

Accuracy and Social Motivations Shape Judgements of (Mis)Information

Steve Rathje<sup>1</sup>  
Jon Roozenbeek<sup>1</sup>  
Jay J. Van Bavel<sup>2</sup>  
Sander van der Linden<sup>1</sup>

<sup>1</sup>Department of Psychology, University of Cambridge, Cambridge CB2 3RQ, United Kingdom

<sup>2</sup>Department of Psychology & Center for Neural Science, New York University, New York, NY 10003

**Author Note**

Steve Rathje: <https://orcid.org/0000-0001-6727-571X>

Jay Van Bavel: <https://orcid.org/0000-0002-2520-0442>

Jon Roozenbeek: <https://orcid.org/0000-0002-8150-9305>

Sander van der Linden: <https://orcid.org/0000-0002-0269-1744>

**Acknowledgments:** We are grateful for support from a Gates Cambridge Scholarship awarded to SR (Grant #OPP1144), a British Academy Postdoctoral Fellowship awarded to JR (#PF21\210010), a John Templeton Foundation Grant (#61378) awarded to JVB and an Infodemic grant awarded to SVL (UK Government, #SCH-00001-3391).

Correspondence concerning this article should be addressed to Steve Rathje (sjr210@cam.ac.uk) and Sander van der Linden ([sander.vanderlinden@psychol.cam.ac.uk](mailto:sander.vanderlinden@psychol.cam.ac.uk))

## ACCURACY INCENTIVES SHAPE BELIEF IN (MIS)INFORMATION

### **Abstract:**

The extent to which belief in (mis)information reflects lack of knowledge versus a lack of motivation to be accurate is unclear. Across four experiments ( $n = 3,364$ ), we motivated US participants to be accurate by providing financial incentives for correct responses about the veracity of true and false political news headlines. This incentive improved accuracy and reduced partisan bias in judgements of headlines by about 30%, primarily by increasing the perceived accuracy of true news from the opposing party ( $d = 0.47$ ). Incentivizing people to identify news that would be liked by their political allies, however, decreased accuracy. Replicating prior work, conservatives were less accurate at discerning true from false headlines than liberals, yet incentives closed the gap in accuracy between conservatives and liberals by 52%. A non-financial accuracy motivation intervention was also effective, suggesting that motivation-based interventions are scalable. Altogether, these results suggest that judgments of (mis)information in part reflect motivational factors.

**Keywords:** Misinformation, Motivated Reasoning, Incentives, News, Social Identity

### Main

Misinformation – which can refer to fabricated news stories, false rumors, conspiracy theories, or disinformation campaigns – can have serious negative effects on society and democracy<sup>1,2</sup>. Numerous studies suggest that misinformation exposure<sup>1</sup> may reduce support for climate change<sup>5,6</sup> and the COVID-19 vaccine<sup>7,8</sup>, and that the mere repetition of misinformation can increase belief in it<sup>9,10</sup>. Anti-vaccination viewpoints are becoming increasingly popular online<sup>11</sup>, and there is widespread belief in misinformation and conspiracy theories about election fraud<sup>12</sup> and COVID-19<sup>13</sup>. There has thus been a growing interest in understanding the psychology of belief in misinformation and how to mitigate its spread<sup>1,2,14–16</sup>.

There are substantial partisan differences in how people judge information to be true or false. People are much more likely to believe news with politically-congruent content<sup>17–20</sup> or news that comes from politically-congruent sources.<sup>21,22</sup> However, there are multiple possible reasons that can explain why this partisan divide exists. One possible explanation is that people tend to engage in politically-motivated cognition<sup>23,24</sup>: although people are often motivated to be accurate, they also have social goals (e.g., group belonging, status, etc.) for holding certain beliefs that can interfere with accuracy goals<sup>17</sup>. Another potential explanation is that partisans have different pre-existing knowledge, or different prior beliefs, as a result of exposure to different partisan news outlets and social media feeds<sup>14</sup>. It is challenging to differentiate between these explanations unless accuracy or social motivations are experimentally manipulated<sup>25–28</sup>.

Several studies have also found that US conservatives tend to believe in and share more misinformation than US liberals<sup>29–35</sup>. One interpretation behind this asymmetry is that US conservatives are exposed to more low-quality information and thus have less accurate political knowledge, perhaps due to US conservative politicians and news media sources sharing less accurate information<sup>36,37</sup>. Another interpretation again focuses on motivation, suggesting that US conservatives may, in some contexts, have greater motivations to believe ideologically or identity-consistent claims that could interfere with their motivation to be accurate<sup>38,39</sup>. Yet, it is difficult to disentangle the causal role of motivation versus prior knowledge without experimentally manipulating motivations.

We examine the causal role of accuracy motives in shaping judgements of true and false political news via the provision of financial incentives for correctly identifying accurate headlines. Prior research about the effect of financial incentives for accuracy has yielded mixed results. For example, previous studies have found that financial incentives to be accurate can reduce partisan bias about politicized issues<sup>40,41</sup> and news headlines,<sup>42</sup> and improve accuracy about scientific information<sup>43</sup>. However, another study found that incentives for accuracy can backfire, increasing belief in false news stories<sup>18</sup>. Incentives also do not eliminate people's tendency to view familiar statements<sup>44</sup> or positions for which they advocate<sup>45</sup> as more accurate, raising questions as to whether incentives can override the heuristics people use to judge truth<sup>46</sup>. These conflicting results motivate the need for a systematic investigation of when and for whom various motivations influence belief.

We also examine whether social identity-based motivations to identify posts that will be liked by one's political in-group interfere with accuracy motivations. On social media, content that appeals to social-identity motivations, such as expressions of out-group derogation, tends to

---

<sup>1</sup> It should be noted that there are some null results regarding the effects of misinformation, such as a small correlational study finding that belief in COVID-19 misinformation was not associated with vaccine hesitancy<sup>3</sup>. In contrast, however, a larger-scale correlational study found that belief in COVID-19 misinformation was a robust negative predictor of intentions to engage in preventative health behavior<sup>4</sup>.

receive higher engagement online<sup>47</sup>. False news stories may be good at fulfilling these identity-based motivations, as false content is often negative about outgroup members<sup>29,48</sup>. The incentive structure of the social media environment draws attention to social motivations (e.g., receiving social approval in the form of likes and shares), which may lead people to give less weight to accuracy motivations online<sup>49,50</sup>.

Finally, we compare the effect of accuracy motivations to the effects of other factors that are regularly invoked to explain the belief and dissemination of misinformation, such as analytic thinking<sup>51</sup> political knowledge<sup>52</sup>, media literacy skills<sup>53</sup>, and affective polarization<sup>48</sup>. By including these variables in the same study, we are able to develop a more complete account of the factors that drive (mis)information belief and sharing<sup>2,15</sup>.

### Overview

Across four pre-registered experiments, including a replication with a nationally representative US sample, we test whether (A) incentives to be accurate improve people's ability to discern between true and false news and (B) reduce partisan bias (Experiment 1). Additionally, we test whether (C) increasing partisan identity motivations by paying people to correctly identify posts that appeal to one's in-group (mirroring the incentives of social media) reduces accuracy, even when paired with accuracy incentives (Experiment 2). Further, (D) we examine whether the effects of incentives are attenuated when partisan source cues are removed from posts (Experiment 3). Then, to test the generalizability of these results and help rule out alternate explanations, we test whether (E) increasing accuracy motivations through a non-financial accuracy motivation intervention also improves accuracy. Finally, in an integrative data analysis, we (F) examine whether motivation helps explain the gap in accuracy between conservatives and liberals, and (G) compare the effects of motivation to the effects of other variables known to predict misinformation susceptibility.

### Results

#### Experiment 1: Incentives Improve Accuracy and Reduce Bias

In Experiment 1, we recruited a politically-balanced sample of 462 US adults via the survey platform Prolific Academic.<sup>54</sup> Participants were shown 16 pre-tested news headlines with an accompanying picture and source (similar to how a news article preview would show up on someone's Facebook feed). In a pre-test, eight headlines (four false and four true) were rated as more accurate by Democrats than Republicans, and eight headlines (four false and four true) were rated as more accurate by Republicans than Democrats<sup>55</sup>. An example of a Democrat-leaning true headline was "Facebook removes Trump ads with symbols once used by Nazis" from *apnews.com*, and an example of a Democrat-leaning false news headline was "White House Chef Quits because Trump Has Only Eaten Fast Food For 6 Months" from *halfwaypost.com*. After seeing each headline, participants were asked "To the best of your knowledge, is the claim in the above headline accurate?" and were then asked "If you were to see the above article on social media, how likely would you be to share it?" See *Methods* for more details.

Half of the participants were randomly assigned to the *accuracy incentives* condition. In this condition, participants were told they would receive a small bonus payment of up to one US dollar based on how many correct answers they could provide regarding the accuracy of the articles. The other half of participants were assigned to a *control* condition in which they were asked the same questions about accuracy and sharing without any incentive to be accurate.

