions. *Political Behavior* 24(2), 117–150.
- Bartels LM** (2014) Remembering to forget: A note on the duration of campaign advertising effects. *Political Communication* 31(4), 532–544.
- Bénabou R and Tirole J** (2002) Self-confidence and personal motivation. *The Quarterly Journal of Economics* 117(3), 871–915.
- Berinsky AJ, Margolis MF, and Sances MW** (2014) Separating the shirkers from the workers? Making sure respondents pay attention on self-administered surveys. *American Journal of Political Science* 58(3), 739–753.
- Berry JM and Sobieraj S** (2013) *The Outrage Industry: Political Opinion Media and the New Incivility*. New York, NY: Oxford University Press.
- Bisgaard M** (2019) How getting the facts right can fuel partisan-motivated reasoning. *American Journal of Political Science* 63(4), 824–839.
- Bolsen T and Druckman JN** (2018) Do partisanship and politicization undermine the impact of a scientific consensus message about climate change? *Group Processes & Intergroup Relations* 21(3), 389–402.
- Bolsen T, Druckman JN, and Cook FL** (2014) The influence of partisan motivated reasoning on public opinion. *Political Behavior* 36(2), 235–262.
- Bolsen T and Palm R** (2019). Motivated reasoning and political decision making. In Redlawsk DP (ed), *Oxford Research Encyclopedia of Politics*. New York, NY: Oxford University Press. <https://doi.org/10.1093/acrefore/9780190228637.013.923>.
- Broockman DE, Kalla JL, and Westwood SJ** (2023) Does affective polarization undermine democratic norms or accountability? Maybe not. *American Journal of Political Science* 67(3), 808–828.
- Bullock JG** (2011) Elite influence on public opinion in an informed electorate. *American Political Science Review* 105(3), 496–515.
- Bullock JG, Gerber AS, Hill SJ, and Huber GA** (2015) Partisan bias in factual beliefs about politics. *Quarterly Journal of Political Science* 10, 519–578.
- Christensen J and Moynihan DP** (2024) Motivated reasoning and policy information: Politicians are more resistant to debiasing interventions than the general public. *Behavioral Public Policy* 8(1), 47–68.
- Coppock A** (2019) Generalizing from survey experiments conducted on Mechanical Turk: A replication approach. *Political Science Research and Methods* 7(3), 613–628.
- Coppock A** (2023). *Persuasion in Parallel: How Information Changes Minds About Politics*. Chicago, IL: University of Chicago Press.
- Coppock A, Gross K, Porter E, Thorson E, and Wood TJ** (2023) Conceptual replication of four key findings about factual corrections and misinformation during the 2020 US election: Evidence from panel survey experiments. *British Journal of Political Science* 53(4), 1328–1341.
- De Dreu CK and Van Knippenberg D** (2005) The possessive self as a barrier to conflict resolution: Effects of mere ownership, process accountability, and self-concept clarity on competitive cognitions and behavior. *Journal of Personality and Social Psychology* 89(3), 345.
- Druckman JN** (2012) The politics of motivation. *Critical Review* 24(2), 199–216.
- Druckman JN, Fein J, and Leeper TJ** (2012) A source of bias in public opinion stability. *American Political Science Review* 106(2), 430–454.
- Druckman JN and McGrath MC** (2019) The evidence for motivated reasoning in climate change preference formation. *Nature Climate Change* 9(2), 111–119.
- Edwards K and Smith EE** (1996) A disconfirmation bias in the evaluation of arguments. *Journal of Personality and Social Psychology* 71(1), 5.
- Fishbein M** (1963) An investigation of the relationships between beliefs about an object and the attitude toward that object. *Human Relations* 16(3), 233–239.
- Flynn D, Nyhan B, and Reifler J** (2017) The nature and origins of misperceptions: Understanding false and unsupported beliefs about politics. *Political Psychology* 38, 127–150.
- Gerber A and Green D** (1999) Misperceptions about perceptual bias. *Annual Review of Political Science* 2(1), 189–210.
- Graham MH and Singh S** (2024) An outbreak of selective Attribution: Partisanship and blame in the COVID-19 pandemic. *American Political Science Review* 118(1), 423–441.
- Green MC, Visser PS, and Tetlock PE** (2000) Coping with accountability cross-pressures: Low-effort evasive tactics and high-effort quests for complex compromises. *Personality and Social Psychology Bulletin* 26(11), 1380–1391.
- Groenendyk E and Krupnikov Y** (2021) What motivates reasoning? A theory of goal-dependent political evaluation. *American Journal of Political Science* 65(1), 180–196.
- Guay B and Johnston CD** (2020) Ideological asymmetries and the determinants of politically motivated reasoning. *American Journal of Political Science* 66, 285–301.
- Guess A and Coppock A** (2020) Does counter-attitudinal information cause backlash? Results from three large survey experiments. *British Journal of Political Science* 50(4), 1497–1515.

- Hart PS and Nisbet EC** (2012) Boomerang effects in science communication: How motivated reasoning and identity cues amplify opinion polarization about climate mitigation policies. *Communication Research* **39**(6), 701–723.
- Iyengar S, Sood G, and Lelkes Y** (2012) Affect, not ideology: A social identity perspective on polarization. *Public Opinion Quarterly* **76**(3), 405–431.
- Jamieson KH and Cappella JN** (2008) *Echo chamber: Rush Limbaugh and the conservative media establishment*. New York, NY: Oxford University Press.
- Jerit J and Barabas J** (2012) Partisan perceptual bias and the information environment. *The Journal of Politics* **74**(3), 672–684.
