Polarization Drives Online Sharing of Accurate Information As Much or More Than Misinformation

Brian Guay[1]
Adam J. Berinsky[2]
Gordon Pennycook[3]
David Rand[4]

[1]Department of Political Science, Stony Brook University
[1]Department of Political Science, Massachusetts Institute of Technology
[3]Department of Psychology, Cornell University
[4]Sloan School of Management, Massachusetts Institute of Technology

March 8, 2024

Global concern over misinformation has fueled a growing literature on what motivates people to share fake news. Recent research identifies polarization as the primary psychological motivation, citing correlations between polarization and sharing fake news on Twitter. However, such findings may not reflect that polarized Americans are more motivated to share fake news, but that they are 1) more exposed to misinformation or 2) more motivated to share news of any kind (real or fake). We disentangle these competing explanations by examining news sharing in an information environment with equal amounts of real and fake content. In contrast to recent work, we find little evidence that polarization is responsible for fake news sharing. In fact, polarization is often associated with sharing *less* fake news. Where polarization is associated with sharing more fake news it is more strongly associated with sharing real news, indicating that polarization is associated with sharing news in general rather than fake news specifically. Critically, polarization is consistently associated with the *normatively positive* outcome of sharing more true than false content. Implications for how to combat the spread of fake news are discussed.

Current version (here)

This working paper has not yet been peer reviewed

Misinformation has become an issue of global concern across issues from election denialism and vaccine hesitancy (Graham & Yair 2023, Loomba et al. 2021) to political violence and erosion of trust in democratic institutions (Ognyanova et al. 2021, Piazza, 2022). In response, a growing body of research has sought to identify effective approaches to reduce the belief in, and sharing of, misinformation (Guess et al., 2020, Pennycook et al., 2021, Badrinathan, 2021). The most widely discussed approaches to combating misinformation seek to equip people with tools to avoid believing and sharing falsehoods. This includes interventions such as fact-check based corrections and warning labels, digital literacy interventions, and prompts that shift people's attention towards accuracy. These approaches are predicated on the assumption that individuals believe and disseminate misinformation primarily due to an absence of information about, or attention to, the veracity of the content.

This assumption has been challenged by an alternative account which we will refer to as the *Polarization Perspective*. By this account, social media users share misinformation because they are primarily concerned with attacking their political opponents and supporting their own viewpoints when deciding which news to share—and thus are happy to share false content that advances their political agenda (Osmundsen et al. 2021, Marie, Altay, & Strickland 2023, Ekstrom & Lai 2020, Pretus et al. 2023, Jenke 2022). This account resonates with popular theoretical narratives which argue that identity and politically motivated reasoning are central to decision-making (Kunda 1990, Kahan 2013, Taber & Lodge 2006). Perhaps most importantly, this alternative account of misinformation sharing suggests that existing information-based approaches to combating misinformation (e.g. fact-checks, literacy tips, and accuracy prompts) will, at best, have limited impact —and that what is needed instead to protect against misinformation is depolarization.

The strongest empirical evidence in support of the Polarization Perspective comes from a recent influential study reporting that political polarization, rather than lack of information or inattention, is the primary predictor of whether or not Twitter users shared any links to fake news websites (Osmundsen et al. 2021). However, there are two conceptual issues which prevent the study's data and analyses from providing clear support for the Polarization Perspective in the context of misinformation. First, more polarized individuals may share more false content on social media not because they are more prone to sharing misinformation, but simply because they are exposed to a greater volume of misinformation. That is, people with different levels of polarization may be equally likely to share misinformation upon encountering it, but highly polarized individuals encounter misinformation more frequently (Mohsen & Rand 2022). We refer to this as the *exposure concern*, because differential exposure to false content may create the appearance that people are differentially *motivated* to share false content, even when they are not.

Second, polarization may be associated with a greater propensity to share political news *in general*, rather than fake news specifically. If people are primarily motivated to promote their own views and attack those of political opponents, there is no shortage of credible, partisan sources of true information with which to arm themselves (Grinberg et al. 2019, Guess et al. 2020). Thus, we might expect polarized people to not only share more politically aligned misinformation, but also share more politically aligned credible information. If this were the case, more polarized individuals would share more information of *all* types simply because they are more politically engaged, rather than actually being particularly psychologically susceptible to sharing *inaccurate* information. As a result, these individuals would not actually share worse content on average than less polarized individuals (i.e would not be less 'discerning' in their sharing behavior). We refer to this as the *discernment concern*, as differential preferences for sharing news in general may give the appearance of differential sharing of fake news specifically.

These two concerns imply that discrepancies in the sharing of false information could be explained by accounts other than the Polarization Perspective, and that prior findings do not necessarily imply that existing misinformation interventions will be ineffective. Without a research design and analysis approach that addresses these concerns, it is impossible to know whether polarization actually makes people particularly susceptible to sharing fake news, and in turn, whether existing interventions are ill-conceived.

Here, we present the results of a large, national survey experiment that addresses both concerns in order to provide a clearer picture of the relationship between polarization and fake news sharing. First, we address the exposure concern by exposing respondents to an equal number of true and false headlines, so that differential exposure cannot drive differential sharing of fake news. Second, we address the discernment concern by examining the relationship between polarization and the *difference in* sharing of true versus false news (i.e., discernment;

Guay et al. 2023)—so that preferences for sharing more news in general are not confused with preferences for sharing more fake news specifically.

Our findings are starkly inconsistent with the Polarization Perspective. First, across six measures of polarization, we find no evidence that polarization is associated with decreased sharing discernment. In fact, we find the opposite: polarization is associated with *greater* sharing discernment: polarized respondents share more real news than fake news. Moreover, we find limited evidence of a positive association between polarization and sharing fake news. Measures of negative out-party affect – those most often employed in discussions of the Polarization Perspective) are associated with sharing *less* fake news; and while measures of positive in-party affect and partisan extremity are associated with sharing more fake news, they are also associated with sharing more *true* news to a similar extent. Thus, we find no evidence that polarization is associated with sharing more fake news than real news. Finally, we observe a similar pattern of findings when reanalyzing Twitter sharing data from Osmundsen et al. (2021): although one out of three polarization measures is associated with sharing news from low quality sources, polarization is approximately 7 times more strongly associated with sharing news from high-quality sources - and thus, more polarized Twitter users are much *more* discerning in their sharing than less polarized users.

Our findings also challenge recent claims that polarization better explains fake news sharing than other politically relevant factors. In particular, the size of the associations between sharing fake news and polarization in our study are either on par with, or smaller than, those of other factors such as age, political knowledge, online political engagement, trolling, and cynicism. Overall, then, our results contradict the claim that polarization is particularly important for explaining the psychological susceptibility to sharing misinformation - and the resulting conclusion that information-based interventions are unlikely to be effective.

Why do People Share Misinformation?

Information Processing and Polarization Perspectives

In response to growing concern about proliferation of false and misleading content on social media (Lazer 2018), and with over half adults in the U.S. accessing news through social media (Shearer 2021), researchers have sought to understand what causes people to share misinformation online. Because the content seen by social media users is determined by what other users share, slowing the spread of misinformation is a primary goal of existing misinformation interventions. Designing these interventions requires understanding *why* people share misinformation. Two broad perspectives have emerged, each with their own implications for the design of misinformation interventions.

The first, which we refer to as the *Information Processing Perspective*[1], posits that people share misinformation due to various deficits in basic information processing, such as a failure to be attentive to accuracy (Pennycook et al. 2020, Pennycook et al. 2021), an overreliance on intuitive impressions (Pennycook & Rand 2019), and/or a lack of underlying knowledge or ability to understand the information being presented (Guess et al. 2020). Although there are several different mechanisms that relate to this perspective, they share the assumption that sharing misinformation is a category of reasoning error (Pennycook 2023) that could, at least in theory, be remedied by improving what people know and/or how they think. Consistent with this perspective, research shows that simply reminding people about accuracy increases the quality of content that people share (Pennycook et al. 2021). Digital literacy tips that provide people with tools to detect false and misleading content have been shown to have a similar effect (Guess et al. 2020, Badrinathan 2021).

The second perspective, which we refer to as the *Polarization Perspective*, posits that people share misinformation because they are motivated to share news that promotes their political views and undermines those of their opponents. This perspective is aligned with research showing that people's consumption of information is often motivated by directional or partisan aims (Kunda 1990, Kahan 2013, Taber & Lodge 2006). Accordingly, people may care more about attacking political opponents than sharing content that is true. Moreover, social media users often operate in ideologically homogeneous information environments in which there is no shortage of ideologically agreeable information to share (Barbera et al. 2015). Recent research observes correlations between affective polarization and believing misinformation (Jenke 2022), while other work finds that people who hold more extreme views on an issue are more likely to share misinformation about it (Marie, Altay, & Strickland 2023, Ekstrom & Lai 2020, Pretus et al. 2023).

---

[1] Also referred to as the "confusion" or "inattention" account by Pennycook et al. 2021 and "Ignorance Theory" by Osmundsen et al. 2021.

Differing Implications for How to Address the Misinformation Problem

Although these two perspectives do not entirely conflict (i.e., both may explain critical elements of misinformation sharing), they have starkly different implications for how to slow the spread of misinformation online. Most interventions aimed at combating the spread of misinformation originate from the Information Processing Perspective and are designed to provide, or draw attention to, information about the accuracy of online content. For instance, many interventions aim to provide information with which social media users can detect misinformation, often via digital literacy training (Badrinathan 2021, Guess et al. 2020, van der Linden et al. 2017). Other interventions involve tagging to social media posts to warn users that they may contain false or misleading misinformation (Clayton et al. 2020). Some interventions instead aim to draw users' attention to the accuracy of content by reminding them of the importance of sharing accurate content online (Pennycook et al. 2021, Pennycook et al. 2020).

The Polarization Perspective, on the other hand, suggests that depolarization is the key to slowing the spread of fake news. Osmundsen et al. (2021) conclude that researchers should not be surprised when existing fact-checking and nudging interventions fail to slow the spread of misinformation, given that they are not addressing the root problem. Instead, they advocate for the admittedly more challenging task of reducing polarization among the population. While there is no research to date testing the effect of depolarization on news sharing, a growing body of work has tested interventions aimed at reducing polarization, with mixed success (Voelkel et al. 2023, Combs et al. 2023).

Concerns With Using Observational Social Media Data

Perhaps the strongest evidence for the Polarization Perspective comes from a recent influential study that uses observational social media data of sharing on Twitter to understand the relationship between polarization and news sharing. Osmundsen et al. (2021) scraped the tweets of 2,337 U.S. survey respondents during a two month period in 2019 and found that polarization predicted fake news sharing more than alternative factors, such as digital literacy, political knowledge, and cognitive reflection. This led them to conclude that "partisan polarization is the primary psychological motivation behind political fake news sharing on Twitter." Other prominent studies use similar approaches to understand who shares misinformation online. For instance, Grinberg et al. (2019) use Twitter data to show that older, Republican, and politically engaged users are more likely to share fake news. Guess et al. (2019) use Facebook data to show that older and Republican users were more likely to share fake news on Facebook during the 2016 campaign.

Using observational social media data (i.e., digital trace data) in this way has obvious benefits for external validity, as it directly observes how social media users behave online. However, in the case of inferring the true nature of the relationship between polarization and news sharing, this form of data introduces two concerns.

Exposure Concern

First, the observational approach does not capture how much fake news social media users are *exposed to misinformation*, which is likely correlated with polarization. The amount of fake news someone shares is a product of both their motivation to share fake news *and* their exposure to fake news (Guay et al. 2023, Trilling et al. 2022). Two people who share 50% of the fake news they encounter will share far different amounts of fake news if one is exposed to 10 fake news articles a month and the other is exposed to 100. Unfortunately, studies that use observational sharing data do not capture exposure (Grinberg et al. 2019, Lazer 2019). While academic researchers can observe the amount of fake news people engage with (e.g., share or 'like'), they cannot typically measure the amount of news articles people see because social media platforms do not make exposure data publicly available. Underscoring the importance of exposure, approaches to approximating exposure indicate that it is a major factor in predicting sharing, and that associations between traits and sharing behavior (e.g. partisanship and fake news sharing) that appear without account for exposure are eliminated once a rough proxy for exposure is entered as a control (Grinberg et al. 2019).

This is particularly problematic for studies aimed at measuring the relationship between polarization and fake news sharing, given that people with high levels of polarization are more exposed to fake news.[2] For instance, Mohsen & Rand (2022) find that among 5,000 Twitter users, polarization (captured using ideological extremity) is positively correlated with both

[2] This is especially problematic given that alternative predictors of sharing fake news (e.g., cognitive reflection, political knowledge, digital literacy) do not necessarily drive exposure in the same way that polarization does. This sets up an uneven comparison, in which polarization may influence sharing through exposure, resulting in a stronger association, but alternative predictors may not.

following elites who make more false claims—implying greater exposure to misinformation—and sharing links to lower quality news sources.

Thus, exposure presents a confound: highly polarized individuals may share more misinformation not because they are more inclined/motivated to share it, but because they are simply more exposed to it. Critically, in such a situation, while it may be descriptively accurate to say that polarized people share more fake news, it is inaccurate to claim that polarization is the *motivation* or *mechanism* for sharing fake news. It may well be that it is simply exposure.

Discernment Concern

The second concern with using observational sharing data to examine the relationship between polarization and fake news sharing is that polarization may be associated with sharing news *in general*, not with sharing fake news specifically. To illustrate, consider two types of people, those who use Twitter daily and those who use Twitter once a year. Descriptively, it will almost certainly be the case that daily users share more false news than yearly users. But daily users will also share more true news—and it seems hard to imagine that anyone would conclude from these data alone that Twitter usage is the motivation behind sharing fake news.

If people are motivated to bolster their political views and attack those of opponents, as suggested by the Polarization Perspective, there is no reason that this should only result in sharing fake news. There is no shortage of true news—hyperpartisan or even entirely accurate— with which to buttress one's own political views and attack those of opponents. In fact, real news is far more prevalent on social media than fake news (Grinberg et al. 2019, Guess et al. 2020). Thus, it seems likely that partisan motives would drive the sharing of true politically concordant news as much or more than the sharing of false politically concordant news. That is, partisan motives seem likely to drive the sharing of politically concordant news regardless of veracity, rather than specifically motivating people to share false politically concordant news.