We first examined whether accuracy incentives improved truth discernment, or the number of true headlines participants rated as true minus the number of false headlines participants rated as true<sup>19</sup>. As predicted, participants in the *accuracy incentives* condition ( $M =$

## ACCURACY INCENTIVES SHAPE BELIEF IN (MIS)INFORMATION

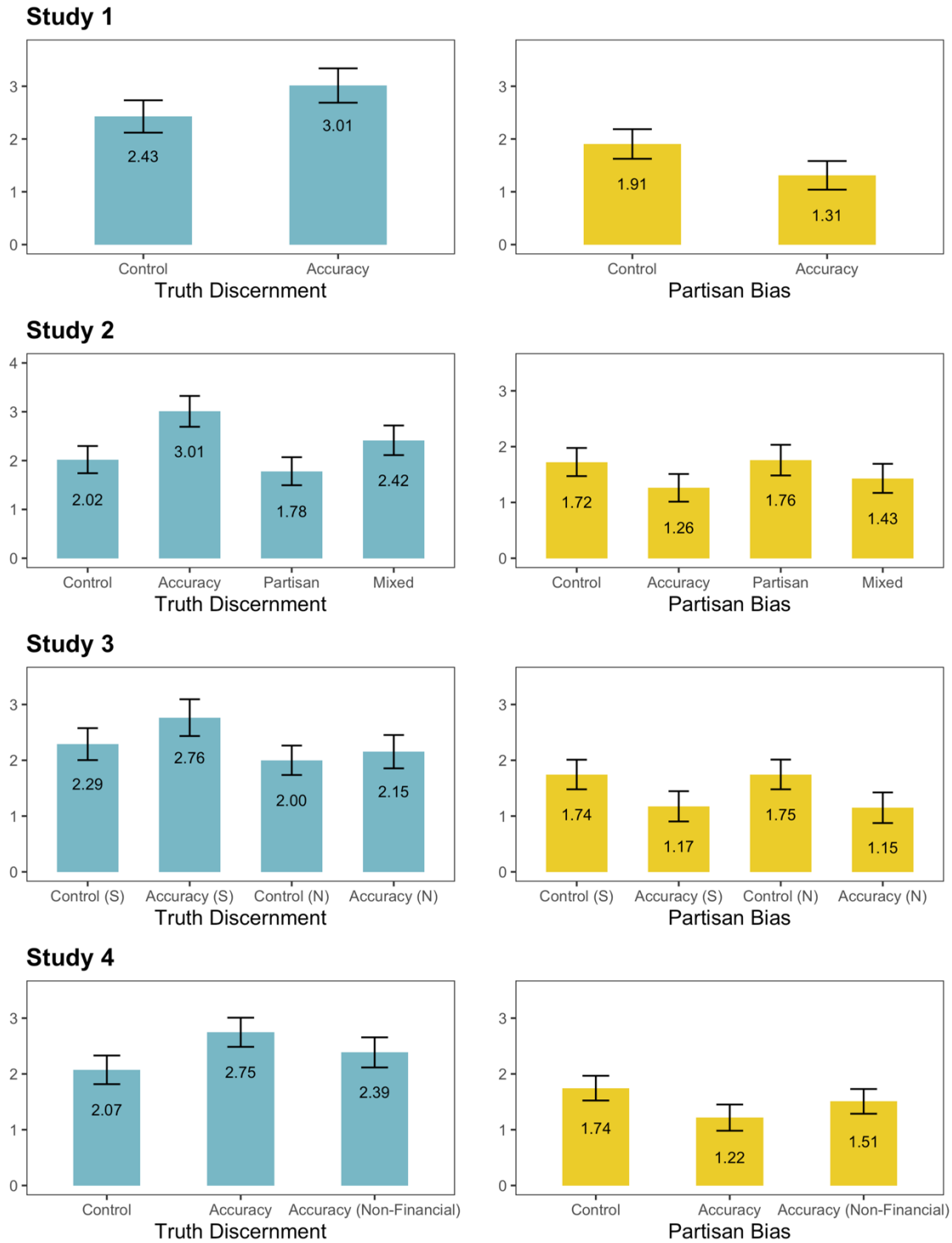
3.01, 95% CI = [2.68, 3.34]) were better at discerning truth than those in the *control* condition ( $M = 2.43$ , 95% CI = [2.12, 2.73]),  $t(457.64) = 2.58$ ,  $p = 0.010$ ,  $d = 0.24$ . In other words, participants answered 11.01 (out of 16) questions correctly in the *accuracy incentives* condition, as opposed to 10.43 (out of 16) questions in the *control* condition.

We next examined whether incentives decreased partisan bias, or the number of politically-congruent headlines participants rated as true minus the number of politically-incongruent headlines participants rated as true. As predicted, partisan bias in accuracy judgements was 31% smaller in the *accuracy incentives* condition ( $M = 1.31$ , 95% CI = [1.04, 1.58]) as compared to the *control* condition ( $M = 1.91$ , 95% CI = [1.62, 2.19]),  $t(495.8) = 3.01$ ,  $p = 0.001$ ,  $d = 0.28$ . Results from all four studies are plotted visually in **Fig 1**.

Additional analysis (See *Supplementary Appendix S1* for extended results) found that the *accuracy incentives* condition increased the percentage of politically-incongruent true headlines rated as true ( $M = 51.53$ , 95% CI = [47.36, 55.70]) as compared to the *control* condition ( $M = 38.25$ , 95% CI = [34.41, 42.08]),  $p < 0.001$ ,  $d = 0.43$ . Incentives did *not* significantly impact judgements of politically-congruent true news, politically-incongruent false news, or politically-congruent false news when controlling for multiple comparisons with Tukey post-hoc tests ( $ps > 0.444$ ). Thus, the effects of incentives were mainly driven by an increased belief in true news from the opposing party.

Finally, we examined whether the incentives influenced sharing discernment, or the number of true headlines shared minus the number of false headlines people intended to share. Interestingly, even though sharing higher-quality articles was not explicitly incentivized, sharing discernment was slightly higher in the *accuracy incentive* condition ( $M = 0.38$ , 95% CI = [0.28, 0.48]) as compared to the *control* condition ( $M = 0.22$ , 95% CI = [0.15, 0.30]),  $t(424.8) = 2.49$ ,  $p = 0.037$ ,  $d = 0.23$ .

# ACCURACY INCENTIVES SHAPE BELIEF IN (MIS)INFORMATION



**Fig. 1.** In **Study 1**, accuracy incentives improved truth discernment and decreased partisan bias in accuracy judgements, primarily by increasing belief in politically-incongruent true news. **Study 2** replicated these findings, but also found that incentives to identify articles that would be liked by one’s political in-group decreased truth discernment – even when paired with the accuracy incentive (the “mixed” condition). **Study 3** further replicated these findings and examined how effect sizes differed with and without source cues (S = source, N = no source). **Study 4** also replicated these findings and found that that a scalable, non-financial accuracy motivation intervention was also able to increase belief in politically-incongruent true news with a smaller effect size. Error bars represent 95% confidence intervals. \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ .

### Experiment 2: Social Motivations Interfere with Accuracy Motivations

In Experiment 2, we aimed to replicate and extend on the results of Experiment 1 by examining whether social motivations to correctly identify articles that would be liked by one's political in-group might interfere with accuracy motives. This condition was meant to mirror the incentive structure of social media whereby people try to share content that will be liked by their friends and followers. We recruited another politically-balanced sample of 998 US adults (see *Methods*). In addition to the *accuracy incentives* and *control condition*, we added a *partisan identity motivation* condition, whereby participants were given a financial incentive to correctly identify articles that would appeal to members of their own political party. Specifically, participants were told that they would receive a bonus payment of up to one dollar based on how accurately they identified articles that would be liked by members of their political party if they shared them on social media. Immediately after answering this question, participants were asked about the accuracy of the article and how likely they would be to share it. Building off of the predictions of the Identity-Based Model of Political Belief<sup>17</sup>, we wanted to examine whether increasing partisan-identity related goals might interfere with accuracy goals. Thus, in a final condition, called the *mixed motivation condition*, participants received a financial incentive of up to one dollar to identify articles that would be liked by one's in-group, followed by an additional financial incentive to accurately identify true and false articles.

We first examined how these motivations influenced truth discernment. Replicating the results of Experiment 1, there was a significant main effect of the *accuracy incentives* condition on truth discernment,  $F(1, 994) = 29.14, p < 0.001, \eta^2_G = 0.03$ , a significant main effect of the *partisan identity* manipulation on truth discernment,  $F(1, 994) = 7.53, p = 0.006, \eta^2_G = 0.01$ , but no significant interaction between the accuracy and the partisan identity manipulation ( $p = 0.237$ ). Tukey HSD post-hoc tests indicated that truth discernment was higher in the accuracy incentives condition ( $M = 3.01, 95\% \text{ CI} = [2.69, 3.32]$ ) compared to the control condition ( $M = 2.02, 95\% \text{ CI} = [1.74, 3.30]$ ),  $p < 0.001, d = 0.41$ . Truth discernment was also higher in the accuracy incentives condition compared to the *partisan identity* condition ( $M = 1.78, 95\% \text{ CI} = [1.49, 2.07]$ ),  $p < 0.001, d = 0.50$ , and the *mixed* condition ( $M = 2.42, 95\% \text{ CI} = [2.11, 2.71]$ ),  $p = 0.029, d = 0.27$ . However, the *mixed* condition did not differ from the control condition ( $p = 0.676$ ), and the *partisan identity* condition also did not significantly differ from the control condition ( $p = 0.241$ ). Taken together, these results suggest that accuracy motivations increase truth discernment, but partisan-identity motives can decrease truth discernment.

