

# Debunking COVID-19 falsehoods: a test of three cognitive predispositions as buffers for partisan influences

Wei Wang

*Annenberg School for Communication, University of Pennsylvania,  
Philadelphia, Pennsylvania, USA, and*

Danny Yihan Jia and Maria Elizabeth Grabe

*Division of Emerging Media Studies, College of Communication, Boston University,  
Boston, Massachusetts, USA*

1

Received 27 August 2025  
Revised 19 December 2025  
Accepted 22 February 2026

## Abstract

**Purpose** – Existing evidence suggests that in a politically polarized country like the United States, partisanship frequently diminishes the effectiveness of fact-checks on falsehoods. Here we tested if debunking sources interact with cognitive predispositions to moderate partisanship-driven resistance to COVID-19 falsehood corrections.

**Design/methodology/approach** – A Qualtrics sample of US adults ( $N = 454$ ) was recruited to experimentally test the likelihood that fact-checker warning labels on a false post versus user-posted debunks of the same falsehoods would moderate partisan receptivity to corrections. We also assessed whether cognitive predispositions (thinking styles, media literacy and trust in public health agencies) moderate debunking sources to buffer against partisan tendencies to believe debunked falsehoods.

**Findings** – Overall, Republicans exhibited stronger truth resistance than Democrats. Counter to expectation, thinking styles exacerbated the partisan divide and media literacy had no effect. However, trust in public health agencies buffered against partisan truth resistance for user but not fact-checker debunks, offering tentative evidence that some Republicans might have a reactive response against fact-checking.

**Originality/value** – Our research suggests that common strategies, such as promoting open-mindedness or media literacy education, are less effective at reducing differences in resistance to debunks driven by partisanship. We suggest that building institutional trust is a promising approach to bridging partisan divides in debunking effectiveness. We also demonstrate the robustness of Republican truth resistance, as this group is associated with low flexibility across levels of each cognitive predisposition.

**Keywords** Fact-checking, Debunking, Partisanship, COVID-19, Actively open-minded thinking, Institutional trust, Media literacy

**Paper type** Research article

The supercharged dissemination of falsehoods on social media (Gruzd and Mai, 2020) has created public concern that pressured social media platforms to adopt more stringent content moderation policies – most notably, the use of fact-check warning labels to flag and debunk questionable posts. Following the 2016 US election, Facebook launched an industry-leading effort to collaborate with third-party fact-checkers to label posts containing disputed information (Mosseri, 2016). In March 2020 alone, the platform displayed labels to more than 40 million COVID-19-related posts, yet 95% of users who viewed labeled content chose not to click on it (Meta, 2021).

© Wei Wang, Danny Yihan Jia and Maria Elizabeth Grabe. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at [Link to the terms of the CC BY 4.0 licence](#).

*Funding statement:* This work was supported in part by the Vaccine Confidence Fund.

*Declaration of interest:* The authors report there are no competing interests to declare.



Online Information Review  
Vol. 50 No. 8, 2026  
pp. 1-19

Emerald Publishing Limited  
e-ISSN: 1468-4535

p-ISSN: 1468-4527

DOI 10.1108/OIR-08-2025-0666

Empirical evidence suggests that debunks are generally successful in reducing susceptibility to falsehoods (see [Walter et al., 2020](#), for a review). However, their effectiveness is highly heterogeneous. Who debunks matters and depends on source expertise ([Zhang et al., 2021](#)) and perceived trustworthiness ([Guillory and Geraci, 2013](#)). Furthermore, factors such as preexisting beliefs and identity ([Jennings and Stroud, 2023](#)), analytical thinking styles ([Schwalbe et al., 2024](#)), media literacy ([Altay et al., 2024](#)) and institutional trust ([Roozenbeek et al., 2020](#)), all affect the extent to which media users engage in motivated reasoning ([Kunda, 1990](#)) when processing a debunk.

To complicate matters further, these sources of heterogeneity in debunk effectiveness might interact with each other to exert influence on misinformation processing – a dynamic that remains largely unexplored in debunking literature. While there is reliable evidence that Republicans tend to regard fact-checkers less favorably than Democrats ([Nyhan and Reifler, 2015](#)), it is less clear what features of a debunking message or which media user predispositions would mitigate this divide. Addressing this gap is normatively important because a well-functioning democracy requires informed citizens from across the political spectrum.

The current study addresses this gap by testing how the source of social media debunks and the cognitive predispositions of media users influence partisan divides in factually unsupported beliefs. We investigated this in the context of COVID-19, where Democrats and Republicans differ markedly both in vaccination rates ([Neely et al., 2022](#); [Ye, 2023](#)) and attitudes about COVID-19 debunks ([Robertson et al., 2020](#); [Jia et al., 2022](#)). Our findings confirmed the robustness of partisanship in driving resistance to content moderation. Surprisingly, instead of narrowing the partisan divide, one of the three cognitive predispositions we tested exacerbated partisan differences in truth resistance, another had no effect and one showed qualified effectiveness.