Beyond leading to an incorrect conclusion about the underlying psychology and motivations, focusing primarily on the relationship between polarization and sharing fake news neglects the larger context of the information environment on social media. Sharing behavior matters because what one user shares determines the quality of the information environment for others—what content others in their network are exposed to. Thus, to determine a user's impact on the information environment requires considering the amount of both true and false news they share. Guay et al. (2023) argue that *sharing discernment*–the difference between the amount of true and false content people share—best reflects this impact.

This also has important implications for intervention efficacy: if polarization caused more sharing of true and false news, then depolarization interventions would not necessarily improve the quality of the information environment. In fact, if polarization was more strongly linked to sharing true compared to false news, depolarization could actually reduce the average quality of news shared. Given the far greater supply of real news than fake news on social media, this could also occur when an intervention has an identical or even weaker effect on reducing sharing of real news than fake news (Guay et al. 2023).

From a measurement perspective, even if one were to consider relationships with both the amount of true and false content sharing, studies using observational social media cannot fully capture discernment due to a lack of data on what content users are exposed to. This lack of exposure data means that we do not know what content users see and choose *not* to share, which can bias measures of discernment. To illustrate, imagine trying to understand which of two people share higher quality news. Both share 10 real articles and 10 fake articles, but Person A is exposed to 80 real articles and 20 fake ones while Person B is exposed to 20 real articles and 80 fake ones. Without data on how many articles they are exposed to, it appears that they are equally discerning between real and fake news: each shares 50% fake. However, proportional to what they are exposed to, person A shares far more fake news (50%, 10/20) than person B (12.5%, 10/80). Similarly, Person B shares far more real news (50%) than person A (12.5%).

A design that addresses both the exposure and discernment concerns exposes people to a set of headlines with a fixed proportion of real and fake headlines, and measures their likelihood of sharing each. By using self-reported sharing intentions (Altay et al., 2020, Petersen et al. 2023, Marie & Petersen 2023, Pennycook et al. 2021, Mosleh et al. 2020) rather than observational sharing data, we are able to hold exposure constant across varying levels of polarization (addressing the exposure concern) and measure discernment (addressing the discernment concern). Exposure is held constant by exposing all respondents to equal amounts of true and false content, and discernment is calculated by computing the difference in sharing both types of content.

Data & Methodology

Sample

We recruited 1,000 U.S. adult respondents in November 2022 from YouGov, which draws respondents from a demographically balanced panel of U.S. survey respondents who are invited to participate in surveys for compensation. All analyses are weighted according to gender, age, race, education, region, and past presidential vote based on data from the U.S. Census Bureau, the 2020 Congressional Election Study, and 2020 election exit polls. A more detailed description of the sampling procedure and weights is included in the Supplementary Materials (2.4).

All analyses were pre-registered (pre-registration is available here). We pre-registered our decision to exclude respondents who engaged in straightlining on survey grid questions—i.e., selecting identical response options to respond to very different questions in a way that is highly implausible for attentive respondents. We also pre-registered our decision to include these inattentive respondents (N = 62) in a supplemental analysis. We report the results of this analysis, which do not meaningfully differ from the main results, in the Supplementary Materials (1.1), as well as a more detailed discussion of our process for screening out inattentive respondents. 84.8% of the sample reported using social media (Facebook, Twitter, Instagram, TikTok, Snapchat, Reddit, WhatsApp, or another social media network).[3]

Sharing Intentions

Respondents were then shown 20 recent news headlines and reported how likely they would be to share each. Instructions read "Next you will be presented with a set of news headlines (20 in total). We are interested in how likely you would be to share these stories online." The subsequent screens each featured a news article as it would appear on social media, with a headline ("e.g., Special Forces Arrest Deep State Dr. Anthony Fauci"), the abbreviated first sentence of the article ("US Special Forces on Saturday scored a major victory in the war…"), the source of the headline (e.g., realrawnews.com), and the original image accompanying the

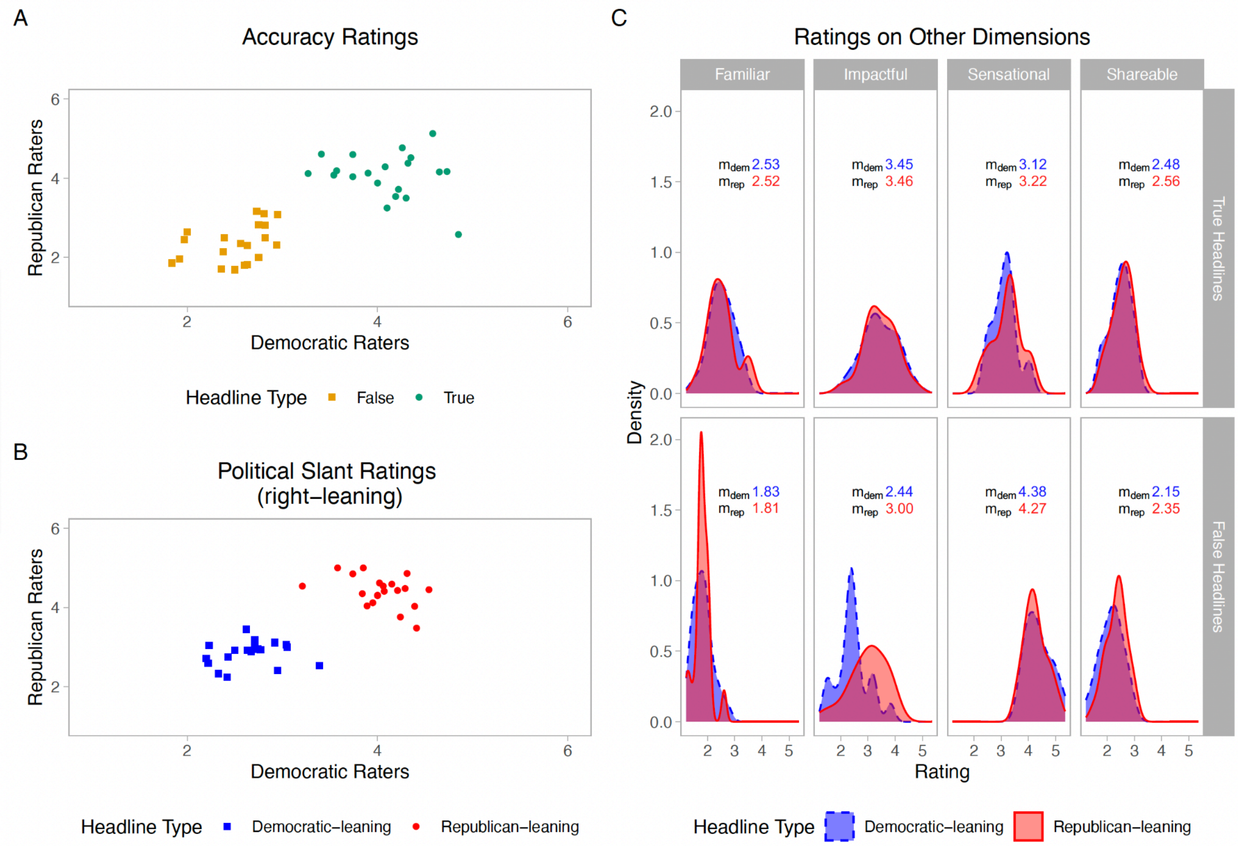
[3] In the Supplementary Materials (1.2) we reproduce our main findings using only respondents who report using social media.

headline when it was published (see the Supplementary Materials 2.3 for all headline stimuli). For each headline, respondents were asked ``If you were to see the above post on Facebook, how likely would you be to share it?'' and answered using a 6-point Likert scale ranging from ``extremely unlikely'' to ``extremely likely.”

Sharing intentions are commonly used in misinformation research (e.g., Altay et al., 2020, Petersen et al. 2023, Marie & Petersen 2023, Pennycook et al. 2021) and headline-level analyses find them to be correlated with actual sharing behavior on social media—and, even more importantly, to show similar correlation patterns with a range of covariates as is observed using actual sharing (Mosleh et al., 2020). An additional strength of this design is that it allows us to know the veracity of the headlines seen by respondents with high certainty, rather than relying on the reputation of news sources as a proxy for headline veracity (Guess et al. 2019, Grinberg et al. 2019, Osmundsen et al. 2021).[4]

[4]Rather than determining the veracity of each headline individually, studies that use observational sharing data often rate the credibility of a domain (e.g., breitbart.com) and apply that rating to all news stories from that domain. This is problematic because fake news websites often publish real content as a means of bolstering their credibility.

Figure 1: Pre-test Headline Ratings



Ratings of headlines from a pre-test on a separate sample. Panels A and B: Respondents rated false headlines less likely to be true than true headlines, and Republican-leaning headlines as more right-leaning than Democratic-leaning headlines. [[Say something about the scale.]] Ratings from Democratic and Republican respondents were similar. Panel C: Density plots of ratings of Democratic-leaning and Republican-leaning headlines as familiar, impactful, sensational, and [[likelihood of sharing]]. Means for democratic-leaning headlines (m_{dem}) and republican-leaning headlines (m_{rep}) are reported.

The 20 headlines viewed by respondents were balanced by partisanship (left-leaning, right-leaning) and veracity (true or false). Each respondent saw 5 true left-leaning headlines, 5 false left-leaning headlines, 5 true right-leaning headlines, and 5 false right-leaning headlines. When referring to headlines used in the study, we use the terms `true` and `false` to describe real and fake headlines, respectively, to underscore the fact that all of the headlines used appeared on actual websites (as opposed to fake headlines that are made up by researchers, e.g, Pereira et al. 2018). In order to increase generalizability, the 20 headlines viewed by each respondent were sampled from a larger set of 40 headlines.[5] We followed Pennycook et al.'s (2021) methodology for selecting headlines, drawing false articles from fact-checking websites (e.g., Snopes.com) and true news articles from a wide variety of mainstream sources.

Polarization

In order to thoroughly measure polarization, we used the three polarization measures used by Osmundsen et al. (2021) (partisan extremity, in-party emotions, out-party emotions) and three additional measures (out-party traits, out-party feeling thermometer, in-party feeling thermometer). The exact wording of all survey questions is included in the Supplementary Materials (2.1-2.3), and all scales were constructed such that high values represent greater levels of polarization. We measured partisan extremity by asking respondents to place themselves on a 7-point scale from "Strong Democrat" to "Strong Republican" and operationalize extremity as the absolute distance from the midpoint. Also following Osmundsen, we asked respondents the extent to which they feel positive emotions (hopeful, enthusiastic, proud) and negative emotions (angry, frustrated, afraid) when they think of Democrats and Republicans. We used these items to create separate measures of positive in-party emotions and negative out-party emotions, where high values on each indicate greater polarization.[6]

We included three additional polarization measures. Respondents rated how they felt about Democrats and Republicans on a feeling thermometer ranging from 0 (cold, unfavorable) to 100 (warm, favorable), which were used to create in-party and out-party feeling thermometer measures (Lelkes and Westwood 2017). Respondents also rated the extent to which certain traits describe members of the political party they do not belong to (i.e., out-party): patriotic, intelligent, honest, open-minded, generous, hypocritical, selfish, and mean (Iyengar, Sood, and Lelkes 2012; Garrett et al. 2014). We used these responses to a mean scale for negative affect toward the out-party by reverse-coding questions measuring positive traits.

Modeling Sharing Discernment

---

[5] We used stratified random sampling to assign headlines to each respondent: each respondent was randomly assigned 5 headlines of each type (true left-leaning, false left-leaning, true right-leaning headlines, and false right-leaning).

[6] To construct the positive in-party emotion scale, we used responses to questions measuring the extent to which respondents feel both positive and negative emotions about the in-party, where the negative emotions were reverse coded. We followed the same approach to construct the negative out-party emotion scale, but reverse-coded positive emotions.

As described above, we use sharing discernment as our primary outcome of interest, which reflects the difference in sharing true versus false news. Guay et al. 2023 outline two types of discernment, based on different methods of calculating this difference. *Additive discernment* reflects the additive difference between sharing true and false news--e.g., if respondents share an average of 10 true and 6 false news articles, additive discernment is 4 (10 - 6 = 4). *Multiplicative discernment* reflects the multiplicative difference between sharing true and false news—in the same example, respondents are 1.7 times more likely to share true news than false news (10 / 6 = 1.7). Since multiplicative discernment is reported in the form of a ratio, 1 indicates no discernment (people are just as likely to share true news as false news), values greater than 1 indicate positive discernment (sharing more true news than false news), and values less than 1 indicate negative discernment (sharing more false news than true news). We consider both types of discernment, modeling additive discernment in the main analysis and multiplicative discernment in the Supplementary Materials (1.3-1.4). Both types of discernment yield similar conclusions about the relationship between polarization and sharing misinformation.

Our primary modeling approach is to predict sharing intentions (on a 0 to 1 scale) with a two-way interaction between a dummy variable for headline veracity (0=false, 1=true) and each predictor of interest (age, cognitive reflection, political unawareness, digital media literacy, trolling, political cynicism, partisanship, and each measure of polarization), estimating separate models for each predictor of interest. Each model also includes two-way interactions between veracity and each of the control variables (gender, income, education, race, and political interest).

All models use survey weights and two-way clustered standard errors, as the observations are nested within both respondents (each respondent rates multiple headlines) and headlines (each headline is seen by multiple respondents). We fit the models using Maximum Likelihood Estimation with a Gaussian distribution and identity link function, which results in parameter estimates equivalent to those estimated using Ordinary Least Squares.[7] We use these models to calculate both types of discernment by predicting sharing intentions while holding variables (e.g., headline veracity and headline partisan lean) at set values. We use simulation-based

[7] We use the *glm* function in R, which simplifies the process of calculating two-way clustered standard errors with survey weights.

inference to construct confidence intervals and perform hypothesis testing (King et al. 2000), using the *Clarify* package in R (Griefer et al. 2023).[8]

Results

Polarization and Sharing Behavior

We begin by considering the relationship between polarization and sharing discernment. We pre-registered our decision to analyze agreeable and disagreeable headlines separately, as the Polarization Perspective predicts that polarization should be associated with lower discernment for both agreeable and disagreeable content for different reasons. For agreeable content polarization should increase the amount of false content shared (assuming that people are already sharing true agreeable content), whereas for disagreeable content, polarization should decrease the amount of true content shared (assuming people are already refraining from sharing false disagreeable content).