We then examined how these motives influenced partisan bias. Replicating the results from Experiment 1, there was a significant main effect of *accuracy incentives* on partisan bias,  $F(1, 994) = 9.01, p = 0.003, \eta^2_G = 0.01$ , but no effect of the *partisan identity* manipulation,  $F(1, 994) = 0.60, p = 0.441, \eta^2_G = 0.00$ , or the interaction between accuracy and the partisan identity manipulation,  $F(1, 994) = 0.27, p = 0.606, \eta^2_G = 0.00$ . Post-hoc tests indicated that there was a non-significant difference in partisan bias between the *accuracy incentives* condition ( $M = 1.26, 95\% \text{ CI} = [1.01, 1.51]$ ) and the *control* condition ( $M = 1.72, 95\% \text{ CI} = [1.47, 1.98]$ ),  $p = 0.062, d = 0.23$  – a 27% decrease in partisan bias. There was a significant difference between the *accuracy incentives* condition and the *partisan identity* motives condition ( $M = 1.76, 95\% \text{ CI} = [1.48, 2.03]$ ),  $p = 0.040, d = 0.24$ . No other post-hoc tests yielded significant differences ( $ps > 0.182$ ).

Follow-up analysis (*Supplementary Appendix S1*) once again indicated that the incentives primarily impacted the percentage of *politically incongruent true* headlines rated as accurate ( $M = 55.61\%, 95\% \text{ CI} = [51.68, 59.54]$ ) when compared to the control condition ( $M = 37.65\%, 95\%$

CI = [33.83, 41.46]),  $p < 0.001$ ,  $d = 0.58$ . The incentives again did not impact congruent true news, incongruent false news, or congruent false news ( $ps > 0.148$ ).

There was no significant effect of accuracy incentives on sharing discernment ( $p = 0.996$ ), diverging from the results of Study 1. However, follow-up analysis (*Supplementary Appendix S1*) indicated that those in the *partisan identity* motivation condition shared more politically-congruent news (either true or false) ( $M = 1.98$ , 95% CI = [1.90, 2.05]) as compared to the control condition ( $M = 1.80$ , 95% CI = [1.74, 1.87]),  $p = 0.015$ ,  $d = 0.21$ . Additionally, those in the *mixed* condition ( $M = 2.02$ , 95% CI = [1.94, 2.10]) shared more politically-congruent news (true or false) as compared to the control condition,  $p < 0.001$ ,  $d = 0.26$ . Thus, prompting participants to identify whether an article will be liked by their political allies – whether or not they are also incentivized to be accurate – appears to indiscriminately increase intentions to share both true and false news that appeals to one’s own political party.

### **Experiment 3: Accuracy Incentives and Source Cues in a Representative Sample**

In Experiment 3, we sought to replicate our prior findings in a nationally representative sample in the United States. We recruited a sample of 921 US participants that was quota-matched to the national distribution on age, gender, ethnicity, and political party. We also tested a potential psychological process underlying the effects of accuracy incentives. Since prior work has found strong effects of source cues<sup>21</sup> on judgements of news headlines, we suspected that people were responding to source cues when making judgements about news. Since true news often contains more recognizable sources with partisan connotations (e.g. “nytimes.com” as opposed to the fake news website “yournewswire.com”)<sup>56</sup>, this may explain why incentives only impacted judgements of true news in Experiments 1 and 2. To test this possibility, we examined the effect of incentives with and without source cues (e.g., a URL name such as “foxnews.com”) present beside the headlines (see *Methods* for more details). Because we wanted to compare the effects of accuracy incentives with and without sources, this study had four conditions: accuracy incentives (with sources), control (with sources), accuracy incentives (without sources), and control (without sources).

Replicating the main results from Experiments 1 and 2, the *accuracy incentives* condition significantly improved truth discernment,  $F(1, 917) = 4.44$ ,  $p = 0.035$ ,  $\eta^2_G = 0.01$ , reduced partisan bias,  $F(1, 917) = 18.21$ ,  $p < 0.001$ ,  $\eta^2_G = 0.02$ , and increased the number of politically-incongruent true articles rated as accurate,  $F(1, 917) = 20.94$ ,  $p < 0.001$ ,  $\eta^2_G = 0.02$ . Thus, accuracy incentives appear to increase accuracy and reduce partisan bias in a large representative sample, suggesting that the results of these experiments likely generalize to the US population as a whole.

Although effect sizes appeared to be descriptively smaller when sources were removed from the headlines (see **Fig. 1** and *Supplementary Appendix S1* for detail), we did not find significant interactions between the main outcome variables and the presence or absence of source cues. However, this study design did not provide strong power to test whether this was not due to chance, since interaction effects can require up to 16 times as much power as main effects<sup>57,58</sup> (see *Methods* for power analysis). Additional analysis using Bayes factors<sup>59</sup> reported in *Supplementary Appendix S1* did not find strong evidence for the absence of interaction effects. Like in Experiment 2, there was once again no significant impact of accuracy incentives on sharing discernment ( $p = 0.906$ ).

### **Experiment 4: The Effect of a Non-Financial Accuracy Motivation Intervention**

In Experiment 4, we replicated the accuracy incentive and control condition in another politically-balanced sample of 983 US adults, but also added a *non-financial accuracy*



*motivation condition.* This *non-financial accuracy motivation condition* was designed to rule out multiple interpretations behind our earlier findings. One mundane interpretation is that participants are merely saying what they believe fact-checkers think is true, rather than answering in accordance with their true beliefs. However, this non-financial intervention does not incentivize people to answer in ways that do not align with their actual beliefs. Additionally, because financial incentives are more difficult to scale to real-world contexts, the non-financial accuracy motivation condition speaks to the generalizability of these results to other, more scalable ways of motivating accuracy.

In the non-financial accuracy condition, people read a brief text about how most people value accuracy and how people think sharing inaccurate content hurts their reputation<sup>60</sup> (See intervention text in *Supplementary Appendix S2*.) People were also told to be as accurate as possible and that they would receive feedback on how accurate they were at the end of the study.

Our main pre-registered hypothesis was that this *non-financial accuracy motivation* condition would increase belief in politically-incongruent true news as relative to the *control* condition. An ANOVA found a main effect of the experimental conditions on the amount of politically-incongruent true news rated as true,  $F(2, 980) = 17.53$ ,  $p < 0.001$ ,  $\eta^2_G = 0.04$ . Supporting our main pre-registered hypothesis, the *non-financial accuracy motivation* condition increased the percentage of politically-incongruent true news stories rated as true ( $M = 43.97$ , 95% CI = [40.59, 47.34]) as compared to the *control* condition ( $M = 35.19$ , 95% CI = [31.93, 38.45]),  $p < 0.001$ ,  $d = 0.29$ . Replicating studies 1-3, the *accuracy incentive* condition also increased perceived accuracy of politically-incongruent true news ( $M = 49.15$ , 95% CI = [45.74, 52.55]),  $p < 0.001$ ,  $d = 0.45$ . The *accuracy incentive* and *non-financial accuracy motivation* condition were not significantly different from one another ( $p = 0.083$ ,  $d = 0.17$ ), though this may be because we did not have enough power to detect a difference. In short, the *non-financial accuracy motivation* manipulation was also effective at increasing belief in politically-incongruent true news, with an effect about 63% as large as the effect of the financial incentive.

Since we expected the *non-financial accuracy motivation* condition to have a smaller effect than the *accuracy incentives* condition, we did not pre-register hypotheses for truth discernment and partisan bias, as we did not anticipate having enough power to detect effects for these outcome variables. Indeed, the *non-financial accuracy motivation* condition did not significantly increase truth discernment ( $p = 0.221$ ) or partisan bias ( $p = 0.309$ ). However, replicating studies 1-3, *accuracy incentives* once again improved truth discernment ( $p = 0.001$ ,  $d = 0.28$ ) and reduced partisan bias ( $p = 0.003$ ,  $d = 0.25$ ). The effect of the *non-financial accuracy motivation* condition was 47% as large as the effect of the *accuracy incentive* for truth discernment and 45% as large for partisan bias. There was also no overall effect of the experimental conditions on sharing discernment ( $p = 0.689$ ). See *Supplementary Appendix S1* for extended results.

Together, these results suggest that a subtler (and also more scalable) accuracy motivation intervention that does not employ financial incentives may be effective at increasing the perceived accuracy of true news from the opposing party, but appears to have a smaller effect size than the stronger financial incentive intervention.