### **Evidence and explanations for the partisan divide in COVID-19 beliefs**

The pathogen prevalence hypothesis supports expectations of social conservatism as a positive correlate for proactive behavior to avoid infection during a pandemic like COVID-19 ([Samore et al., 2021](#)). In partisan terms, this implies Republicans should have embraced COVID-19 protective measures more fully than Democrats. Yet, evidence points to the opposite: Democrats were more likely to mask ([Howard, 2022](#)), comply with social distance recommendations ([Gratz et al., 2021](#)) and vaccinate ([Ruiz and Bell, 2021](#)), driven in part by partisan divides in COVID-19 beliefs. This paradox has drawn research attention from theoretical perches and levels of analysis ranging from algorithmic ([Cardenal et al., 2019](#)) to self-selected falsehood exposure ([Robertson et al., 2023](#)) and from ecologically produced epistemological tendencies ([Conway et al., 2020](#)) to partisan-aligned proclivities ([Samore et al., 2021](#)). These studies consistently point to political identity-driven motivated reasoning and cognitive predispositions as key explanations for partisan divides, and warrant closer consideration here.

First, partisanship shapes the way people process information such that when exposed to evidence-based information that is incongruent with their partisan identity, they can strengthen beliefs that are directly contrary to the evidence ([Van Bavel and Pereira, 2018](#)), a phenomenon rooted in dissonance avoidance ([Festinger, 1957](#)). In the context of COVID-19 messages, studies have shown that partisan identity eclipses life-saving, scientifically sound information in driving attitudes, beliefs and preventive behavior ([Gadarian et al., 2021](#)). To no surprise, responses to COVID-19 fact-checking are also influenced by motivated reasoning ([Lu and Zhong, 2022](#)). Originating from the political realm during the mid-2000s ([Szewach et al., 2022](#)), the murkiness of separating political opinion from fact sparked mostly Republican hostilities toward fact-checkers ([Graves, 2016](#); [Walter et al., 2020](#)). By the outbreak of COVID-19, these corrective mechanisms were widely seen as tools of cancel culture to gag

political expression. Partisan antagonism toward fact-checking persisted despite scientific evidence, leading to low fact-checking efficacy (Krause *et al.*, 2020; Saltz *et al.*, 2021).

Second, predispositions including cognitive style (Meyer *et al.*, 2024), media literacy (Hameleers, 2022) and institutional trust (Goldstein and Wiedemann, 2022) have been investigated as correlates of partisanship in resisting debunking systems.

### Partisan divides in debunk credibility perceptions

Meta-analyses reveal that debunks, also referred to as corrective information (Hameleers, 2024), content moderation and veracity or contextual labeling (Morrow *et al.*, 2022), are generally effective. Citizen beliefs become more factually consistent, even after a single exposure to fact-checking (Walter *et al.*, 2020; Zhao *et al.*, 2023), but the format, content and mechanisms of debunks affect their immediate and long-term effectiveness (Chan *et al.*, 2017; Ecker *et al.*, 2020).

Most pertinent to our study, credibility perceptions of debunking sources are associated with partisanship, evidenced by deep divides in US public opinion about content regulation on social media platforms. Compared to Democrats and Independents, a higher percentage of Republicans perceive platforms as being too tough in removing content and biased in censoring conservative over liberal political views (Knight Foundation and Gallup, 2020; Vogels, 2022). Relatedly, Republican elites and activists are more likely to share unsubstantiated information in the first place, necessitating more debunks of their claims (Garrett and Bond, 2021; Moseleh *et al.*, 2024). Yet, Republicans often interpret the disproportionate fact-checking responses to their posts as fact-checkers discriminating against conservative ideology (Graves, 2016). As a result, US liberals show greater awareness of the positive effects and find more use for fact-checking than conservatives (Robertson *et al.*, 2023). We thus expect that Republicans will be more likely than Democrats and Independents to perceive falsehoods as truthful despite a fact-check warning label attached to them.

As a baseline comparison, a debunk message from a typical social media user is less likely to provoke a partisan response. Whereas fact-checkers are politically polarizing, individual users on Facebook should not be, as long as they do not exhibit explicit partisan affiliations. Furthermore, related research on fact-checking messages that convey crowdsourced user consensus suggests that such messages can reduce backlash against debunked messages (Giudice, 2010) and even outperform scientific expert sources in reducing partisan motivated reasoning (Moon *et al.*, 2023). We thus expect that Republicans will be as likely as Democrats and Independents to perceive user-authored debunking messages as truthful. Based on what we know about partisanship, attitudes about fact-checkers and user debunks, we propose the first hypothesis:

- H1.* The falsehood debunking source (W) would moderate the relationship between political partisanship (X) and truth resistance (Y) such that (a) user-debunked messages will elicit similar levels of truth resistance among Republicans and non-Republicans and (b) Republicans will have higher truth resistance to messages debunked by independent fact-checkers than non-Republicans.