Figure 2 illustrates the relationship between polarization and sharing discernment. We first consider ideologically agreeable headlines (Panels A & B). Panel A reports the change in mean sharing intentions associated with a one standard deviation increase in each measure of

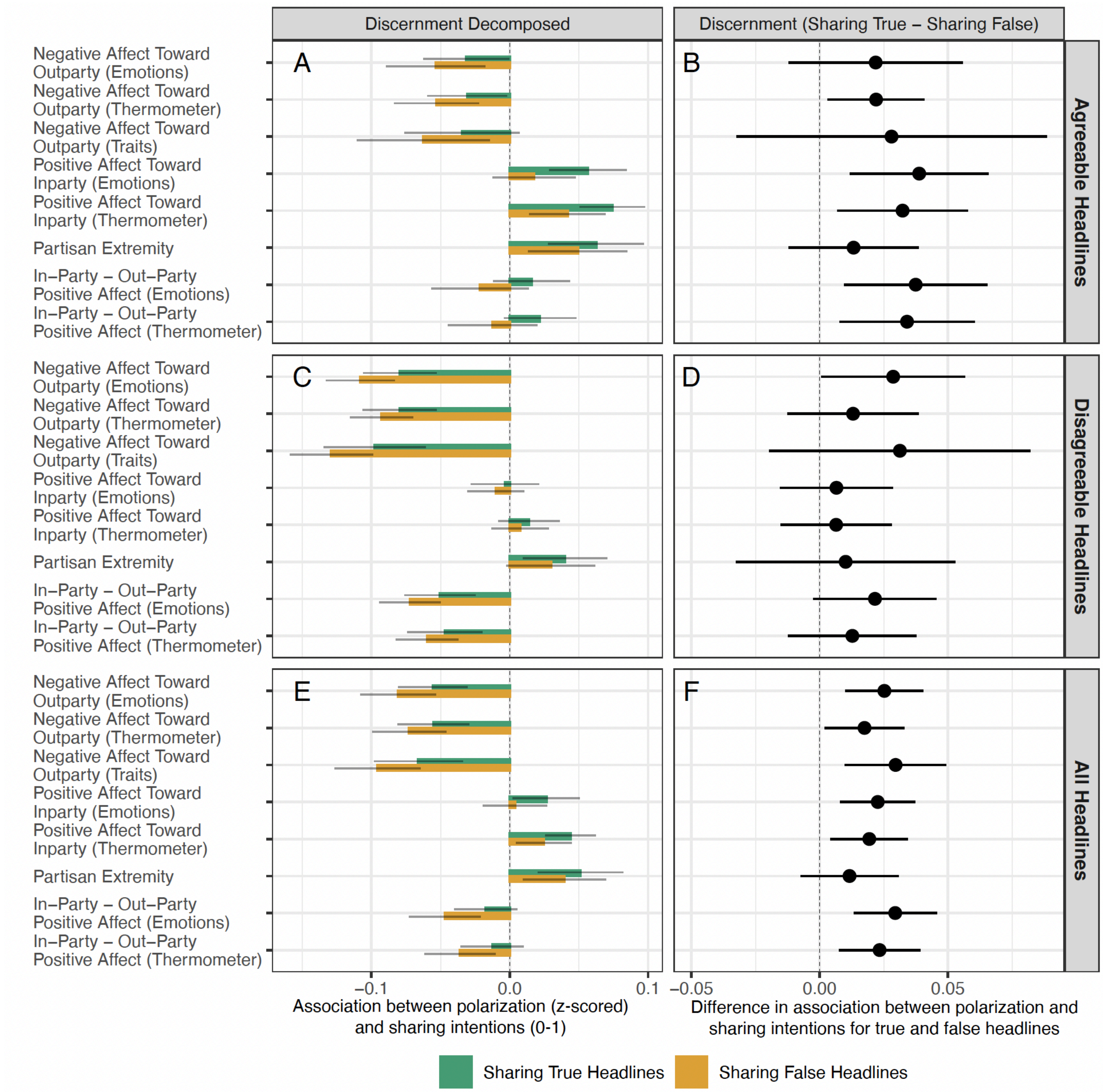
[8] Like bootstrapping, simulation-based inference calculates a quantity of interest n times (typically n > 1,000) and uses the resulting distribution of quantities to construct confidence intervals and perform hypothesis testing. Here, the quantity of interest is (mean of predicted sharing intentions for true headlines) / (mean of predicted sharing intentions for false headlines) for multiplicative discernment, and (mean of predicted sharing intentions for true headlines) - (mean of predicted sharing intentions for false headlines) for additive discernment. In the bootstrapping approach, a model is run on n randomly sampled subsets of the data and quantities are computed using each set of parameter estimates. In simulation-based inference, one model is run on the full dataset, producing a single set of coefficients (parameter estimates and standard errors). Then, n sets of parameter estimates are sampled from a normal distribution with a mean and standard deviation given by the original parameter estimates and standard errors, respectively. Quantities of interest are then calculated using each set of simulated point estimates, and the resulting distribution of quantities is used to calculate point estimates, construct confidence intervals, and perform hypothesis testing (e.g., the mean of the distribution is the point estimate, the standard deviation is the standard error, and the 2.5th and 97.5th quantiles represent the 95% confidence interval). For multiplicative discernment, the quantity of interest is the ratio of predicted sharing intention for true articles to the predicted sharing intention for false articles. For additive discernment, the quantity of interest is the difference in predicted sharing intentions for true articles and false articles. Note that the point estimate and confidence intervals obtained in this method for additive discernment are nearly identical to the OLS parameter estimate for the two-way interaction between headline veracity and each characteristic of interest (e.g., polarization).}

polarization.[9] For instance, the top row of Panel A illustrates that a one standard deviation increase in negative out-party emotions is associated with a 0.033 decrease in sharing true news and a 0.054 decrease in sharing false news. Sharing discernment is the difference between these and is plotted in the first row of Panel B (0.033 - 0.054 = -0.021). Here, polarization is associated directionally, though not significantly ($p = 0.145$), with *greater* discernment.

This pattern of a directionally positive, but not statistically significant, relationship between polarization and sharing discernment holds for all three measures of negative out-party affect. In all cases, polarization is associated with sharing less news of any kind (true and false), but the relationship between polarization and sharing false headlines is approximately twice that of polarization and sharing true headlines. Thus, more polarized individuals exhibit greater sharing discernment because they are less likely to share news of any kind, and especially false news.

[9] From a model predicting sharing intentions (0-1) with each measure of polarization, controlling for gender, income, education, race, and political interest. Standard errors are clustered at the respondent and headline level. We ran separate models for true and false headlines.

Figure 2: Polarization and Sharing Discernment, By Headline Agreeableness



Panels A, C, & E: Association between each polarization measure (z-scored) and news sharing intentions (measured on a 0-1 scale), separately for agreeable (A) and disagreeable (C) headlines, and for all headlines (E). Panels B, D, & F: The association between polarization and sharing discernment, i.e. how much more respondents share true news than false news. Each value in Panels B, D, & F is equivalent to the difference between the corresponding values in Panels A, C, & E. For example, the difference between the green and yellow bars in the first row of Panel A (-0.033 - -0.054) is approximately equivalent to the parameter estimate reported in the first row of Panel B (0.021). All models use survey weights and control for gender, income, education, race, and political interest. Horizontal lines represent 95% confidence intervals.

For measures of positive in-party affect and partisan extremity, the relationship between polarization and sharing discernment is also directionally positive, and statistically significant in one case. Here, polarization is associated with sharing *more* news of any kind, and especially more true news. For instance, a one standard deviation increase in positive in-party emotions is associated with a 0.038 increase in sharing discernment, resulting from a 0.056 increase in sharing true content and a 0.018 increase in sharing false content (0.056 - 0.018 = 0.038). More polarized individuals exhibit greater sharing discernment because they are more likely to share news of any kind, and especially true news.

These findings suggest that when holding exposure constant, the relationship between polarization and news sharing is starkly inconsistent with the Polarization Perspective. First, across six measures of polarization, we find no evidence that polarization is associated with lower sharing discernment. In fact, we consistently find that polarization is directionally (and in one case, significantly) associated with normatively *positive* behavior: greater sharing discernment. Second, we find mixed evidence regarding the association between polarization and sharing of false content: for half of our polarization measures the relationship is positive and for half the relationship is negative. We consider potential explanations for this in the discussion below. Third, in all cases, the relationship between polarization and sharing false news is directionally the same as it is for true news. In other words, polarization is associated with news sharing behavior in general, not false or true news sharing in particular.

While we expected the mechanism underlying the relationship between polarization and discernment to vary by headline agreeableness, we observe an almost identical pattern of findings for disagreeable headlines in Panels C and D. Here again, the association between polarization and sharing discernment is directionally positive, though not statistically significant. Negative out-party affect is associated with sharing less news, and sharing less false content in particular. The associations between positive in-party affect and sharing intentions are not significant for true or false content. As was the case for agreeable headlines, partisan extremity is associated with sharing more content, especially more true content.

Given a lack of differences in our findings across agreeable and disagreeable headlines, we conduct a *post hoc* analysis that pools across headline agreeableness to increase statistical power, and report our findings in Panels E & F. Increasing power by pooling across agreeable and disagreeable headlines clarifies the relationship between polarization and sharing discernment: polarization is associated with greater sharing discernment across all measures of polarization. This relationship is statistically significant for four of the six measures at the alpha = .05 level (p values for negative out-party feeling thermometer and partisan extremity are 0.051 and 0.225, respectively).

Re-analyzing Osmundsen et al. (2021)

One potential explanation for the difference between our conclusions and those of Osmundsen et al. (2021) is that the types of headlines viewed by respondents in our study are somehow different from those shared by respondents in theirs. To account for this, and any other design differences between these studies that may have contributed to any differences in findings, we re-analyze their data with the approach used above.

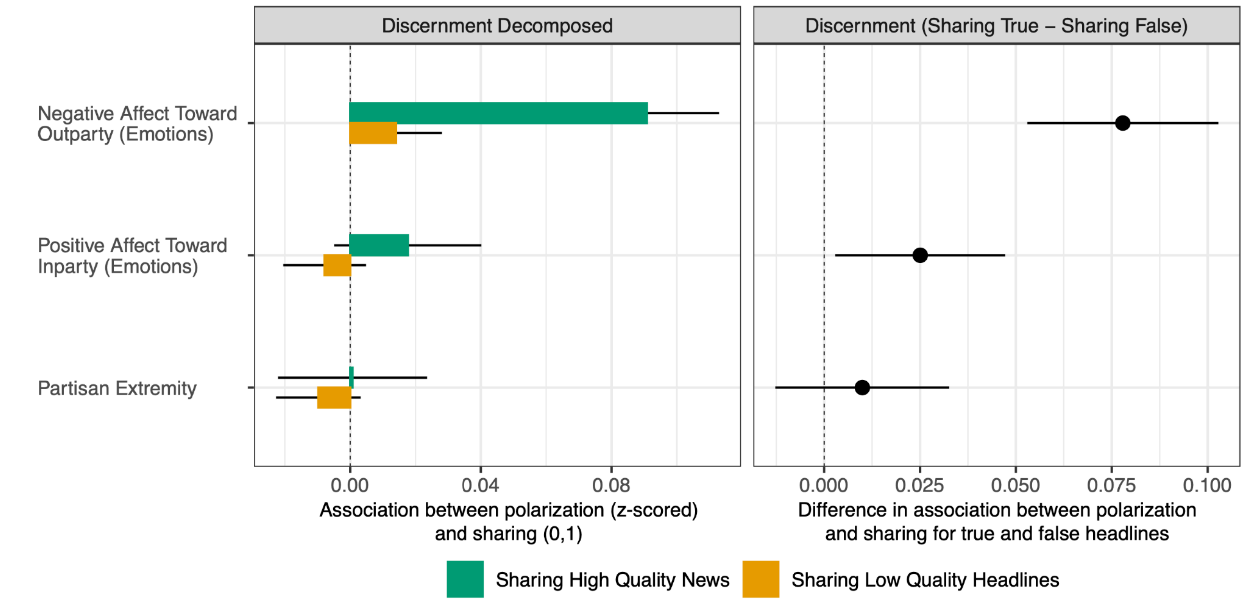
As discussed previously, Osmundsen's study differs from ours in its use of Twitter data to observe what users actually share online, preventing an accurate measure of sharing discernment due to a lack of information about what news users are exposed to. Nonetheless, we use a proxy for discernment, calculating how the association between polarization and news sharing varies for true and false content.[10] We follow Osmundsen et al. in modeling the association between polarization and whether or not respondents shared any links to news sources (1 = shared at least one link, 0 = shared no links), separately for high quality (i.e., true) and low quality (i.e., false) sources. We then use the simulation-based inference approach described above to calculate the difference in this association between high quality and low quality sources.

As reported in Figure 3, when analyzed in a manner analogous to our study, Osmunden et al.'s data tell a very similar story to that found in our data and described above. Polarization—both negative out-party affect and positive in-party affect—is associated with *greater* sharing discernment in their data as well as ours. In particular, the association between negative out-party affect and sharing is 6.5 times larger for high quality news outlets than low quality news outlets. Interestingly, only one measure of polarization, negative out-party affect, is associated with being more likely to share any low quality news, while positive in-party affect

[10] This approach makes the assumption that exposure to true and false content is constant across levels of polarization, which, as discussed above, is unlikely to hold.

and partisan extremity are not significantly associated with sharing low quality news (and are directionally opposite of each other).

Figure 3: Polarization and Sharing Discernment, Using Data From Osmundsen et al. (2021)



Left panel: Association between each polarization measure and the predicted probability of sharing true and false news content. Right panel: The association between polarization and sharing discernment, i.e. how much more respondents share true news than false news.