### **Integrative Data Analysis**

To generate more precise estimates of our effects, we pooled data from all four studies<sup>2</sup> to conduct an integrative data analysis (IDA).<sup>61</sup> For the IDA, we only used the 16 news headlines

---

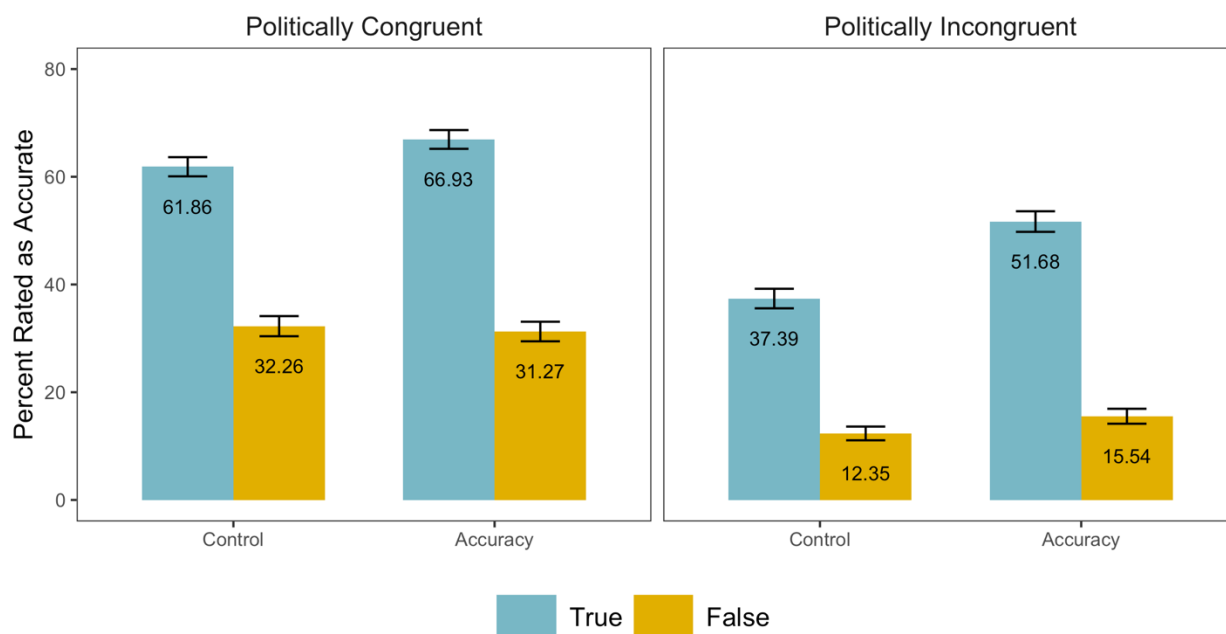
<sup>2</sup> We did not have any studies in the file drawer on this topic, meaning that our estimate was not influenced by publication bias.

## ACCURACY INCENTIVES SHAPE BELIEF IN (MIS)INFORMATION

that were used in all four studies, and only included the *accuracy incentives* and *control* conditions that were used in all four studies.

Incentives had the largest positive effect on the perceived accuracy of politically-incongruent true news,  $p < .001$ ,  $d = 0.47$ ; and a smaller positive effect on the perceived accuracy of politically-congruent true news,  $p = 0.001$ ,  $d = 0.17$ . Incentives did not significantly affect belief in politically-incongruent false news,  $p = 0.163$ ,  $d = 0.13$ , or belief in politically-congruent false news,  $p = 0.993$ ,  $d = -0.04$  (See **Fig. 2**) after adjusting for multiple comparisons with Tukey post-hoc tests. Analysis for each individual item revealed that incentives significantly increased belief in all true items, but they did not significantly decrease belief in any false items (though they significantly increased belief in one false item). More details are reported in *Supplementary Appendix S1*, and a headline-level analysis is reported in *Supplementary Appendix S3*. Additional analysis using Bayes Factors reported in *Supplementary Appendix S4* found strong evidence that incentives impacted belief in both politically-congruent and politically-incongruent true news, but found inconsistent evidence that they affected belief in false news.

While effects on sharing discernment were inconsistent across studies, the IDA found that there was a small positive effect of the incentive on sharing discernment,  $t(2020.20) = 2.19$ ,  $p = 0.029$ ,  $d = 0.10$ . Finally, people spent slightly more time on each headline in the accuracy incentives condition,  $t(818.53) = 2.34$ ,  $p = 0.019$ ,  $d = 0.16$ , indicating that incentives may have led people to put more effort into their responses.

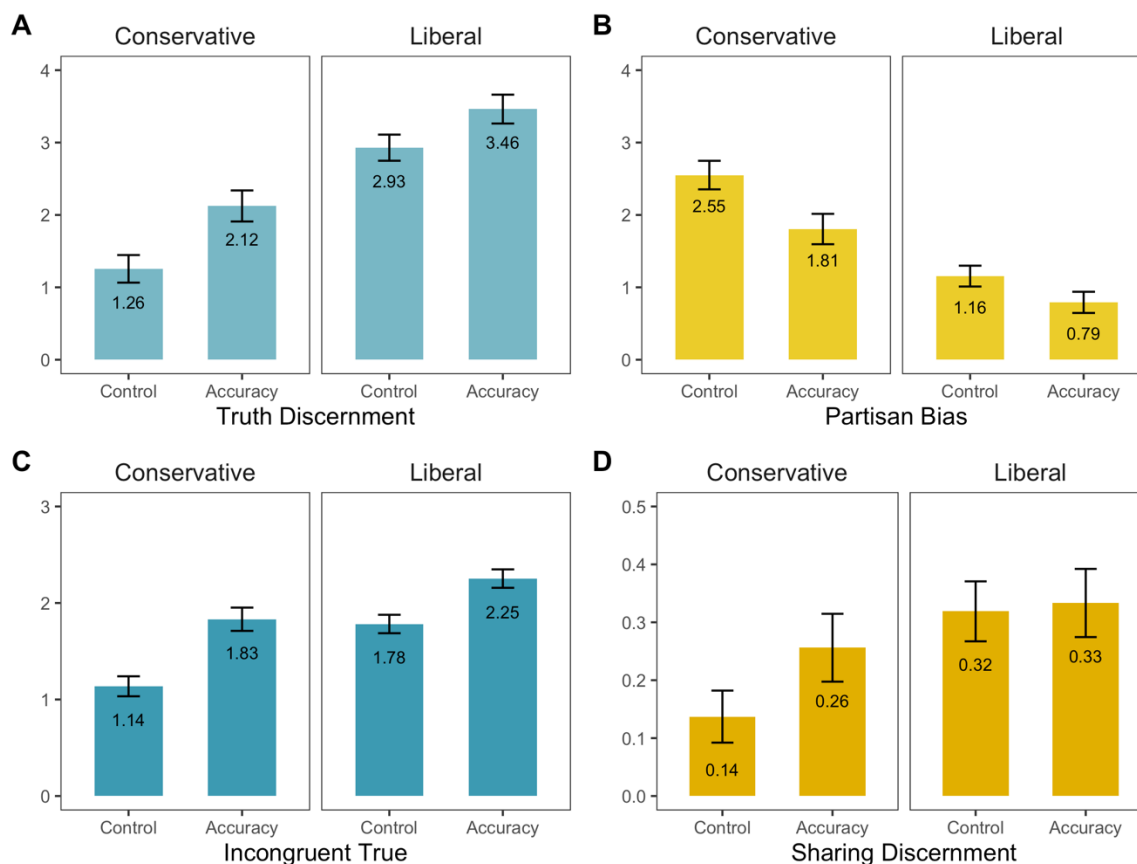


**Fig 2.** Integrative data analysis results (with data from all four studies,  $n = 2,092$ ) broken up by headline type. Incentives had a large effect on belief in politically-incongruent true news, and also had an effect on politically-congruent true news. Incentives did not have a significant effect on politically-congruent or politically-incongruent false news when controlling for multiple comparisons. Headline-level analysis revealed that incentives increased belief in all 8 true items, but did not decrease belief in a single false item (See *Supplementary Appendix S3* for item-level analysis).

### Incentives Reduce the Accuracy Gap Between Liberals and Conservatives

## ACCURACY INCENTIVES SHAPE BELIEF IN (MIS)INFORMATION

Replicating prior work<sup>29–34</sup>, conservatives were worse at discerning between true and false headlines than liberals. Conservatives answered about 9.26 (out of 16) questions correctly when not incentivized to be accurate and liberals answered 10.93 questions out of 16 correctly when unincentivized – a 1.67-point difference, 95% CI = [1.41, 1.94],  $t(1035.69) = 12.53$ ,  $p < .001$ ,  $d = 0.77$ . But, when conservatives were incentivized to be accurate, they answered 10.12 questions correctly—making the gap between incentivized conservatives and unincentivized liberals 0.81 points, 95% CI [0.53, 1.09],  $t(951.91) = 5.65$ ,  $p < .001$ ,  $d = 0.35$ . In other words, paying conservatives less than a dollar to correctly identify news headlines as true or false reduced the gap in performance between conservatives and (unincentivized) liberals by 51.50%. Incentives also considerably reduced the gap between conservatives and liberals in terms of partisan bias, sharing discernment, and belief in politically-incongruent true news. More detail is reported in *Supplementary Appendix S1* and plotted visually in **Fig. 3**. Altogether, these results suggest that a substantial portion of US conservatives’ tendency to believe and share less accurate news reflects a lack of motivation to be accurate rather than lack of knowledge alone.



**Fig 3.** Conservatives were worse at truth discernment as compared to liberals (**Panel A**). They also showed more partisan bias (**Panel B**), less belief in politically-incongruent true news (**Panel C**), and worse sharing discernment (**Panel D**). However, incentives closed the gap between conservatives and liberals for all these outcome variables by more than half, suggesting that conservatives’ greater tendency to believe in and share (mis)information may in part reflect a lack of motivation to be accurate (instead of lack of knowledge or ability alone).