- Kahan DM, Peters E, Dawson EC, and Slovic P** (2017) Motivated numeracy and enlightened self-government. *Behavioral Public Policy* **1**(1), 54–86.
- Key VO** (1966) *The Responsible Electorate: Rationality in Presidential Voting, 1936–1960*. Cambridge, MA: Harvard University Press.
- Khanna K and Sood G** (2018) Motivated responding in studies of factual learning. *Political Behavior* **40**, 79–101.
- Kim JW** (2024) “Replication Data for: “Evidence Can Change Partisan Minds but Less So in Hostile Contexts””, <https://doi.org/10.7910/DVN/AWJOYM>, Harvard Dataverse, V1.
- Kim JW, Guess A, Nyhan B, and Reifler J** (2021) The distorting prism of social media: How self-selection and exposure to incivility fuel online comment toxicity. *Journal of Communication* **71**(6), 922–946.
- Kim JW and Kim E** (2021) Temporal selective exposure: How partisans choose when to follow politics. *Political Behavior* **43**, 1663–1683.
- Klar S** (2014) Partisanship in a social setting. *American Journal of Political Science* **58**(3), 687–704.
- Klein E** (2020) *Why We’re Polarized*. New York, NY: Simon and Schuster.
- Kunda Z** (1990) The case for motivated reasoning. *Psychological Bulletin* **108**(3), 480.
- Lavine HG, Johnston CD, and Steenbergen MR** (2012) *The Ambivalent Partisan: How Critical Loyalty Promotes Democracy*. Oxford University Press.
- Leeper TJ** (2014) The informational basis for mass polarization. *Public Opinion Quarterly* **78**(1), 27–46.
- Leeper TJ and Slothuus R** (2014) Political parties, motivated reasoning, and public opinion formation. *Political Psychology* **35**, 129–156.
- Levendusky M** (2013) *How Partisan Media Polarize America*. Chicago, IL: University of Chicago Press.
- Little AT** (2021) Detecting motivated reasoning. OSF Preprints. July 15. doi: [10.31219/osf.io/b8tvk](https://doi.org/10.31219/osf.io/b8tvk).
- Little AT, Schnakenberg KE, and Turner IR** (2020) Motivated reasoning and democratic accountability. *American Political Science Review* **116**, 1–17.
- Lodge M and Taber CS** (2000) Three steps toward a theory of motivated political reasoning. In Lupia A, McCubbins MD and Popkin SL (eds), *Elements of Reason: Cognition, Choice, and the Bounds of Rationality*, Chapter 9. Cambridge, UK: Cambridge University Press, pp. 182–213.
- Lord CG, Ross L, and Lepper MR** (1979) Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology* **37**(11), 2098.
- MacKuen M, Wolak J, Keele L, and Marcus GE** (2010) Civic engagements: Resolute partisanship or reflective deliberation. *American Journal of Political Science* **54**(2), 440–458.
- Mischel W, Ebbesen EB, and Zeiss AM** (1976) Determinants of selective memory about the self. *Journal of Consulting and Clinical Psychology* **44**(1), 92.
- Nyhan B, Porter E, Reifler J, and Wood TJ** (2020) Taking fact-checks literally but not seriously? The effects of journalistic fact-checking on factual beliefs and candidate favorability. *Political Behavior* **42**, 939–960.
- Nyhan B and Reifler J** (2010) When corrections fail: The persistence of political misperceptions. *Political Behavior* **32**(2), 303–330.
- Peterson E and Iyengar S** (2021) Partisan gaps in political information and information-seeking behavior: Motivated reasoning or cheerleading? *American Journal of Political Science* **65**(1), 133–147.
- Porter E, Wood TJ, and Bahador B** (2019) Can presidential misinformation on climate change be corrected? Evidence from internet and phone experiments. *Research & Politics* **6**(3), 2053168019864784.
- Prior M, Sood G, Khanna K et al.** (2015) You cannot be serious: The impact of accuracy incentives on partisan bias in reports of economic perceptions. *Quarterly Journal of Political Science* **10**(4), 489–518.
- Pyszczynski T and Greenberg J** (1987) Toward an integration of cognitive and motivational perspectives on social inference: A biased hypothesis-testing model. In *Advances in Experimental Social Psychology* **20**, 297–340.
- Redlawsk DP, Civettini AJ, and Emmerson KM** (2010) The affective tipping point: Do motivated reasoners ever ‘get it’? *Political Psychology* **31**(4), 563–593.
- Remnick D** (2017) The tribalism of American politics. *New Yorker*. Available from <https://www.newyorker.com/video/watch/the-tribalism-of-american-politics-2017-05-04>, accessed 1 November 2024.
- Taber CS, Cann D, and Kucsova S** (2009) The motivated processing of political arguments. *Political Behavior* **31**(2), 137–155.
- Taber CS and Lodge M** (2006) Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science* **50**(3), 755–769.

- Tappin BM, Berinsky AJ, and Rand DG** (2023) Partisans' receptivity to persuasive messaging is undiminished by countervailing party leader cues. *Nature Human Behaviour* 7(4), 568–582.
- Tappin BM, Pennycook G, and Rand DG** (2020) Thinking clearly about causal inferences of politically motivated reasoning: Why paradigmatic study designs often undermine causal inference. *Current Opinion in Behavioral Sciences* 34, 81–87.
- Wehner P** (2020) Why Trump supporters can't admit who he really is. *The Atlantic*.
- Wood T and Porter E** (2019) The elusive backfire effect: Mass attitudes' steadfast factual adherence. *Political Behavior* 41(1), 135–163.
- Zimmermann F** (2020) The dynamics of motivated beliefs. *American Economic Review* 110(2), 337–363.