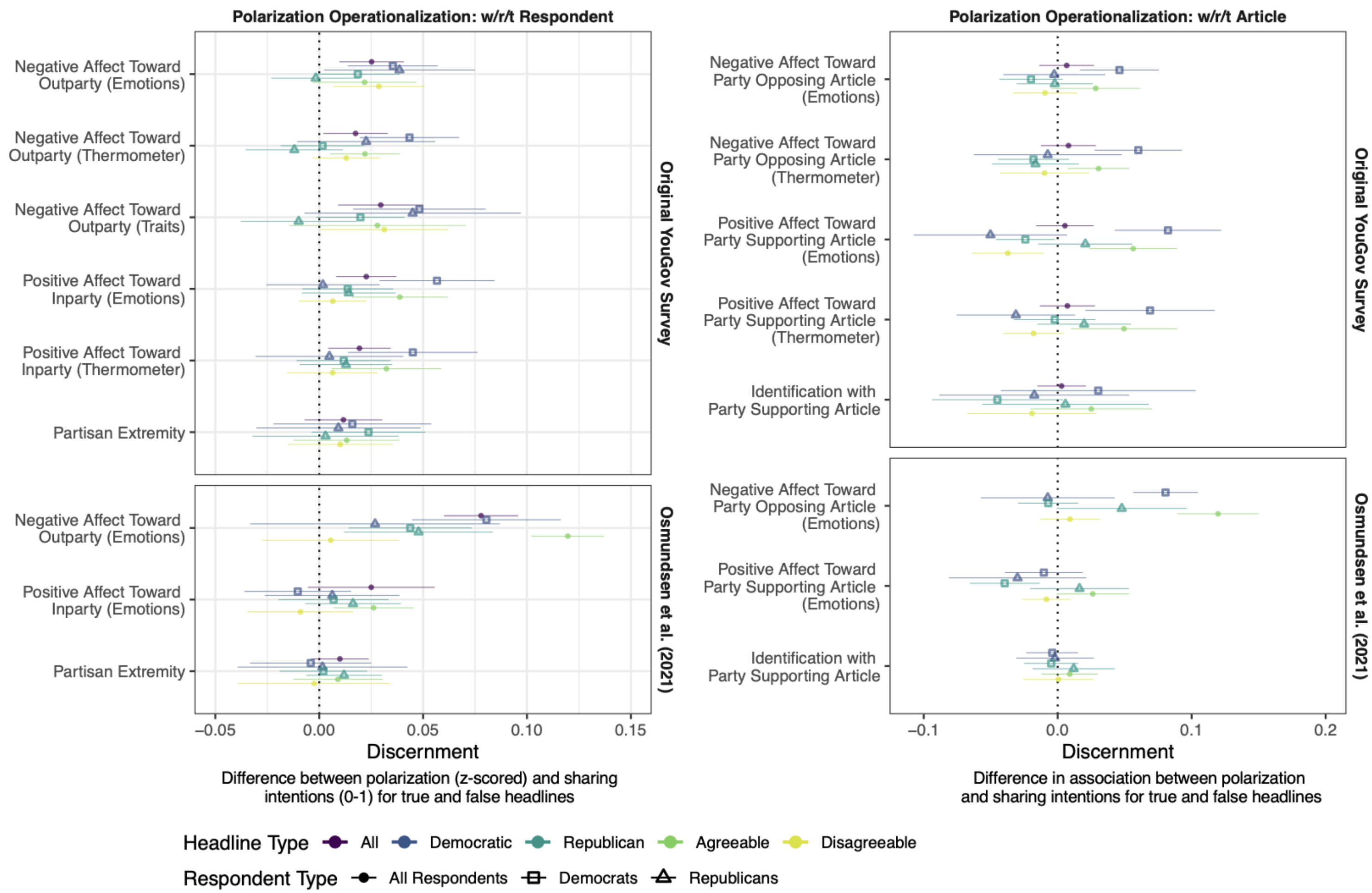
Robustness Checks

Although we find no evidence that polarization is ever associated with decreased sharing discernment across six measures of polarization in our data and three measures of polarization in Osmundsen et al.'s (2021) data, there are other reasonable ways of analyzing the data. In this section, we conduct post hoc analyses testing the robustness of these findings to different methods of operationalizing polarization and subsetting data.

Specifically, we test the effect of two alternative analysis decisions used by Osmundsen et al. (2021). First, we employ an alternative operationalization of polarization. In the preceding analysis, each polarization measure reflects the extent to which the individual respondent is polarized–the degree to which they view their own party favorably and the opposing party unfavorably. In other words, polarization is operationalized with respect to the survey respondent. Another way of operationalizing polarization is with respect to the news article, where polarization reflects how favorably a respondent views the party that is supported by the news article and how unfavorably a respondent views the party that is opposed by the news article.[11] Second, whereas we previously subset the analysis by headline agreeableness (i.e., whether the headline aligns with the respondent's political views) we can also subset the analysis by headline party (i.e., whether the headline aligns with the Democratic or Republican party).

[11] Both operationalizations use the same survey measures of polarization, but code them differently.

Figure 4: Alternative Operationalizations of Polarization



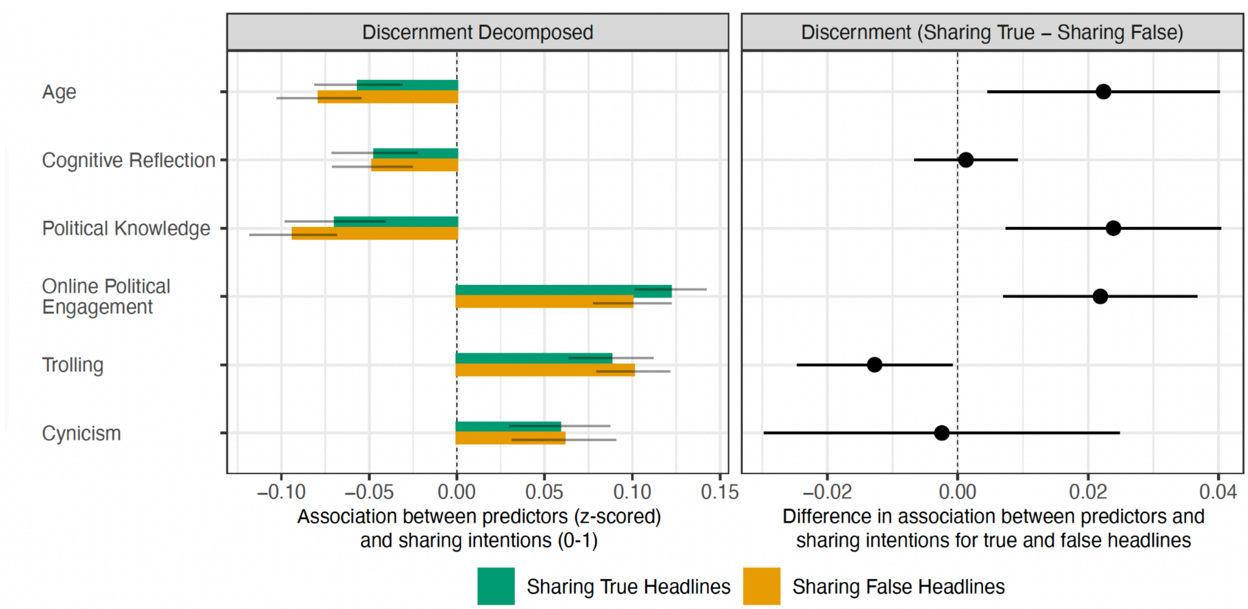
Coefficients representing the association between polarization and sharing discernment for multiple operationalizations of polarization and subsets of the data, for data collected from the original Yougov survey and Osmundsen et al. (2021). Vertical lines represent 95% confidence intervals. Trait measures of affective polarization are available only for out-party ratings, and thus are not included when operationalizing polarization with respect to the article seen by respondents (since in-party articles would have no rating).

Figure 4 reports the association between polarization and sharing discernment for each method of operationalizing polarization and subsetting the data, both for our data and Osmundsen et al.'s data. For data from our YouGov survey, the association between polarization and sharing discernment is significantly positive for 23/77 specifications (30%), significantly negative for 2/77 specifications (3%), and not significantly different from zero for 52/77 specifications (67%) .

For data from Osmundsen et al. (2021), the association is significantly positive for 11/39 specifications (28%), significantly negative for 1/39 specifications (3%), and not significantly different from zero for 27/39 specifications (69%). Thus, no matter how polarization is operationalized, the preponderance of evidence from both studies suggests that polarization is not associated with worse sharing discernment, and may be associated with better sharing discernment.

Other Predictors of Sharing Discernment

Figure 5: Other Predictors of Sharing Discernment



Left panel: Association between each predictor (z-scored) and news sharing intentions (measured on a 0-1 scale). Right panel: The association between each predictor and sharing discernment, i.e. how much more respondents share true news than false news. All models use *survey weights and control for gender, income, education, race, and political interest. Horizontal lines represent 95% confidence intervals.*

Contrary to findings from past work, Osmundsen et al. (2021) observed that age (Nagler and Tucker 2019), political knowledge, online political engagement, cognitive reflection (cite), trolling (Buckels, Trapnell, and Paulhus 2014), and cynicism (Petersen, Osmundsen, and Arceneaux 2018) had little to no association with news sharing, leading them to conclude that polarization is the primary psychological driver of sharing fake news and raise doubts about the efficacy of existing interventions that seek to slow the spread of misinformation. We included measures of these factors in our survey as well.[12] These measures serve both to evaluate the influence of non-polarization predictors of sharing behavior using our alternative design and to contextualize the size of the associations we observe between polarization and sharing behavior.

Figure 5 illustrates the association between these factors and sharing. Several things stand out. First, contrary to Osmundsen et al.'s finding that polarization is the primary correlate of sharing fake news, the size of these relationships in Figure 5 is on par with the size of associations between polarization and sharing (Figure 2). For instance, the largest association between polarization and sharing in Figure 2E is a 0.096 decrease in sharing false news for a one standard deviation increase negative affect toward out-party. The largest association between a non-polarization factor and sharing is for online political engagement, and is approximately the same size, but in the opposite direction (-0.122). The average absolute size of the association between polarization variables and sharing true or false news 0.046 (Panel 2C), 0.022 for discernment (Panel 2D), whereas the average absolute size of the association between non-polarization variables and sharing true or false news 0.077 (Figure 5, left), 0.014 for discernment (Figure 5, right).

[12] Citing work by Gil de Zúñiga et al. (2014) and Feezell (2016), Osmundsen et al. measured digital media literacy (reversed coded in their analysis and referenced as digital media illiteracy) by asking respondents how frequently they engage in various online activities. However, these measures are measures of online political engagement and are not intended to be used as measures of digital literacy (see Gil de Zuniga et al. 2014 and Feezell 2016). We used identical survey questions, but reference it as online political engagement (Guess & Munger 2023). The full list of activities measured is included here: starting or joining a political group or group supporting a cause on a social networking site, posting online personal views related to politics or campaigning, sharing someone else's political post to other people online, emailing a national, state, or local government official about an issue of personal importance.

Moreover, unlike Osmundsen et al. (2021), these associations are in the direction expected from previous work. Age, cognitive reflection, and political knowledge are all associated with sharing less false content, but also less true content, resulting in negative (in the case of age and political knowledge) or null (in the case of cognitive reflection) associations with sharing discernment. Unsurprisingly, people who engage with politics online are more likely to share content of all kinds—but especially true content–leading to a positive association with discernment. Trolling and cynicism are associated with sharing more false content, though surprisingly also more true content, resulting in a small negative (trolling) or no (cynicism) association with discernment. It is worth noting that, as in the case of polarization, the relationship between each of these non-polarization factors and sharing is directionally the same for true and false news, indicating that these factors are related to news sharing as a whole and not fake news sharing in particular.

Discussion

Addressing the growing misinformation problem requires understanding why people spread fake news online. Some recent research suggests that polarization is the main driver of sharing fake news. We provide a rigorous test of this explanation using a research strategy that addresses two design concerns with previous work employing observational social media data: polarized individuals may share more fake news because they are exposed to more such news, or because they are inclined to share more news of *all* kinds – false and true alike.

Our findings are starkly inconsistent with the perspective that polarization underlies a psychology of particular susceptibility to sharing misinformation. We find no evidence that polarization is associated with sharing more false news than true news (i.e., lower sharing discernment). Indeed, in many cases, we find that polarization is associated with *increased* sharing discernment. Thus, polarization is often associated with normatively positive outcomes when it comes to the quality of news that people intend to share online. Importantly, we find the relationship between polarization and news sharing is directionally the same for real and fake news. That is, when polarization is associated with sharing more fake news, it is also associated with sharing more real news, and vice versa. Interestingly, even absent any consideration of discernment, we do not find consistent support for the claim that polarization is associated with sharing more fake news: negative out-party affect is associated with sharing more fake (and real) news, while positive in-party affect is associated with sharing *less* fake (and real) news.

Our findings have implications for how researchers and social media firms should seek to slow the spread of fake news online. First, our findings cast doubt on the suggestion that when informational treatments, such as fact-checks or digital literacy interventions, fail it is because the deficits those interventions seek to remedy do not address the root causes of misinformation sharing (Osmundsen et al. 2021). Further work on the development of such interventions is therefore a worthy pursuit. Second, our findings cast some doubt on recent claims that we can slow the spread of fake news by decreasing political polarization among the public (Osmundsen et al. 2021, Jenke 2022). In fact, given that we find polarization to be associated as much or more with sharing real news, efforts to depolarize the public would likely also result in a reduction of sharing true content *more than false content*. This finding, in conjunction with the fact that real news is far more prevalent on social media than fake news (Grinberg et al. 2019, Guess et al. 2020), means that efforts to depolarize the public might also dramatically reduce the amount of real political content users are exposed to (Guay et al. 2023). While depolarization is surely a worthy goal to mitigate rising levels of political hostility, violence, and distrust, it is not clear that such efforts would necessarily improve the veracity of the information environment on social media.

Of course, our study is not without limitations. The most significant limitation is shared with all studies aimed at understanding why people share fake news, namely the inability to draw causal inferences. While we and others assume that polarization is causally prior to intentions to share news, it may be that sharing news increases polarization, especially when that news is fake or misleading. Future work should leverage the growing number of successful interventions that reduce political polarization (Voelkel et al. 2023) to test the effect of polarization on news sharing behavior. Ideally, this work would also experimentally manipulate factors like information deficits and attention so that various explanations can be compared simultaneously in a causal framework.

In addition, our approach faces a common generalizability constraint inherent to the growing body of research that exposes participants to misinformation within the context of a survey rather than measuring their actual interactions with misinformation on social media. However, in this case we believe the gains in internal validity from controlling exposure far outweigh the gains in external validity from using observational social media data. Without controlling exposure, it is impossible to accurately measure discernment nor rule out the alternative explanation that polarized individuals share more fake news because they are exposed to more of it.

Despite these limitations, we have good reason to have confidence in the generalizability of our findings. First, recent work finds that the type of self-reported sharing intentions used in this study are correlated with actual sharing activity on social media (Mosleh et al. 2020). Second, our re-analysis of more generalizable observational social media data showed a similar pattern of findings to ours. While a lack of data on what content users are exposed to prevents us from adequately measuring discernment using observational social media data, we re-analyzed data from Osmundsen et al. (2021) using a proxy for discernment. We found that polarization is associated with sharing more real news and fake news, but that this relationship is approximately seven times stronger for real news.