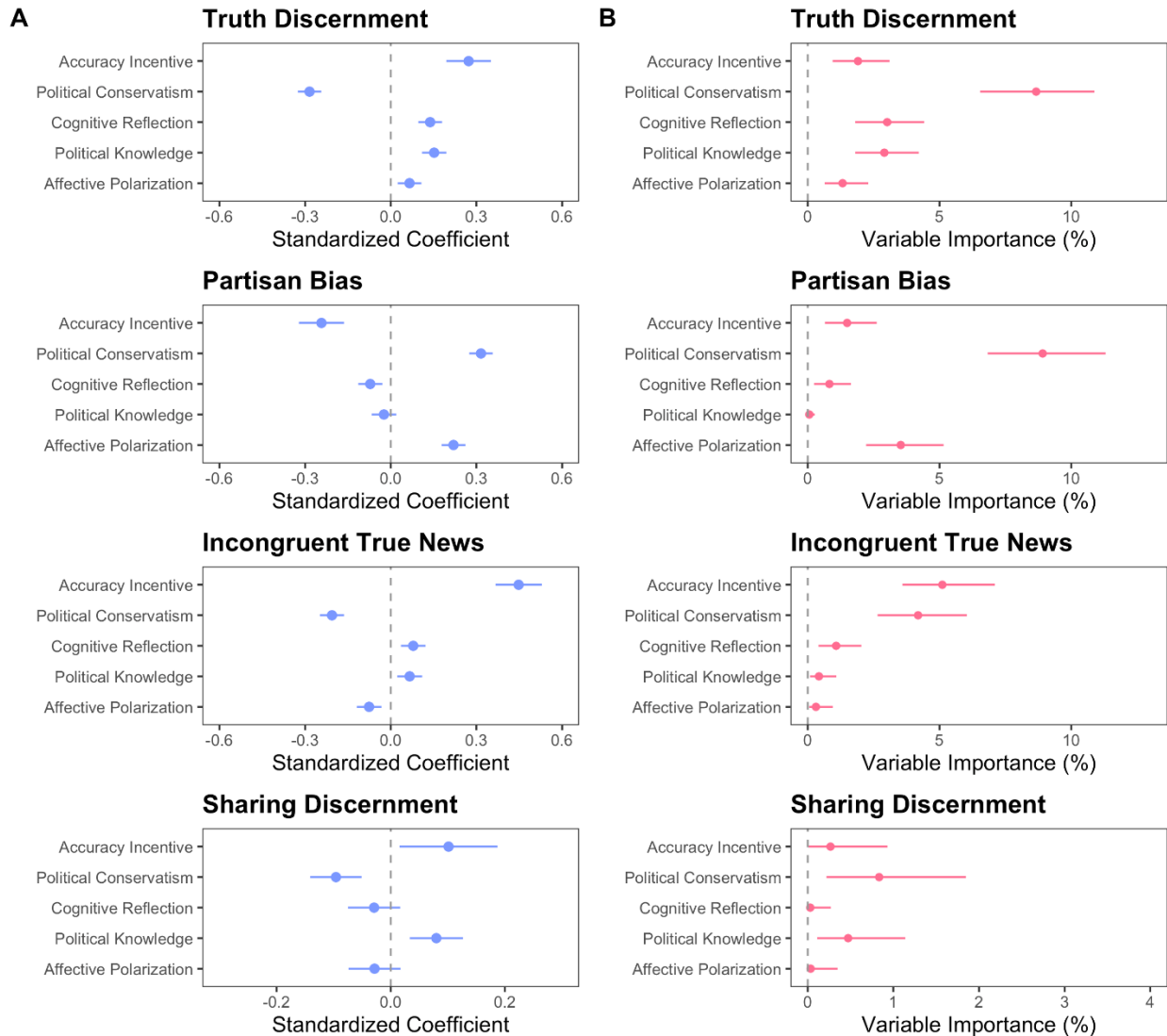
Importantly, the incentives improved truth discernment for both liberals,  $d = 0.23$ ,  $p < 0.001$ , and conservatives,  $d = 0.40$ ,  $p < 0.001$  (see *Supplementary Appendix S5* for table of effect sizes broken down by political affiliation). Descriptively, the effect sizes for our intervention

were larger for conservatives than liberals, which diverges from other misinformation interventions that tend to show larger effect sizes for liberals<sup>62,63</sup>. Furthermore, political ideology (liberal vs. conservative) was a significant moderator of belief in incongruent true news,  $p = 0.033$ , and partisan bias,  $p = 0.029$ , (though this moderation effects was not significant for truth discernment,  $p = 0.095$ , or sharing discernment,  $p = 0.061$ ) such that the effects of incentives appeared to be larger for conservatives than liberals. The effect of the incentives on truth discernment was not significantly moderated by cognitive reflection, political knowledge, or affective polarization ( $ps < 0.182$ ). However, even though we had a large sample, we were still slightly underpowered to detect these interaction effects (see power analysis in *Methods*), and supplemental Bayesian analyses also did not find strong evidence for the significant moderation effects (*Supplementary Appendix S11*), so these interaction effects should be interpreted with caution.

#### **Relative Importance of Accuracy Incentives**

In each experiment, we measured other individual difference variables known to be predictive of truth discernment, such as cognitive reflection, political knowledge, partisan animosity, as well as demographic variables, such as age, education, and gender. We ran a multiple regression analysis on our IDA with all of these variables included in the model (**Fig. 4, Panel A**). To compare the relative importance of each of these predictors, we also ran a relative importance analysis using the “lmg” method<sup>64</sup>, which calculates the relative contribution of each predictor to the  $R^2$  (**Fig. 4, Panel B**). Full models and relative importance analyses are in *Supplementary Appendix S6* and *S7*.

# ACCURACY INCENTIVES SHAPE BELIEF IN (MIS)INFORMATION



**Fig. 4.** In **A**, multiple regression results for the main outcome variables: truth discernment, partisan bias, belief in incongruent true news, and sharing discernment. Standardized beta coefficients are plotted for ease of interpretation. In **B**, variable importance estimates (LMG values) with bootstrapped confidence intervals are shown to examine the estimated percentage contribution of each predictor to the  $R^2$ .

Political conservatism and accuracy incentives were among the most important predictors for many of the key outcome variables, although confidence intervals were large and overlapping for the relative importance analysis (See *Supplementary Appendix S4*). While prominent accounts of misinformation sharing claim that partisanship and politically motivated cognition play a limited role in the belief and sharing of misinformation as compared to other factors (such as reflection or inattention)<sup>14,65</sup>, our results indicate that motivation and partisan identity or ideology are indeed very important factors.

Our data point to the importance of broad theoretical accounts of (mis)information belief and sharing that integrate motivation and partisan identity with other variables<sup>2,14,15,28,66</sup>. Indeed, an investigation using cognitive modeling found that a broad model of misinformation belief that included multiple factors (such as partisan identity, cognitive reflection, and more) performed

better at predicting acceptance of misinformation than other models that included fewer variables<sup>67</sup>.

### Discussion

Across four experiments ( $n = 3,364$ ), we find that increasing people's motivation to be accurate via a small financial incentive of up to one-dollar improved accuracy in discerning between true and false news and decreased the partisan divide in belief in news by about 30%. These effects were driven primarily by an increased belief in politically-incongruent true news ( $d = 0.47$ ), and no significant effects were found for false news (which people encounter relatively infrequently online<sup>68</sup>). Furthermore, providing people with an incentive to identify articles that would be liked by their political in-group reduced accuracy and increased intentions to share politically-congruent true and false news. Thus, social or partisan identity goals appear to interfere with accuracy goals. Additionally, a non-financial accuracy motivation intervention that asked people to be accurate, provided people feedback about their accuracy, and emphasized the social norm and reputational benefits of being accurate, significantly increased the perceived accuracy of politically-incongruent true news ( $d = 0.29$ ). This illustrates that accuracy motivation interventions can be applied at scale.

These results make two key theoretical contributions. First, they suggest that partisan differences in news judgements do not simply reflect differences in factual knowledge<sup>14</sup>. Instead, our data suggest that a substantial portion of this partisan divide can be attributed to a lack of motivation to be accurate. While there have been debates about whether partisan differences in belief reflect differing prior beliefs versus politically-motivated cognition<sup>25,26</sup>, our studies provide causal evidence for the effect of motivation on belief. Along with other research<sup>41,69,70</sup>, these findings suggest that survey data about belief in (mis)information should not be taken at face value, because people answer survey questions differently when they are highly motivated to be accurate. However, judgements of false headlines appeared to be unaffected by accuracy motivations, suggesting that other factors may play a more prominent role in people's assessment of false news as compared to true news.

Second, while a number of studies have observed that American conservatives tend to be more susceptible to misinformation than liberals<sup>29-34</sup>, our studies find that the gap in accuracy between liberals and (unincentivized) conservatives closes by more than half when conservatives are motivated to be accurate. Future work could examine whether this asymmetry arises due to the dynamics of partisan identity, party leadership, and social norms in the United States during this specific political climate, or if it reflects broader differences between liberals and conservatives that can be observed across cultures<sup>71,72</sup>.

These results also have practical implications for interventions<sup>73,74</sup>. Accuracy incentives improved the accuracy of people's judgements, and an integrative data analysis found that this effect may have spilled over into intentions to share more accurate articles. However, the effect on sharing intentions was small and inconsistent across studies. This may be in part because people were asked about accuracy before being asked about sharing intentions, and past research has found that merely asking people about accuracy can improve the accuracy of sharing intentions<sup>65</sup>. Further, making partisan-identity motivations salient increased the sharing of both politically-congruent false (and true) news. Thus, interventions and social media design features should aim to both *increase* accuracy motivations and *decrease* motivations to share inaccurate content that receives high social reward. While effects were only found for false (and not true) headlines, people tend to encounter blatantly false news very infrequently<sup>68</sup>, leading some to

## ACCURACY INCENTIVES SHAPE BELIEF IN (MIS)INFORMATION

suggest that increasing trust in reliable news is more important than reducing belief in falsehoods<sup>75</sup> and that researchers should employ a broad definition of misinformation<sup>76</sup>.