### Can cognitive predispositions moderate partisan-based perceptions of message veracity?

The COVID-19 pandemic provided a unique opportunity to study individual-level vulnerability to falsehoods. Here we pursue three potential moderators (Active Open-minded Thinking about Evidence, media literacy and trust in public health agencies) of the Republican tendency to view COVID-19 falsehoods as more credible than Democrats – even when the information is debunked.

*Active open-minded thinking about evidence (AOT-E)*

Thinking styles can offer a reliable explanation for individual variance in susceptibility to unreliable information. The AOT-E index measures openness to evidence that contradicts existing beliefs (Baron, 2019). High openness to belief-oppositional evidence (Pennycook and Rand, 2019) buffers against motivated reasoning and political loyalties (Stanovich and Toplak, 2019). Conversely, low AOT-E is associated with belief rigidity and motivated reasoning tendencies to safeguard existing beliefs (Haran et al., 2013).

The effectiveness of AOT-E as a cognitive buffer against false information (Pennycook and Rand, 2019; Pennycook et al., 2023) was demonstrated in the context of controversial political and scientific issues. These findings cast doubt about the strength of partisanship as a driver of motivated reasoning. High AOT-E was identified as a more reliable predictor of resistance to false information beliefs than partisanship (Pennycook and Rand, 2019; Erlich et al., 2023). Yet, there are also reports of no such effect (Eichmeier and Stenhouse, 2019) or even of fortified beliefs in falsehoods at high AOT-E levels (Kahan and Corbin, 2016).

In the COVID-19 realm, mind habits like heuristic processing (Kim and Kim, 2020), low analytical thinking (Gligorić et al., 2021) and cognitive rigidity (Meyer et al., 2024) predicted belief in COVID-19 conspiracies. The only known experiment that examined thinking style as a moderator of political identity effects on COVID-19 falsehood beliefs did not produce significant results (Lee et al., 2023a). Although there are no known studies examining the intersection of debunking source, partisanship and AOT-E, there is evidence that debunking methods (correction strength and explanatory depth) mattered less than AOT-E in accepting corrective messages about falsehoods (Martel et al., 2021). Grounded in literature on AOT-E's moderation of partisan-based motivated reasoning and the expectations of H1, we formulated the following research question, visualized in Figure 1:

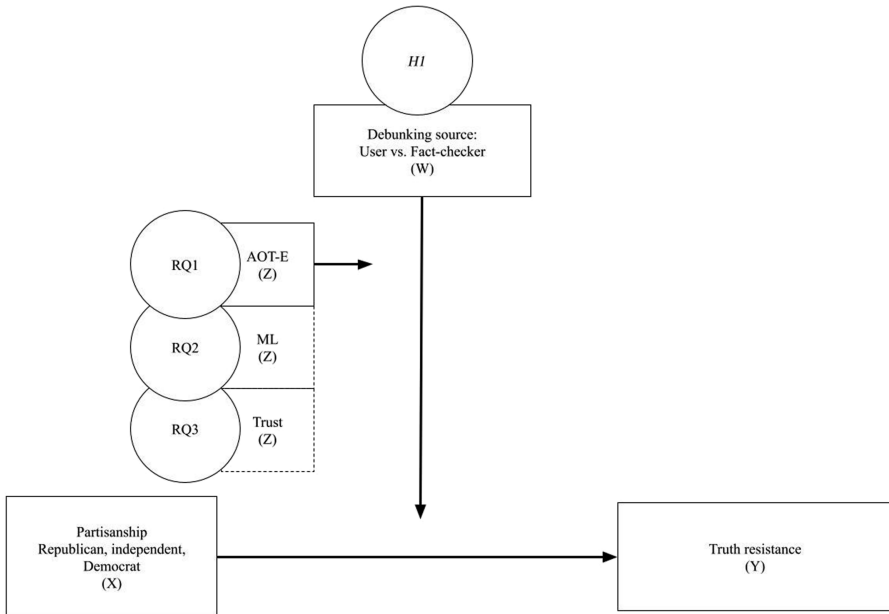


Figure 1. Visual map of hypotheses

RQ1. Will AOT-E levels moderate the moderated relationship of W on XY such that partisan differences in truth resistance will decrease for both debunking sources at high compared to low AOT-E levels?