References

Badrinathan, Sumitra. “Educative Interventions to Combat Misinformation: Evidence from a Field Experiment in India.” *American Political Science Review* 115, no. 4 (2021): 1325–41.

Barberá, Pablo. “Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data.” *Political Analysis* 23, no. 1 (2015): 76–91.

Clayton, Katherine, Spencer Blair, Jonathan A Busam, Samuel Forstner, John Glance, Guy Green, Anna Kawata, et al. “Real Solutions for Fake News? Measuring the Effectiveness of General Warnings and Fact-Check Tags in Reducing Belief in False Stories on Social Media.” *Political Behavior* 42 (2020): 1073–95.

Garrett, R Kelly, Shira Dvir Gvirsman, Benjamin K Johnson, Yariv Tsfati, Rachel Neo, and Aysenur Dal. “Implications of Pro-and Counterattitudinal Information Exposure for Affective Polarization.” *Human Communication Research* 40, no. 3 (2014): 309–32.

Graham, Matthew H, and Omer Yair. “Expressive Responding and Trump’s Big Lie.” *Political Behavior*, 2023, 1–26.

Greifer, N, S Worthington, S Iacus, and G King. “Clarify: Simulation-Based Inference for Regression Models.” *R Package Version* 3, no. 0 (2023).

Grinberg, Nir, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson, and David Lazer. “Fake News on Twitter during the 2016 US Presidential Election.” *Science* 363, no. 6425 (2019): 374–78.

Guay, Brian, Gordon Pennycook, David Rand, and others. “Examining Partisan Asymmetries in Fake News Sharing and the Efficacy of Accuracy Prompt Interventions,” 2022.

Guess, Andrew M., and Kevin Munger. "Digital literacy and online political behavior." Political science research and methods 11, no. 1 (2023): 110-128.

Guess, Andrew M, Michael Lerner, Benjamin Lyons, Jacob M Montgomery, Brendan Nyhan, Jason Reifler, and Neelanjan Sircar. “A Digital Media Literacy Intervention Increases Discernment between Mainstream and False News in the United States and India.” *Proceedings of the National Academy of Sciences* 117, no. 27 (2020): 15536–45.

Iyengar, Shanto, Gaurav Sood, and Yphtach Lelkes. “Affect, Not Ideology: A Social Identity Perspective on Polarization.” *Public Opinion Quarterly* 76, no. 3 (2012): 405–31.

Jenke, Libby. “Affective Polarization and Misinformation Belief.” *Political Behavior*, 2023, 1–60.

Kahan, Dan M. “Ideology, Motivated Reasoning, and Cognitive Reflection.” *Judgment and Decision Making* 8, no. 4 (2013): 407–24.

King, Gary, Michael Tomz, and Jason Wittenberg. “Making the Most of Statistical Analyses: Improving Interpretation and Presentation.” *American Journal of Political Science*, 2000, 347–61.

Kunda, Ziva. “The Case for Motivated Reasoning.” *Psychological Bulletin* 108, no. 3 (1990): 480.

Lazer, David. “Studying Human Attention on the Internet.” *Proceedings of the National Academy of Sciences* 117, no. 1 (2020): 21–22.

Lelkes, Yphtach, and Sean J Westwood. “The Limits of Partisan Prejudice.” *The Journal of Politics* 79, no. 2 (2017): 485–501.

Loomba, Sahil, Alexandre de Figueiredo, Simon J Piatek, Kristen de Graaf, and Heidi J Larson. “Measuring the Impact of COVID-19 Vaccine Misinformation on Vaccination Intent in the UK and USA.” *Nature Human Behaviour* 5, no. 3 (2021): 337–48.

Marie, Antoine, Sacha Altay, and Brent Strickland. “Moralization and Extremism Robustly Amplify Myside Sharing.” *PNAS Nexus* 2, no. 4 (2023): pgad078.

Mosleh, Mohsen, Gordon Pennycook, and David G Rand. “Self-Reported Willingness to Share Political News Articles in Online Surveys Correlates with Actual Sharing on Twitter.” *Plos One* 15, no. 2 (2020): e0228882.

Mosleh, Mohsen, and David G Rand. “Measuring Exposure to Misinformation from Political Elites on Twitter.” *Nature Communications* 13, no. 1 (2022): 7144.

Osmundsen, Mathias, Alexander Bor, Peter Bjerregaard Vahlstrup, Anja Bechmann, and Michael Bang Petersen. “Partisan Polarization Is the Primary Psychological Motivation behind Political Fake News Sharing on Twitter.” *American Political Science Review* 115, no. 3 (2021): 999–1015.

Pennycook, Gordon, Ziv Epstein, Mohsen Mosleh, Antonio A Arechar, Dean Eckles, and David G Rand. “Shifting Attention to Accuracy Can Reduce Misinformation Online.” *Nature* 592, no. 7855 (2021): 590–95.

Pennycook, Gordon, Jonathon McPhetres, Yunhao Zhang, Jackson G Lu, and David G Rand. “Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention.” *Psychological Science* 31, no. 7 (2020): 770–80.

Pennycook, Gordon, and David G Rand. “Lazy, Not Biased: Susceptibility to Partisan Fake News Is Better Explained by Lack of Reasoning than by Motivated Reasoning.” *Cognition* 188 (2019): 39–50.

Pereira, Andrea, and Jay Van Bavel. “Identity Concerns Drive Belief in Fake News.” *Unpublished Manuscript: Https://Psyarxiv. Com/7vc5d*, 2018.

Sirlin, Nathaniel, Ziv Epstein, Antonio A Arechar, and David G Rand. “Digital Literacy Is Associated with More Discerning Accuracy Judgments but Not Sharing Intentions,” 2021.

Taber, Charles S, and Milton Lodge. “Motivated Skepticism in the Evaluation of Political Beliefs.” *American Journal of Political Science* 50, no. 3 (2006): 755–69.

Trilling, Damian, Juhi Kulshrestha, Claes De Vreese, Denis Halagiera, Jakub Jakubowski, Judith Möller, Cornelius Puschmann, Agnieszka Stępińska, Sebastian Stier, and Cristian Vaccari. “Is Sharing Just a Function of Viewing? The Sharing of Political and Non-Political News on Facebook.” *Journal of Quantitative Description: Digital Media* 2 (2022).

Van der Linden, Sander, Anthony Leiserowitz, Seth Rosenthal, and Edward Maibach. “Inoculating the Public against Misinformation about Climate Change.” *Global Challenges* 1, no. 2 (2017): 1600008.

Supplementary Materials

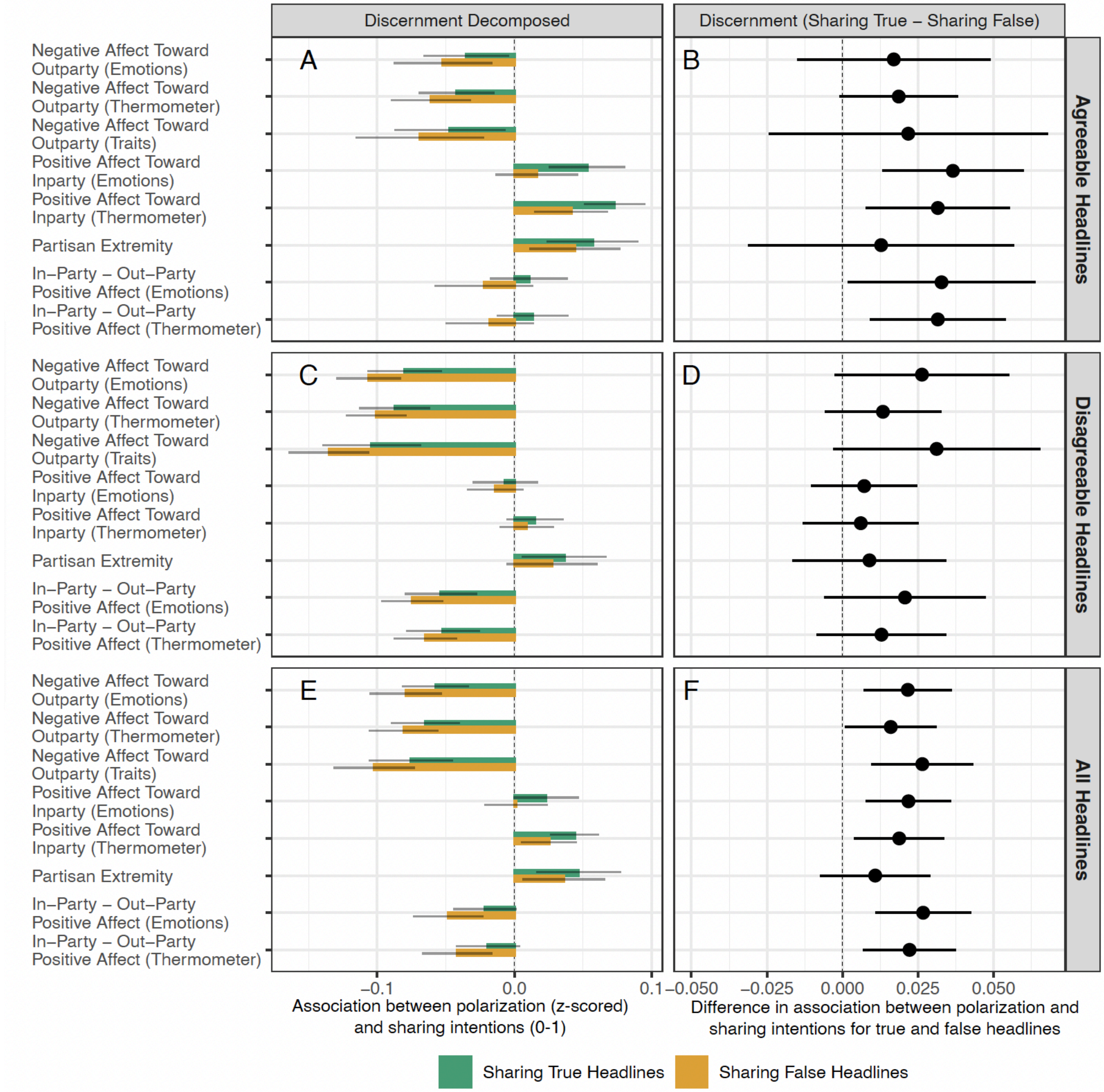
Table of Contents

1. Robustness Checks

1.1 Main Analyses of with Low Attention Respondents Included

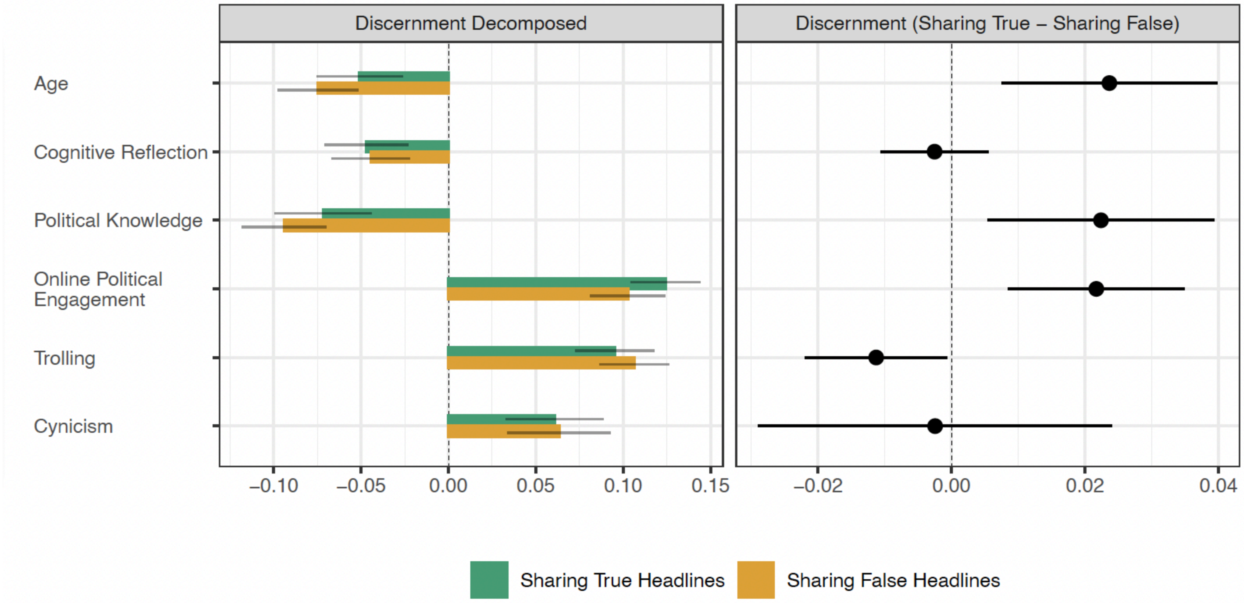
We pre-registered our decision to exclude respondents who engaged in straight lining on survey grid questions—i.e., selecting identical response options to respond to very different questions in a way that is highly implausible for attentive respondents. The three affective polarization grids were used to identify straight-liners in our survey, as each contains both positive and negative items describing each party. We also pre-registered our decision to include these inattentive respondents (N = 62) in a supplemental analysis.

Polarization and Sharing Discernment, By Headline Agreeableness (Low Attention Respondents Included)



Panels A, C, & E: Association between each polarization measure (z-scored) and news sharing intentions (measured on a 0-1 scale), separately for agreeable (A) and disagreeable (C) headlines, and for all headlines (E). Panels B, D, & F: The association between polarization and sharing discernment, i.e. how much more respondents share true news than false news. Each value in Panels B, D, & F is equivalent to the difference between the corresponding values in Panels A, C, & E. For example, the difference between the green and yellow bars in the first row of Panel A (-0.033 - -0.054) is approximately equivalent to the parameter estimate reported in the first row of Panel B (0.021). All models use survey weights and control for gender, income, education, race, and political interest. Horizontal lines represent 95% confidence intervals.