One limitation of this work is that survey experiments have unknown ecological validity. To maximize ecological validity, we used real, pre-tested news headlines in the format in which they would be regularly encountered on social media websites such as Facebook. Additionally, self-reported sharing intentions are highly correlated with real online news sharing<sup>77</sup>, and a field experiment suggests that priming accuracy can improve news sharing decisions on Twitter<sup>65</sup>, illustrating that results from survey experiments on misinformation can translate to the field. Another potential limitation is that there are multiple ways to interpret the effects of financial incentives. For instance, people may be guessing what they think fact-checkers believe to earn money, rather than expressing their true beliefs. However, this interpretation is unlikely to explain the full effect, since a subtle non-financial accuracy motivation intervention had similar (albeit smaller) effects. Furthermore, supplementary analysis found that an extremely small percentage of participants reported answering in ways that did not accord with their true beliefs to receive money (See *Supplementary Appendix S1*).

### **Conclusions**

There is a sizable partisan divide in the kind of news liberals and conservatives believe in, and conservatives tend to believe in and share more false news than liberals. Yet, these differences are not immutable. Motivating people to be accurate improves accuracy about the veracity of (true but not false) news headlines, reduces partisan bias, and closes a substantial portion of the gap in accuracy between liberals and conservatives. Theoretically, these results identify accuracy and social motivations as key factors in driving news belief and sharing. Practically, these results suggest that shifting motivations may be a useful strategy for improving the quality of the news content that people consume and share online.

## Methods

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the experiment. The research methods were approved by the University of Cambridge Psychology Ethics Committee (Protocol #PRE.2020.110). These studies were pre-registered. Stimuli, Qualtrics survey files, anonymized data, analysis code, and all pre-registrations are available on our OSF page: <https://osf.io/75sqf>.

### Experiment 1

**Participants.** The experiment launched on November 30, 2020. We recruited 500 participants via the survey platform Prolific Academic<sup>54</sup>. Specifically, we recruited 250 conservative participants and 250 liberal participants from the US via Prolific Academic's demographic pre-screening service to ensure the sample was politically balanced. Our a priori power analysis indicated that we would need 210 participants to detect a medium effect size of  $d = 0.50$  at 95% power, though we doubled this sample size to account for partisan differences and oversampled to account for exclusions. 511 participants took our survey. Following our pre-registered exclusion criteria, we excluded 32 participants who failed our attention check (or did not get far enough in the experiment to reach our attention check), and an additional 17 participants who said they responded randomly at some time during the experiment. This left us with a total of 462 participants (194 M, 255 F, 12 Trans/Nonbinary; age:  $M = 35.85$ ,  $SD = 13.66$ ; Politics: 253 Democrats, 201 Republicans). The Experiment 1 pre-registration is available here: <https://aspredicted.org/blind.php?x=gk9xg5>.

**Materials.** The materials were 16 pre-tested true and false news headlines from a large pre-tested sample of 225 news headlines<sup>55</sup>. In total, eight of these news headlines were false, and eight of the news headlines were true. Because we were interested in whether accuracy incentives would reduce partisan bias, we specifically selected headlines that had a sizable gap in perceived accuracy between Republicans and Democrats as reported in the pre-test, as well as headlines that were not outdated (the pre-test was conducted a few months before the first experiment). Specifically, we chose eight headlines (four false and four true) that Democrats rated as more accurate than Republicans in the pre-test, and eight headlines (four false and four true) that Republicans rated as more accurate than Democrats. See *Supplementary Appendix S8* for example stimuli and the OSF page for full materials.

### Procedure

**News Evaluation Task.** Participants were shown these 16 news headlines, along with an accompanying picture and source (similar to how a news article preview would show up on someone's Facebook feed), and asked "To the best of your knowledge, is the claim in the above headline accurate?" on a scale from 1 ("extremely inaccurate") to 6 ("extremely accurate"). Afterwards, they were asked "If you were to see the above article on social media, how likely would you be to share it?" on a scale from 1 ("extremely unlikely") to 6 ("extremely likely").

**Accuracy Incentives Manipulation.** Half of the participants were randomly assigned to a *control condition*, in which we explained the news evaluation task, but we did not provide any information about a bonus payment. The other half were assigned to an *accuracy incentives condition*. In this condition, we explained the news evaluation task, and then told participants they would receive a "bonus payment of up to \$1.00 based on how many correct answers [they] provide regarding the accuracy of the articles. Correct answers are based on the expert evaluations of non-partisan fact-checkers." Specifically, they received one dollar for answering 15 out of 16 questions correctly, and fifty cents for answering 13 out of 16 questions correctly. Since we measured accuracy on a continuous scale, we told participants that "if the headline



describes a true event, either ‘slightly accurate,’ ‘moderately accurate,’ or ‘extremely accurate’ constitute correct responses. Similarly, if the headline describes a false event, either ‘extremely inaccurate,’ ‘moderately inaccurate,’ or ‘slightly inaccurate’ constitute ‘correct’ responses.” In other words, the continuous scale was measured dichotomously for the purposes of giving financial incentives. Participants were also notified that all other questions would not affect their bonus payment. See *Supplementary Materials S2* or the OSF for full manipulation text.

**Other Measures.** We gave participants a 3-item cognitive reflection task<sup>51</sup>. We measured participants’ political knowledge using a 5-item scale<sup>48</sup> and in-group love/out-group hate with feeling thermometers<sup>78</sup>. See *Supplementary Appendix S9* and the OSF for question text. These measures were repeated across all studies.

**Analysis.** For truth discernment, partisan bias, and sharing discernment, independent samples t-tests were used. While we asked participants to rate the truth of headlines on a continuous scale, these variables were recoded as dichotomous for analysis because the financial incentive only rewarded participants based on whether they correctly identified a headline as true or false. Since we did not clearly specify this in the Experiment 1 pre-registration (but did for Experiments 2-4), we show the results with a continuous coding in *Supplementary Appendix S10*. The continuous coding did not change the conclusions of our studies.

To test what types of headlines were affected by the incentives, we ran a 2 (accuracy incentive vs. no incentive) X 2 (politically congruent vs. politically incongruent) X 2 (true headlines vs. false headlines) mixed-design ANOVA with the percent of articles rated as accurate as the dependent variable, and then followed up with Tukey HSD post-hoc tests. Extended analyses are in *Supplementary Appendix S1*.

## Experiment 2

**Participants.** The experiment launched on January 22, 2021. We aimed to recruit 1000 total participants (250 per condition) via the survey platform Prolific Academic, though we over-sampled and recruited 1,100 to account for exclusion criteria. We chose this sample size because a power analysis revealed that we needed at least 216 participants per condition to detect the smallest effect size ( $d = 0.24$ ) at 0.80% power using a one-tailed t-test (although two-tailed tests were used for all analysis). Once again, we used Prolific’s pre-screening platform to recruit 550 liberals and 550 conservatives from the United States, and 1,113 participants took our survey. Following our pre-registered exclusion criteria, we excluded 76 participants who failed our attention check (or did not finish enough of the survey to reach the attention check) and an additional 39 participants who said they responded randomly at some point during the experiment. This left us with a total of 998 participants in total (463 M, 505 F, 30 transgender/non-binary/other; age:  $M = 36.17$ ,  $SD = 13.94$ ; politics: 568 liberals, 430 conservatives). This experiment was also pre-registered (pre-registration available here: [https://aspredicted.org/blind.php?x=/FKF\\_15L](https://aspredicted.org/blind.php?x=/FKF_15L)).

**Social Incentives & Mixed Incentives Manipulations.** In the new *partisan identity* condition, participants were first asked before the experiment to report the political party with which they identify. Then, they were told that they would receive a bonus payment of up to \$1.00 based on how accurately they identified information that would be liked by members of their political party if they shared it on social media. Bonuses were awarded based on how closely participants’ answers matched partisan alignment scores from a pre-test<sup>48</sup>. Before each question about accuracy and sharing, participants were asked “If you shared this article on social media, how likely is it that it would receive a positive reaction from [your political party] (e.g., likes, shares, and positive comments)?” In the *mixed* condition, participants were first given financial

incentives for both correctly identifying whether the article would be liked by a member of their political party, and were then asked about accuracy and given incentives for identifying whether the article was accurate. See *Supplementary Appendix S2* for full intervention text.

**Analysis.** To understand the impact of accuracy and partisan identity motivations on truth discernment and partisan bias, we ran 2 (accuracy incentive vs. control) X 2 (partisan identity vs. control) ANOVAs and followed up on the results using Tukey HSD post-hoc tests. To test what types of headlines were affected by the incentives, we ran a 2 (accuracy vs. control) X 2 (partisan identity vs. control) X 2 (politically congruent vs. politically incongruent) X 2 (true headlines vs. false headlines) mixed-design ANOVA with the percentage of articles rated as accurate as the dependent variable, and then followed up with Tukey HSD post-hoc tests.