Non-Polarization Predictors of Sharing Discernment (Low Attention Respondents Included)

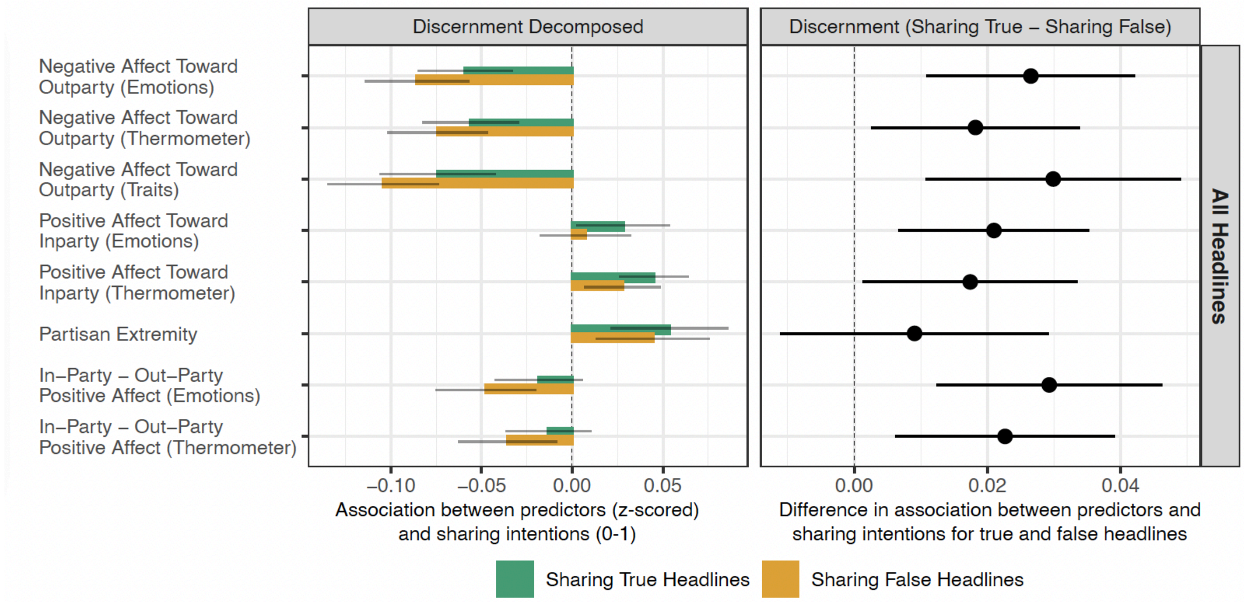


Left panel: Association between each polarization measure (z-scored) and news sharing intentions (measured on a 0-1 scale). Right panel: The association between polarization and sharing discernment, i.e. how much more respondents share true news than false news. All models use survey weights and control for gender, income, education, race, and political interest. Horizontal lines represent 95% confidence intervals.

1.2 Main Analyses with Only Social Media Users

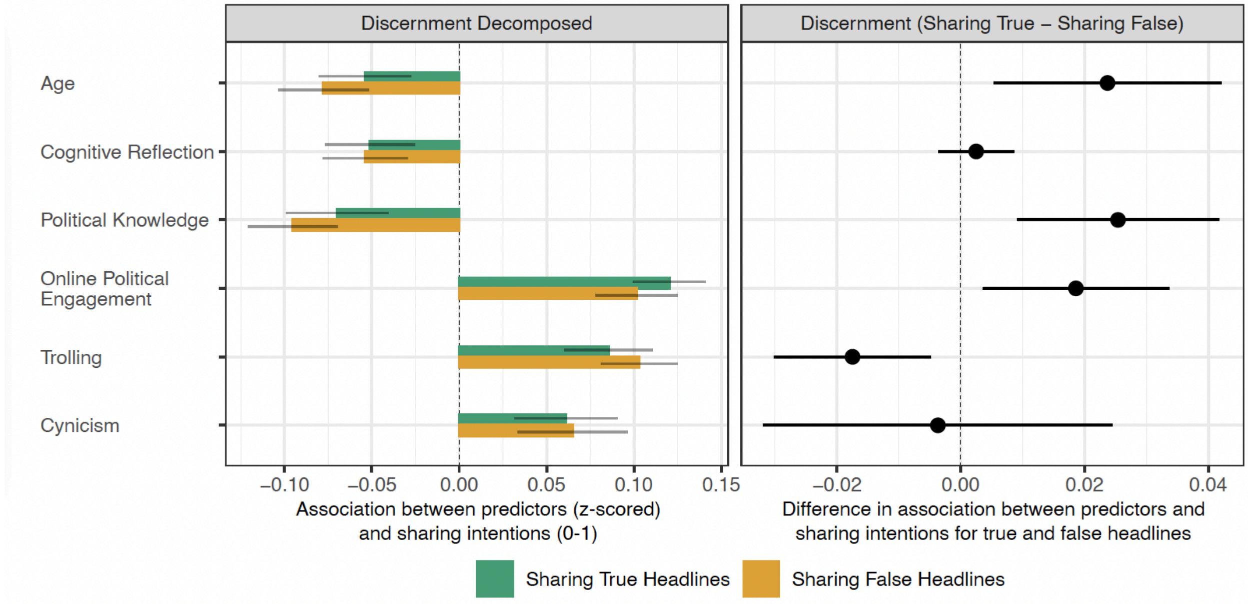
Respondents reported whether they used any social media websites (Facebook, Twitter, Instagram, TikTok, Snapchat, WhatsApp, Reddit, Other, None of these). 848 respondents reported using at least one social media website and 152 reported using none of them. In the main analysis we include respondents regardless of whether they use social media. Here we replicate our main analysis with respondents who reported using social media. The results are nearly identical to the main analysis conducted with the full sample.

Polarization and Sharing Discernment (Social Media Users)



Left panel: Association between each polarization measure (z-scored) and news sharing intentions (measured on a 0-1 scale). Right panel: The association between polarization and sharing discernment, i.e. how much more respondents share true news than false news. All models use survey weights and control for gender, income, education, race, and political interest. Horizontal lines represent 95% confidence intervals.

Non-Polarization Predictors of Sharing Discernment (Social Media Users)

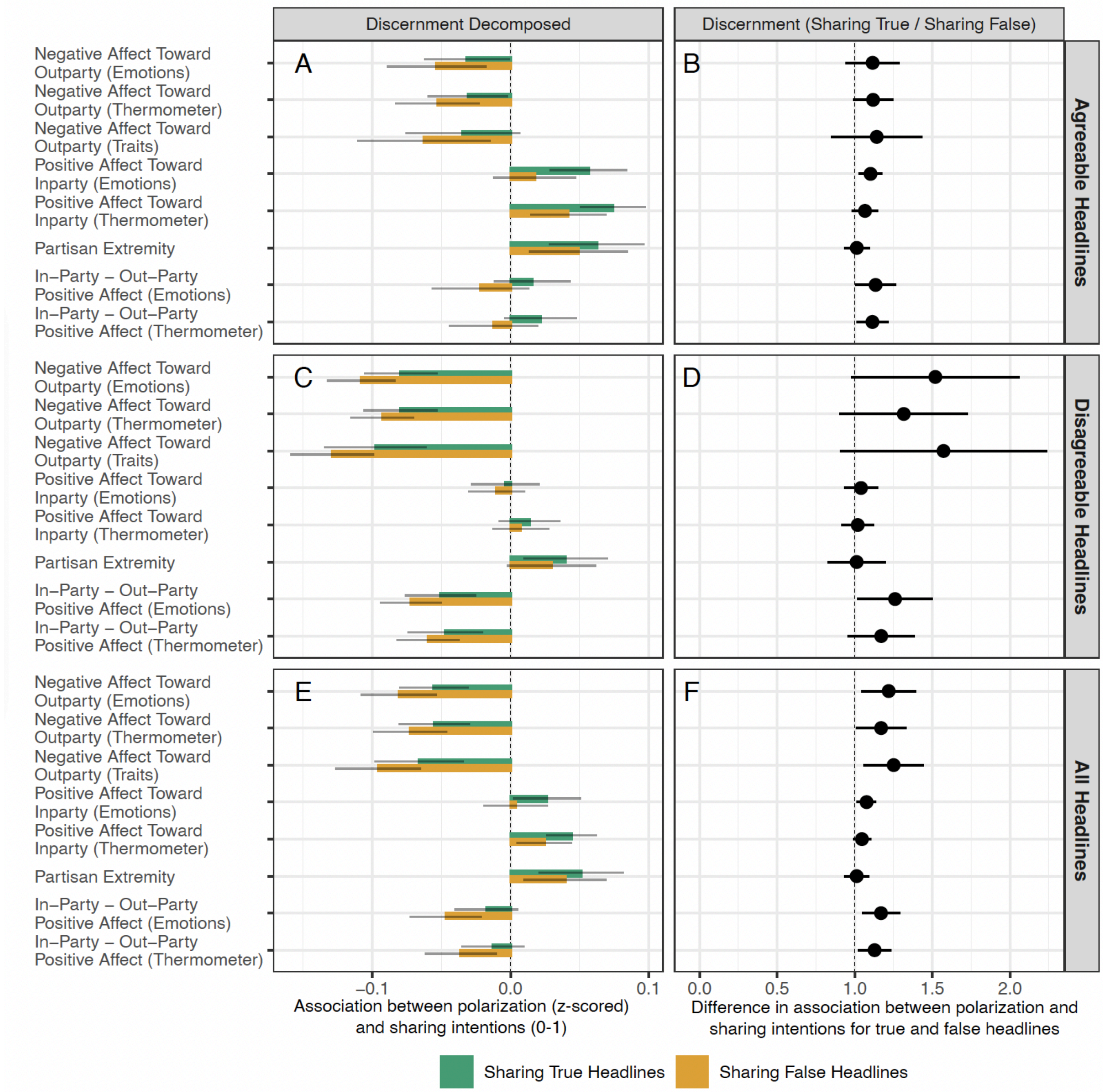


Left panel: Association between each predictor (z-scored) and news sharing intentions (measured on a 0-1 scale). Right panel: The association between each predictor and sharing discernment, i.e. how much more respondents share true news than false news. All models use survey weights and control for gender, income, education, race, and political interest. Horizontal lines represent 95% confidence intervals.

1.3 Main Analysis with Multiplicative Discernment

We use sharing discernment as our primary outcome of interest, which reflects the difference in sharing true versus false news. Guay et al. 2023 outline two types of discernment, based on different methods of calculating this difference. *Additive discernment* reflects the additive difference between sharing true and false news--e.g., if respondents share an average of 10 true and 6 false news articles, additive discernment is 4 (10 - 6 = 4). *Multiplicative discernment* reflects the multiplicative difference between sharing true and false news-–in the same example, respondents are 1.7 times more likely to share true news than false news (10 / 6 = 1.7). Since multiplicative discernment is reported in the form of a ratio, 1 indicates no discernment (people are just as likely to share true news as false news), values greater than 1 indicate positive discernment (sharing more true news than false news), and values less than 1 indicate negative discernment (sharing more false news than true news). We present the additive discernment models in the manuscript and present the multiplicative discernment models here.

Polarization and Sharing Discernment, By Headline Agreeableness (Multiplicative Discernment)



Panels A, C, & E: Association between each polarization measure (z-scored) and news sharing intentions (measured on a 0-1 scale), separately for agreeable (A) and disagreeable (C) headlines, and for all headlines (E). Panels B, D, & F: The association between polarization and sharing discernment, i.e. how much more respondents share true news than false news. Each value in Panels B, D, & F is equivalent to the ratio of the corresponding values in Panels A, C, & E. All models use survey weights and control for gender, income, education, race, and political interest. Horizontal lines represent 95% confidence intervals.

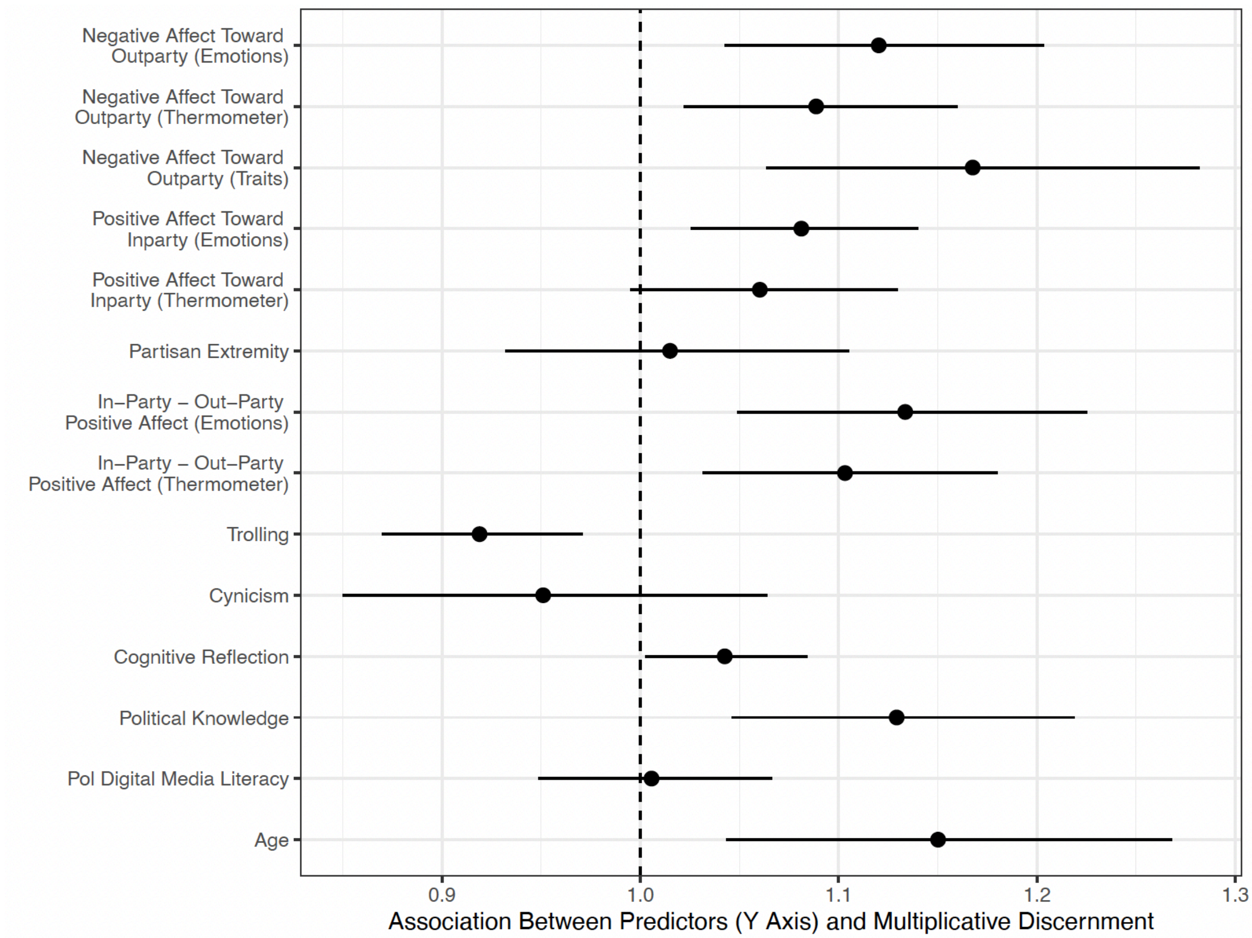
Non-Polarization Predictors of Sharing Discernment (Multiplicative Discernment)

Left panel: Association between each predictor (z-scored) and news sharing intentions (measured on a 0-1 scale). Right panel: The association between each predictor and sharing discernment, i.e., the ratio of shared true news to shared false news. All models use survey weights and control for gender, income, education, race, and political interest. Horizontal lines represent 95% confidence intervals.

1.4 Alternative Approach to Modeling Multiplicative Discernment: Deviation from Pre-Analysis Plan

There are two ways to model multiplicative discernment using non-binary outcomes. The first is using quasi-poisson models. Because the quasi-poisson model uses a log-link function, exponentiated interaction terms are interpreted as ratios. In our case, the ratio of shared true news to shared false news. However, the use of quasi-poisson models outside of modeling overdispersed count variables is rare. Therefore, as described in the manuscript, we adopt a more straightforward approach: modeling sharing intentions with an interaction between polarization and headline veracity, then using simulation-based inference to calculate the ratio that the exponentiated interaction term from the quasi-poisson model provides. The two approaches yield nearly identical point estimates and standard errors to what two figures immediately above. Because we pre-registered the former approach, we include this analysis below.

Associations with Multiplicative Discernment (From Quasipoisson Models)

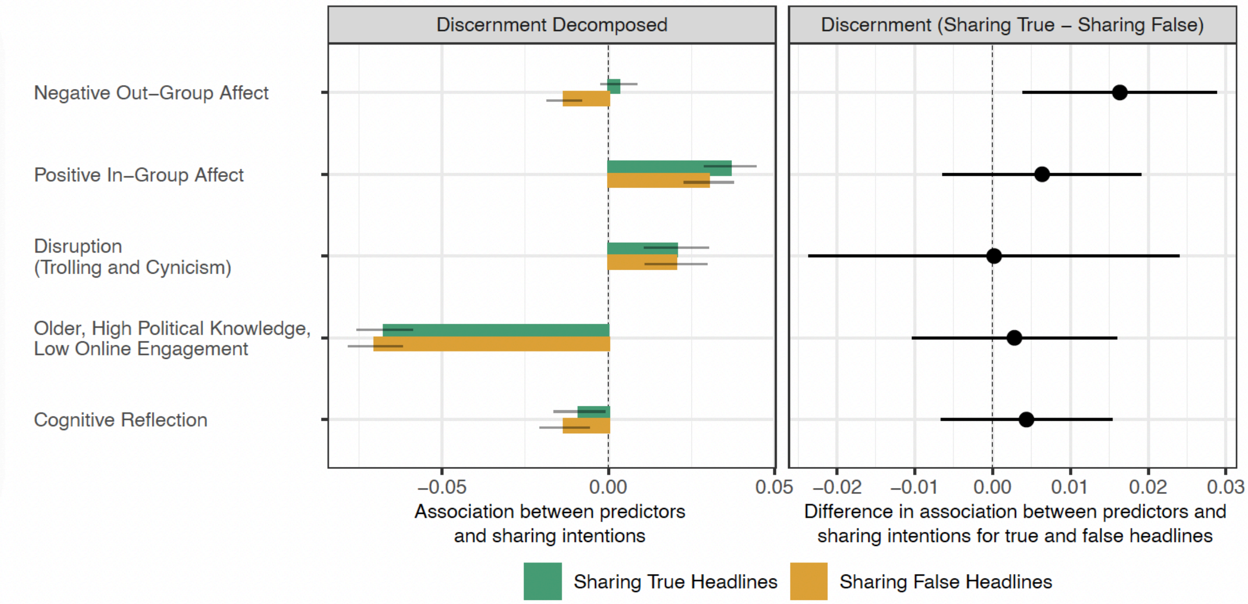


The association between polarization and sharing discernment (multiplicative), i.e. how much more respondents share true news than false news. All models use survey weights and control for gender, income, education, race, and political interest. Horizontal lines represent 95% confidence intervals.

1.5 Principal Components Analysis

In the main analysis, separate models are run for each of the polarization and non-polarization predictors of sharing. While each model controls for gender, income, education, race, and political interest, the models do not control for the other predictors of sharing. Here we model sharing intentions with all of these predictors, as well as the standard set of control variables. Rather than entering all of the individual variables into one model, we ran a principal components analysis (PCA) to reduce the number of model inputs. We conducted a Horn's Parallel Analysis of Principal Components using the *Paran* package in R to determine the number of components to retain. We ran separate principal components analysis for the main affective polarization variables (positive in-party affect and negative out-party affect using the emotions and feeling thermometer measures, and negative out-party affect using the traits measure, disruption variables (trolling and cynicism), and ignorance variables (age, cognitive reflection, political knowledge, and online political engagement). We retained two components for affective polarization, one component for disruption, and two components for ignorance. We then modeled sharing intentions with all of these components as predictors, as well as the standard set of control variables.

PCA Analysis



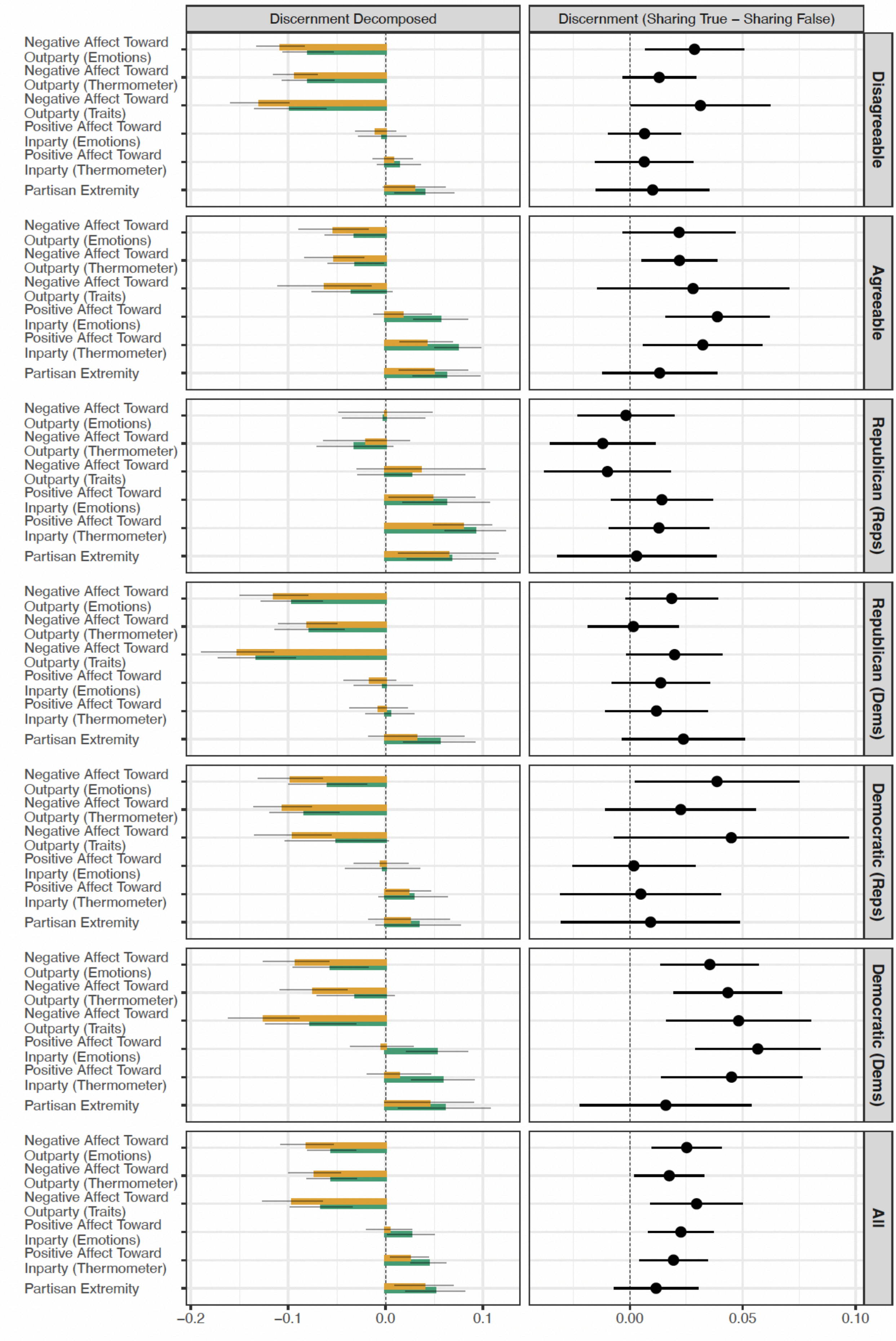
The associations between the components and sharing true and false content are similar to those presented in the main analysis, except that while there remains a significant negative association between negative out-party affect and sharing false content, there is no significant association between negative out-party affect and sharing true content. There remains a significant positive association between negative out-party affect and sharing discernment. There is no significant association between the other components and sharing discernment.

1.6 Alternative Approach to Modeling Osmundsen et al.'s Data

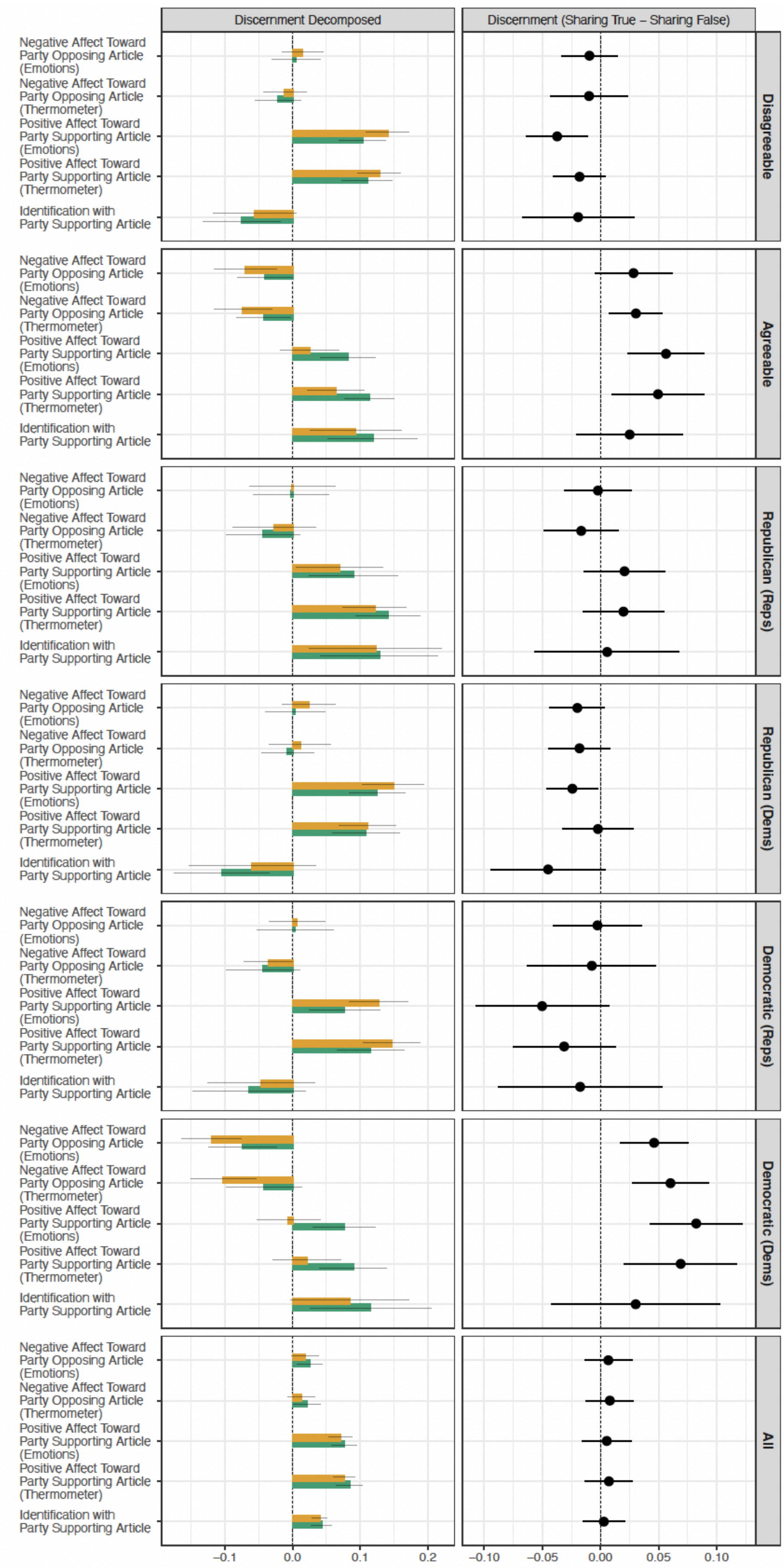
We will also run these models with a non-binary dependent variable: modeling the proportion of shared content that is false rather than a dummy for whether a respondent shared any false content at all.