### Experiment 3

**Participants.** The experiment launched on June 13, 2021. We aimed to recruit a nationally representative sample (quota-matched to the US population distribution by age, ethnicity, and gender) of 1,000 participants via the survey platform Prolific. As in studies 1 and 2, we ensured that the nationally-representative sample was politically balanced, or half liberal and half conservative. 1,055 total participants took the survey. Then, we once again excluded 95 participants who failed our attention check (or did not make it to that point in the survey), as well as 39 participants who said they were responding randomly at some point in the survey. This left us with a total of 921 participants (439 M, 470 F, 12 transgender/non-binary/other; age:  $M = 40.07$ ,  $SD = 14.67$ ; politics: 542 liberals, 379 conservatives). This experiment was also pre-registered (pre-registration available at: [https://aspredicted.org/7M2\\_9K9](https://aspredicted.org/7M2_9K9)).

**Materials.** We once again used the same 16 pre-tested true and false news headlines in addition to eight extra true and false news items from the same pre-test. For consistency, we report the results of the 16 news items in the manuscript, but we also report the results for the full set of 24 items in the *Supplementary Appendix S3*, which did not change our conclusions.

**Manipulations.** In addition to the accuracy incentive and control condition, participants were assigned to identical accuracy incentive and control conditions *without source cues* present on the stimuli. In these conditions, the sources (e.g., “nytimes.com”) were greyed out, so participants could only make assessments of the stimuli based on the photo and headline alone (see *Supplementary Materials S8* for examples).

**Analysis.** To understand the impact of accuracy incentives and source cues on truth discernment and partisan bias, we ran 2 (accuracy vs. control) X 2 (source vs. no source) ANOVAs and followed up on the results using Tukey HSD post-hoc tests. To test what types of headlines were affected by the incentives, we ran a 2 (accuracy vs. control) X 2 (source vs. no source) X 2 (politically congruent vs. politically incongruent) X 2 (true headlines vs. false headlines) mixed-design ANOVA with the percent of articles rated as accurate as the dependent variable, and then followed up with Tukey HSD post-hoc tests.

**Power Analysis for Interaction Effects.** Based on the effect sizes of Study 2 and the principle that 16 times the sample size is needed to detect an attenuated interaction effect<sup>57,58</sup>, a power analyses conducted after we ran the study found that we needed roughly 1536 participants to detect an interaction for the amount of politically-incongruent news rated as true, 2560 participants to detect an interaction effect for truth discernment, and 7488 participants to detect an interaction effect for partisan bias with 80% power. Thus, this particular design was underpowered to detect whether accuracy incentives interacted with source cues.

### Experiment 4

**Participants.** This experiment launched on May 25, 2022. We aimed to recruit a total of 1000 participants (roughly 333 per condition) via the platform Prolific academic. We chose this sample size as a power analysis found that we would 312 per condition to detect the smallest effect size found in the previous study ( $d = 0.26$ ) with 90% power. Additionally, we wanted relatively high power because we expected the effect of the non-financial accuracy motivation condition to be smaller than that of the financial incentive condition. We used Prolific's pre-screening platform to recruit a sample that was balanced by politics and gender. 1007 participants took our survey. Following our pre-registered exclusion criteria, we excluded 17 participants who failed our attention check (or did not finish enough of the survey to reach the attention check) and an additional 8 participants who said they responded randomly at some point during the experiment. This left us with a total of 993 participants in total (486 M, 483 F, 30 transgender/non-binary/other; age:  $M = 41.46$ ,  $SD = 15.06$ ; politics: 507 liberals, 476 conservatives). This experiment was also pre-registered (pre-registration available here: [https://aspredicted.org/86W\\_BY4](https://aspredicted.org/86W_BY4)).

**Materials.** We once again used the same 16 pre-tested true and false news headlines extra “misleading” news headlines.

**Analysis.** Following our pre-registered analysis plan, we ran a 1-way (accuracy vs. control vs. non-financial accuracy motivation) ANOVA with the percent of incongruent-true articles rated as true as the dependent variable, followed up by Tukey post-hoc tests. We also ran 1-way ANOVAs with truth discernment and partisan bias and DVs and followed up with post-hoc tests.

### **Integrative Data Analysis**

**Analysis.** We conducted moderation analysis on the pooled dataset by testing for an interaction between the condition and political ideology (liberal vs. conservative) in a linear regression. To test the relative importance of each predictor, we ran a relative importance analysis using the “reliampo” package in R. Bootstrapped confidence intervals were calculated for “lmg” variables using 1,000 bootstraps.

**Power Analysis for Moderation Effects.** Using effect sizes from the integrative data analysis and the principle that 16 times the sample size is needed to detect an attenuated interaction effect<sup>57,58</sup>, a post-hoc power analysis found that we needed 2336 participants to detect an interaction effect for the amount of politically-incongruent news rated as true, 5984 participants to detect an interaction effect for truth discernment, 7488 for partisan bias, and 50,336 to detect an interaction for sharing discernment. Thus, moderation effects should be interpreted with caution.

**Signal Detection Analysis.** As another robustness check, we also conducted supplemental analysis using signal detection modeling<sup>19</sup>. This analysis found that incentives increased participants' discrimination between true and false news (for both politically-congruent and politically incongruent headlines), and also increased the threshold by which people accepted politically-incongruent headlines as true (See *Supplementary Appendix S12*). In sum, analysis using signal detection modeling yielded highly similar results to our main analysis.

## ACCURACY INCENTIVES SHAPE BELIEF IN (MIS)INFORMATION

**Data Availability:** Anonymized data, Qualtrics files, and stimuli are available on the Open Science Framework (OSF): <https://osf.io/75sqf>.

## ACCURACY INCENTIVES SHAPE BELIEF IN (MIS)INFORMATION

**Code Availability:** The R code necessary to reproduce our results is available on the Open Science Framework (OSF): <https://osf.io/75sqf>.

## References

1. Lewandowsky, S., Ecker, U. K. & Cook, J. Beyond misinformation: Understanding and coping with the “post-truth” era. *Journal of applied research in memory and cognition* **6**, 353–369 (2017).
2. Van Bavel, J. J. *et al.* Political psychology in the digital (mis) information age: A model of news belief and sharing. *Social Issues and Policy Review* **15**, 84–113 (2021).
3. Kreps, S., Dasgupta, N., Brownstein, J. S., Hswen, Y. & Kriner, D. L. Public attitudes toward COVID-19 vaccination: The role of vaccine attributes, incentives, and misinformation. *npj Vaccines* **6**, 1–7 (2021).
4. Pavlović, T. *et al.* Predicting attitudinal and behavioral responses to COVID-19 pandemic using machine learning. *PNAS Nexus* (2022).
5. Biddlestone, M., Azevedo, F. & van der Linden, S. Climate of conspiracy: A meta-analysis of the consequences of belief in conspiracy theories about climate change. *Current Opinion in Psychology* 101390 (2022).
6. Van der Linden, S., Leiserowitz, A., Rosenthal, S. & Maibach, E. Inoculating the public against misinformation about climate change. *Global Challenges* **1**, (2017).
7. Pierri, F. *et al.* Online misinformation is linked to early COVID-19 vaccination hesitancy and refusal. *Scientific reports* **12**, 1–7 (2022).
8. Loomba, S., de Figueiredo, A., Piatek, S. J., de Graaf, K. & Larson, H. J. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nature Human Behaviour* 1–12 (2021).
9. Dechêne, A., Stahl, C., Hansen, J. & Wänke, M. The truth about the truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review* **14**, 238–257 (2010).
10. Pennycook, G., Cannon, T. & Rand, D. G. Prior exposure increases perceived accuracy of fake news. (2018).
11. Johnson, N. F. *et al.* The online competition between pro-and anti-vaccination views. *Nature* **582**, 230–233 (2020).
12. Pennycook, G. & Rand, D. G. Examining false beliefs about voter fraud in the wake of the 2020 Presidential Election. *The Harvard Kennedy School Misinformation Review* (2021).
13. Roozenbeek, J. *et al.* Susceptibility to misinformation about COVID-19 around the world. *Royal Society open science* **7**, 201199 (2020).
14. Pennycook, G. & Rand, D. G. The psychology of fake news. *Trends in cognitive sciences* (2021).
15. van der Linden, S. *et al.* How Can Psychological Science Help Counter the Spread of Fake News? *The Spanish Journal of Psychology* **24**, (2021).
16. Robertson, C., Pretus, C., Rathje, S., Harris, E. A. & Van Bavel, J. J. How Social Identity Shapes Conspiratorial Belief. (2022).
17. Van Bavel, J. J. & Pereira, A. The partisan brain: An Identity-based model of political belief. *Trends in cognitive sciences* **22**, 213–224 (2018).
18. Aslett, K. *et al.* Measuring belief in fake news in real-time. in *CEUR Workshop Proceedings* vol. 2890 (CEUR-WS, 2021).
19. Batailler, C., Brannon, S. M., Teas, P. E. & Gawronski, B. A signal detection approach to understanding the identification of fake news. *Perspectives on Psychological Science* **22**, (2021).
20. Gawronski, B. Cognitive Sciences. *Trends in Cognitive Sciences* **25**, 723 (2021).