1.7 Robustness Check Analysis with Discernment Decomposition

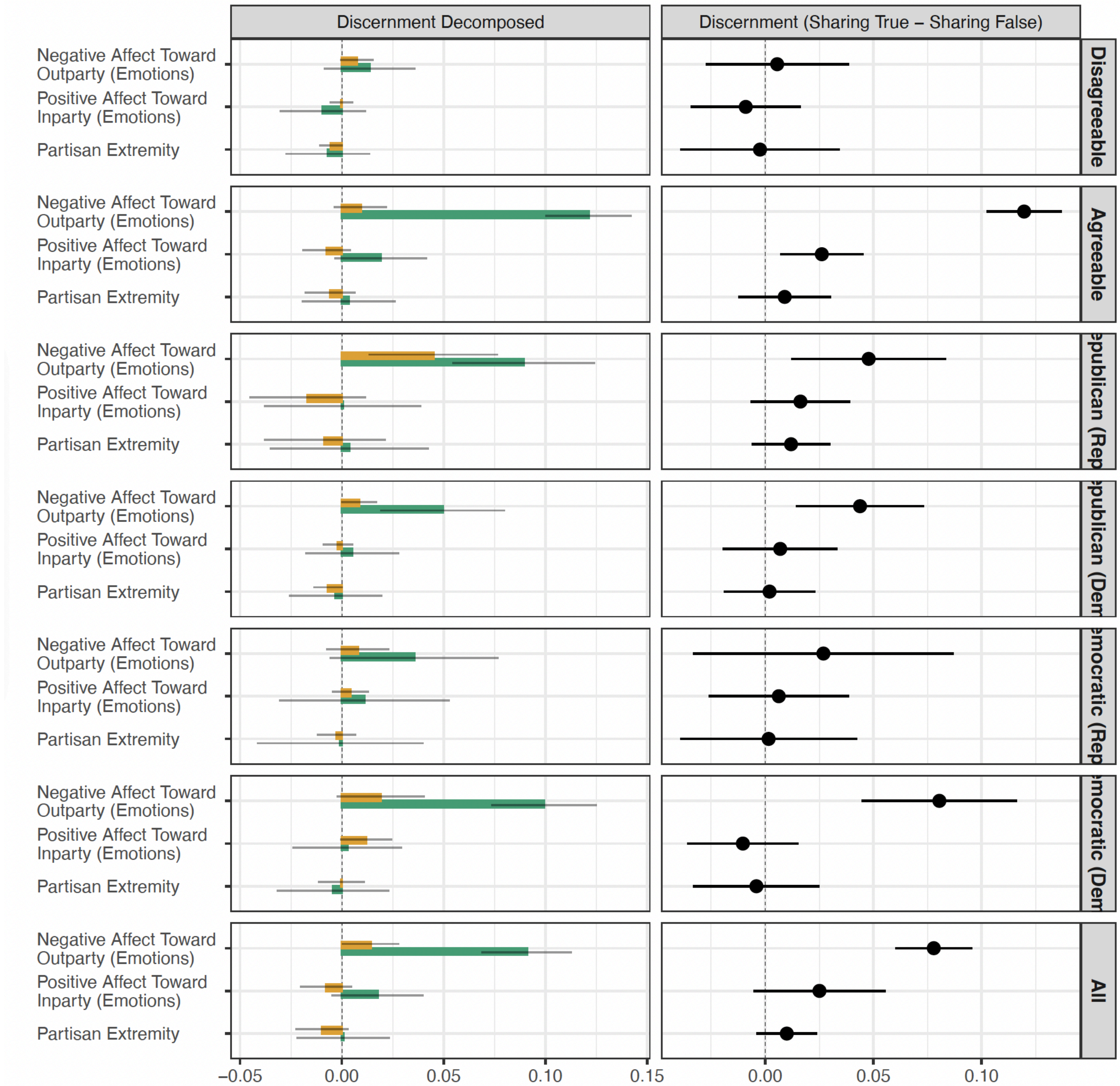
Original YouGov Survey Data, Polarization Operationalized w/r/t Respondent


Discernment Decomposed
Discernment (Sharing True – Sharing False)
Negative Affect Toward Outparty (Emotions)
Negative Affect Toward Outparty (Thermometer)
Negative Affect Toward Outparty (Traits)
Positive Affect Toward Inparty (Emotions)
Positive Affect Toward Inparty (Thermometer)
Partisan Extremity
Disagreeable
Agreeable
Republican (Reps)
Republican (Dems)
Democratic (Reps)
Democratic (Dems)
All
−0.2
−0.1
0.0
0.1
0.00
0.05
0.10

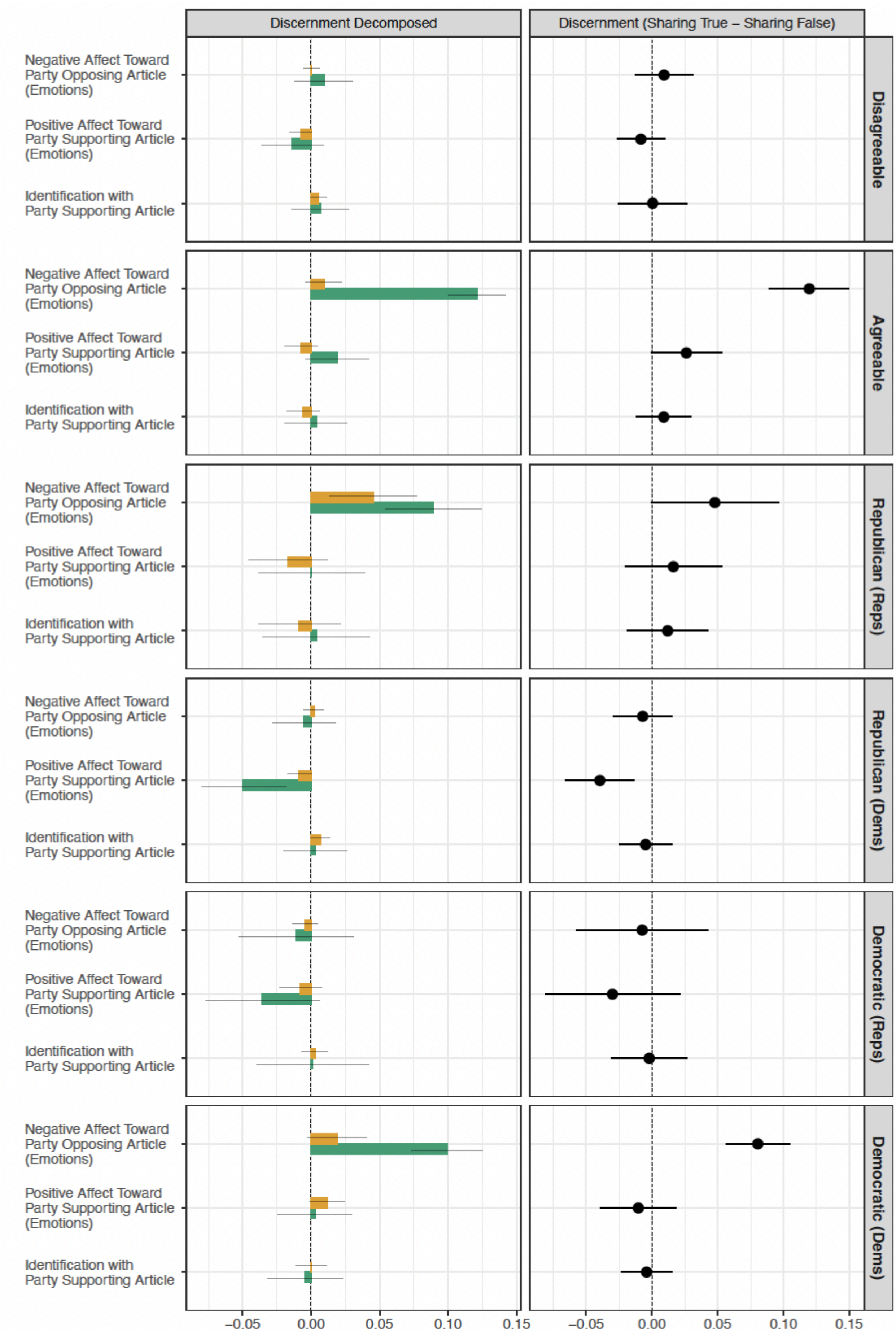
Original YouGov Survey Data, Polarization Operationalized w/r/t Article


Discernment Decomposed
Discernment (Sharing True – Sharing False)
Negative Affect Toward Party Opposing Article (Emotions)
Negative Affect Toward Party Opposing Article (Thermometer)
Positive Affect Toward Party Supporting Article (Emotions)
Positive Affect Toward Party Supporting Article (Thermometer)
Identification with Party Supporting Article
Disagreeable
Agreeable
Republican (Reps)
Republican (Dems)
Democratic (Reps)
Democratic (Dems)
All
−0.1
0.0
0.1
0.2
−0.10
−0.05
0.00
0.05
0.10

Osmundsen et al. Data, Polarization Operationalized w/r/t Respondent



Osmundsen et al. Data, Polarization Operationalized w/r/t Article


Discernment Decomposed
Discernment (Sharing True – Sharing False)
Negative Affect Toward Party Opposing Article (Emotions)
Positive Affect Toward Party Supporting Article (Emotions)
Identification with Party Supporting Article
Disagreeable
Agreeable
Republican (Reps)
Republican (Dems)
Democratic (Reps)
Democratic (Dems)
−0.05
0.00
0.05
0.10
0.15

2. Survey Materials
2.1 Polarization Predictors of Sharing
Emotions

- What do you feel when you think about Democrats?
    - o Grid with the following rows: Angry, Frustrated, Afraid, Hopeful, Enthusiastic, Proud, and columns: 1 (Not at all), 2, 3, 4, 5, 6, 7 (Very strongly)
- What do you feel when you think about Republicans?
    - o Grid with the following rows: Angry, Frustrated, Afraid, Hopeful, Enthusiastic, Proud, and columns: 1 (Not at all), 2, 3, 4, 5, 6, 7 (Very strongly)

What do you feel when you think about [Democrats \ Republicans].?

	1 (Not at all)	2	3	4	5	6	7 (Very strongly)
Angry							
Frustrated							
Afraid							
Hopeful							
Enthusiastic							
Proud							

- Traits (only asked about out-party)
    - o Now we'd like to know more about what you think about Democrats[Republicans]. Below, we've given a list of words that some people might use to describe them. For each item, please indicate how well you think it applies to Democrats[Republicans]. Grid with the following rows: Patriotic, Intelligent, Honest, Open-minded, Generous, Hypocritical, Selfish, Mean, and columns: Not at all well, Not too well, Somewhat well, Very well, Extremely well

Now we'd like to know more about what you think about [Democrats/Republicans]. Below, we've given a list of words that some people might use to describe them. For each item, please indicate how well you think it applies to [Democrats/Republicans].

	Not at all well	Not too well	Somewhat well	Very well	Extremely well
Patriotic					
Intelligent					
Honest					
Open-minded					

Generous					
Hypocritical					
Selfish					
Mean					

- Feeling Thermometers
    - o We'd like you to rate how you feel towards Democrats and Republicans on a scale of 0 to 100, which we call a "feeling thermometer." On this feeling thermometer, ratings between 0 and 49 degrees mean that you feel unfavorable and cold (with 0 being the most unfavorable/coldest). Ratings between 51 and 100 degrees mean that you feel favorable and warm (with 100 being the most favorable/warmest). A rating of 50 means you have no feelings on way or the other.

      How would you rate your feeling toward Democrats?
      (horizontal slider with labels "Very unfavorable/cold" and "Very favorable/warm" on the left and right ends, respectively. The slider was set to zero at default and respondents were required to move it. A number indicating the slider's position (e.g., 91) appeared underneath as it moved.

2.2 Non-Polarization Predictors of Sharing

- Political Knowledge:
    - o Whose responsibility is it to decide if a law is constitutional or not?: Congress, The President, The Supreme Court
    - o Whose responsibility is it to nominate judges to Federal Courts?: Congress, The Supreme Court, The President
    - o Do you know what job or political office is currently held by Nancy Pelosi? Is it: Justice of the Supreme Court, Secretary of the Treasury, Governor of California, Senate Majority Leader, Speaker of the House
    - o Do you know what job or political office is currently held by Janet Yellen? Is it: Secretary of the Treasury, Secretary of State, House Republican Leader, Attorney General, Justice of the Supreme Court
- Cognitive Reflection
    - o If you're running a race and you pass the person in second place, what place are you in?
    - o A farmer had 15 sheep and all but 8 died. How many are left?
    - o Emily's father has three daughters. The first two are named April and May. What is the third daughter's name?
    - o How many cubic feet of dirt are there in a hole that is 3' deep x 3' wide x 3' long?
- Online Political Engagement
    - o How often do you do the following?

      Grid rows: Start or join a political group or group supporting a cause on a social networking site, Post online personal views related to politics or campaigning, Share someone else's political post to other people online, Email a national, state, or local government official about an issue of personal importance.

Grid columns: 1 (Never), 2, 3, 4, 5, 6, 7 (Several times a day)

- Trolling:
  - I like to troll people in forums or the comments section of websites.
  - I have sent people to shock websites for the fun of it.(Strongly agree, somewhat agree, slightly agree, neither agree nor disagree, Slightly disagree, Somewhat disagree, Strongly disagree)
- Cynicism
  - How much do you agree or disagree with the following statements.

    Grid rows: They are using our tax money well in Washington, Democracy in the U.S. functions well, Washington is perfectly able to solve problems in our society, Politics in the United States considers the interest of the people.

    Grid columns: 1 (Strongly Disagree), 2, 3, 4, 5, 6, 7 (Strongly Agree)

2.3 Headline Rating Task

The instructions for the headline rating task read: Next you will be presented with a set of news headlines (20 in total). We are interested in how likely you would be to share these stories online.

After seeing each headline, respondents were asked: “If you were to see the above post online, how likely would you be to share it?” with the following response options: Extremely unlikely, Moderately unlikely, Slightly unlikely, Slightly likely, Moderately likely, Extremely likely

The headlines appeared as they would on a social media website, accompanied by an image and the name of the source. The headlines are available on Open Science Framework, sorted by partisan-lean (pro Republican, pro Democratic) and veracity (true, false).

2.4 Survey Sample

The following text is from the YouGov codebook:

YouGov interviewed 1174 respondents who were then matched down to a sample of 1000 to produce the final dataset. The interviews were conducted between November 18, 2022 - November 28, 2022. The respondents were matched to a sampling frame on gender, age, race, and education. The frame was constructed by using a politically representative "modeled frame" of US adults, based upon the American Community Survey (ACS) public use microdata file, public voter file records, the 2020 Current Population Survey (CPS) Voting and Registration supplements, the 2020 National Election Pool (NEP) exit poll, and the 2020 CES surveys, including demographics and 2020 presidential vote. The matched cases were weighted to the sampling frame using propensity scores. The matched cases and the frame were combined and a logistic regression was estimated for inclusion in the frame. The propensity score function included age, gender, race/ethnicity, years of education, presidential vote 2020, own or rent, and region. The propensity scores were grouped into deciles of the estimated propensity score in the frame and post-stratified according to these deciles. The weights were then post-stratified on 2020 Presidential vote choice, and a three-way stratification of gender, age (4-categories), and education (4-categories), a three-way stratification of gender, age (4-categories), and race

(4-categories), and a two-way stratification of race (4-categories), and education (4-categories) to produce the final weight.