21. Traber, C. S. & van der Linden, S. Birds of a feather are persuaded together: Perceived source credibility mediates the effect of political bias on misinformation susceptibility. *Personality and Individual Differences* **185**, 111269 (2022).
22. van der Linden, S., Panagopoulos, C. & Roozenbeek, J. You are fake news: political bias in perceptions of fake news. *Media, Culture & Society* **42**, 460–470 (2020).
23. Kunda, Z. The case for motivated reasoning. *Psychological bulletin* **108**, 480 (1990).
24. Taber, C. S. & Lodge, M. Motivated skepticism in the evaluation of political beliefs. *American journal of political science* **50**, 755–769 (2006).
25. Druckman, J. N. & McGrath, M. C. The evidence for motivated reasoning in climate change preference formation. *Nature Climate Change* **9**, 111–119 (2019).
26. Tappin, B. M., Pennycook, G. & Rand, D. G. 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 (2020).
27. Bayes, R., Druckman, J. N., Goods, A. & Molden, D. C. When and how different motives can drive motivated political reasoning. *Political Psychology* **41**, 1031–1052 (2020).
28. van der Linden, S. Misinformation: susceptibility, spread, and interventions to immunize the public. *Nat Med* 1–8 (2022) doi:10.1038/s41591-022-01713-6.
29. Garrett, R. K. & Bond, R. M. Conservatives' susceptibility to political misperceptions. *Science Advances* **7**, eabf1234 (2021).
30. Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B. & Lazer, D. Fake news on Twitter during the 2016 US presidential election. *Science* **363**, 374–378 (2019).
31. Guess, A., Nagler, J. & Tucker, J. Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science advances* **5**, eaau4586 (2019).
32. Lawson, M. A. & Kakkar, H. Of pandemics, politics, and personality: The role of conscientiousness and political ideology in the sharing of fake news. *Journal of Experimental Psychology: General* (2021).
33. Pereira, A. & Van Bavel, J. Identity concerns drive belief in fake news. (2018).
34. van der Linden, S., Panagopoulos, C., Azevedo, F. & Jost, J. T. The paranoid style in American politics revisited: an ideological asymmetry in conspiratorial thinking. *Political Psychology* **42**, 23–51 (2021).
35. Roozenbeek, J. *et al.* Susceptibility to misinformation is consistent across question framings and response modes and better explained by myside bias and partisanship than analytical thinking. *Judgment and Decision Making* **17**, 547–573 (2022).
36. Evanega, S., Lynas, M., Adams, J., Smolenyak, K. & Insights, C. G. Coronavirus misinformation: quantifying sources and themes in the COVID-19 'infodemic'. *JMIR Preprints* **19**, 2020 (2020).
37. Mosleh, M. & Rand, D. Falsehood in, falsehood out: A tool for measuring exposure to elite misinformation on Twitter. (2021).
38. Jost, J. T., Glaser, J., Kruglanski, A. W. & Sulloway, F. J. Political conservatism as motivated social cognition. *Psychological bulletin* **129**, 339 (2003).
39. Baron, J. & Jost, J. T. False equivalence: Are liberals and conservatives in the United States equally biased? *Perspectives on Psychological Science* **14**, 292–303 (2019).
40. Bullock, J. G. & Lenz, G. Partisan bias in surveys. *Annual Review of Political Science* (2019).

41. Prior, M., Sood, G. & Khanna, K. You cannot be serious: The impact of accuracy incentives on partisan bias in reports of economic perceptions. *Quarterly Journal of Political Science* **10**, 489–518 (2015).
42. Jakesch, M., Koren, M., Evtushenko, A. & Naaman, M. The role of source and expressive responding in political news evaluation. in *Computation and Journalism Symposium* (2019).
43. Panizza, F. *et al.* Lateral reading and monetary incentives to sort out scientific (dis) information. (2021).
44. Speckmann, F. & Unkelbach, C. Monetary incentives do not reduce the repetition-induced truth effect. *Psychonomic Bulletin & Review* **29**, 1045–1052 (2022).
45. Melnikoff, D. E. & Strohming, N. The automatic influence of advocacy on lawyers and novices. *Nature Human Behaviour* **4**, 1258–1264 (2020).
46. Brashier, N. M. & Marsh, E. J. Judging truth. *Annual review of psychology* **71**, 499–515 (2020).
47. Rathje, S., Van Bavel, J. J. & van der Linden, S. Out-group animosity drives engagement on social media. *Proceedings of the National Academy of Sciences* **118**, (2021).
48. OSMUNDTSEN, M., BOR, A., VAHLSTRUP, P. B., BECHMANN, A. & PETERSEN, M. B. Partisan polarization is the primary psychological motivation behind political fake news sharing on Twitter. *American Political Science Review* 1–17 (2021).
49. Ren, Z. B., Dimant, E. & Schweitzer, M. E. Social Motives for Sharing Conspiracy Theories. *Eugen and Schweitzer, Maurice E., Social Motives for Sharing Conspiracy Theories (September 8, 2021)* (2021).
50. Brady, W. J., Crockett, M. J. & Van Bavel, J. J. The MAD Model of Moral Contagion: The Role of Motivation, Attention, and Design in the Spread of Moralized Content Online. *Perspectives on Psychological Science* 1745691620917336 (2019).
51. Pennycook, G. & Rand, D. G. Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition* (2018) doi:10.1016/j.cognition.2018.06.011.
52. Vegetti, F. & Mancosu, M. The impact of political sophistication and motivated reasoning on misinformation. *Political Communication* **37**, 678–695 (2020).
53. Guess, A. M. *et al.* 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**, 15536–15545 (2020).
54. Peer, E., Rothschild, D. M., Evernden, Z., Gordon, A. & Damer, E. MTurk, Prolific or panels? Choosing the right audience for online research. *Choosing the right audience for online research (January 10, 2021)* (2021).
55. Pennycook, G., Binnendyk, J., Newton, C. & Rand, D. A practical guide to doing behavioural research on fake news and misinformation. (2020).
56. Pennycook, G. & Rand, D. G. Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences* **116**, 2521–2526 (2019).
57. Gelman, A. You need 16 times the sample size to estimate an interaction than to estimate a main effect. *Statistical Modeling, Causal Inference, and Social Science* (2018).
58. Blake, K. R. & Gangestad, S. On attenuated interactions, measurement error, and statistical power: Guidelines for social and personality psychologists. *Personality and Social Psychology Bulletin* **46**, 1702–1711 (2020).



59. Wetzels, R., van Ravenzwaaij, D. & Wagenmakers, E.-J. Bayesian analysis. *The encyclopedia of clinical psychology* 1–11 (2014).
60. Altay, S., Hacquin, A.-S. & Mercier, H. Why do so few people share fake news? It hurts their reputation. *new media & society* 1461444820969893 (2019).
61. Curran, P. J. & Hussong, A. M. Integrative data analysis: the simultaneous analysis of multiple data sets. *Psychological methods* **14**, 81 (2009).
62. Rathje, S. Letter to the editors of Psychological Science: Meta-analysis reveals that accuracy nudges have little to no effect for US conservatives: Regarding Pennycook et al. (2020). *Psychological Science* (2022).
63. Pretus, C. *et al.* The role of political devotion in sharing partisan misinformation. (2021).
64. Tonidandel, S. & LeBreton, J. M. Relative importance analysis: A useful supplement to regression analysis. *Journal of Business and Psychology* **26**, 1–9 (2011).
65. Pennycook, G. *et al.* Shifting attention to accuracy can reduce misinformation online. *Nature* 1–6 (2021) doi:10.1038/s41586-021-03344-2.
66. Robertson, C. E., Pretus, C., Rathje, S., Harris, E. & Van Bavel, J. J. How Social Identity Shapes Conspiratorial Belief. *Current Opinion in Psychology* 101423 (2022).
67. Borukhson, D., Lorenz-Spreen, P. & Ragni, M. When Does an Individual Accept Misinformation? An Extended Investigation Through Cognitive Modeling. *Computational Brain & Behavior* 1–17 (2022).
68. Guess, A. M., Nyhan, B. & Reifler, J. Exposure to untrustworthy websites in the 2016 US election. *Nature human behaviour* **4**, 472–480 (2020).
69. Bishop, G. F. *The illusion of public opinion: Fact and artifact in American public opinion polls*. (Rowman & Littlefield Publishers, 2004).
70. Edwards, A. L. The social desirability variable in personality assessment and research. (1957).
71. Imhoff, R. *et al.* Conspiracy mentality and political orientation across 26 countries. *Nature Human Behaviour* 1–12 (2022).
72. Jost, J. T., van der Linden, S., Panagopoulos, C. & Hardin, C. D. Ideological asymmetries in conformity, desire for shared reality, and the spread of misinformation. *Current Opinion in Psychology* **23**, 77–83 (2018).
73. Roozenbeek, J., Sander, van der L., Goldberg, B., Rathje, S. & Lewandowsky, S. Psychological inoculation improves resilience against misinformation on social media. *Science Advances* (In Press).
74. Bak-Coleman, J. B. *et al.* Combining interventions to reduce the spread of viral misinformation. *Nature Human Behaviour* 1–9 (2022).
75. Acerbi, A., Altay, S. & Mercier, H. Research note: Fighting misinformation or fighting for information? (2022).
76. Traber, C. S. Misinformation: broaden definition to curb its societal influence. *Nature* **606**, 653–653 (2022).
77. Moseh, M., Pennycook, G. & Rand, D. G. Self-reported willingness to share political news articles in online surveys correlates with actual sharing on Twitter. *Plos one* **15**, e0228882 (2020).
78. Druckman, J. N. & Levendusky, M. S. What do we measure when we measure affective polarization? *Public Opinion Quarterly* **83**, 114–122 (2019).