### *Media literacy (ML)*

There is growing public support for primary and secondary curricula to include critical skills to navigate the information-rich contemporary media environment (Jolls, 2015; Bulger and Davison, 2018). This encouragement is driven by evidence that ML skills enhance the ability to distinguish reliable from misleading or false content.

In the fast-growing body of ML research, the construct has been explicated as encompassing awareness and perceived importance dimensions (Koc and Barut, 2016; Vraga et al., 2015). Perceived importance contributes significantly to scrutiny of information (Vraga et al., 2015). Consequently, individuals with high ML tend to have better discernment between true and false information (Altay et al., 2024; Chan, 2024), reducing the influence of partisanship. Su et al. (2022) found that self-perceived ML significantly weakened the main and indirect effects of social media news use on COVID-19 misperceptions when controlling for political orientation.

How individuals with different ML levels respond to varying debunking sources remains unclear. Yet, higher ML levels have been linked to more fact-checking engagement (Lee and Ramazan, 2021). In terms of receptivity to debunking messages, Hameleers (2022) found that the combination of ML interventions and fact-checks effectively counters agreement with fake news. It is possible that self-perceived ML not only improves the detection of falsehoods but also encourages open-mindedness to corrective content, potentially buffering the partisan divide. Thus, we expect that high ML would mitigate partisan differences in resisting falsehood corrections. Yet, short on known existing evidence for this expectation, we posed the following research question:

RQ2. Will ML levels moderate the moderated relationship of W on XY such that partisan differences in truth resistance would decrease for both debunking sources at high compared to low ML levels?

### *Trust in public health agencies*

Healthy skepticism increases online information verification behavior (Fletcher and Park, 2017), but a detrimental consequence of distrust is belief in conspiracies and falsehoods (Albarracín, 2022). Specifically, low trust in science and scientists is associated with a higher likelihood of believing COVID-19 conspiracies in the United States (Agle and Xiao, 2021), and low trust in health institutions is associated with higher vaccine hesitancy in the United Kingdom, driven in part by the belief that the virus was lab-made (Jennings et al., 2021).

It is well-documented that Republicans and Democrats exhibit asymmetric levels of trust in health and scientific institutions. Republicans show greater distrust in science and traditional information sources compared to other political identity groups, highlighting a significant partisan divide (Gauchat, 2012; Tyson and Kennedy, 2023). By contrast, Democrats are significantly more likely to have confidence in science compared to other partisans (Lee, 2021). Moreover, low institutional trust is one explanation for why Republican-leaning counties recorded lower compliance with stay-at-home orders than Democrat-leaning counties (Goldstein and Wiedemann, 2022). On the flipside of this asymmetry, though, lies the potential that building up trust may also bridge the partisan gap. Despite strong correlations between partisanship and COVID-19 beliefs, institutional trust tends to be the stronger predictor of health-promoting outcomes (Goldstein and Wiedemann, 2022). Given that trust may bridge partisan differences in COVID-19 attitudes, we expect it to also bridge partisan gaps in truth resistance:

RQ3. Will institutional trust levels moderate the moderated relationship of W on XY such that partisan differences in truth resistance will decrease for both debunking sources at high compared to low trust levels?

**Materials and methods**

We fielded a web-based experiment via Qualtrics on a sample of US adults who are active Facebook users. The study was approved by the Indiana University Institutional Review Board. Data collection occurred in January 2023.

**Sample.** A nationally proportionate sample of 692 recruited participants was reduced to 454, using three exclusion criteria: speeding (below four-minute completion), responding with no variation across batteries with six or more items and incompleteness. Participants were between 18 and 60 ( $M = 40.1, SD = 9.9$ ) years old; 54% self-identified as female and 0.4% as nonbinary. Race proportions were 63.7% White, 8.8% Black, 11.9% Hispanic or Latino, 4.0% Asian, 9.3% multi-racial and 0.7% other. High school or less was the highest level of education for 44.7%.

**Design.** In a within-subjects experiment, we manipulated debunking source at two levels: a fact-check warning label on a COVID-19 falsehood (called fact-checker debunk) versus a content rebuttal of the falsehood by a platform user (called user debunk). This factor was used as the key moderator (W) in subsequent models. Party identity was used as the independent variable (X), predicting truth resistance (Y). In addition, three separate models were tested to assess the potential of AOT-E, ML and trust in public health agencies as moderators (Z) of debunking methods (W).

**Stimuli.** We created mock Facebook posts to represent the user versus fact-checker (see Figure 2) debunk conditions. The user debunk condition featured a rebuttal of a widely trafficked falsehood, whereas the fact-checker debunk condition presented the falsehood with a Facebook-consistent fact-check warning. There were eight topical posts, each bimodal, containing text and a graphic. Each participant saw one version of all eight posts, resulting in four true (user debunk) and four false (fact-checker debunk) post

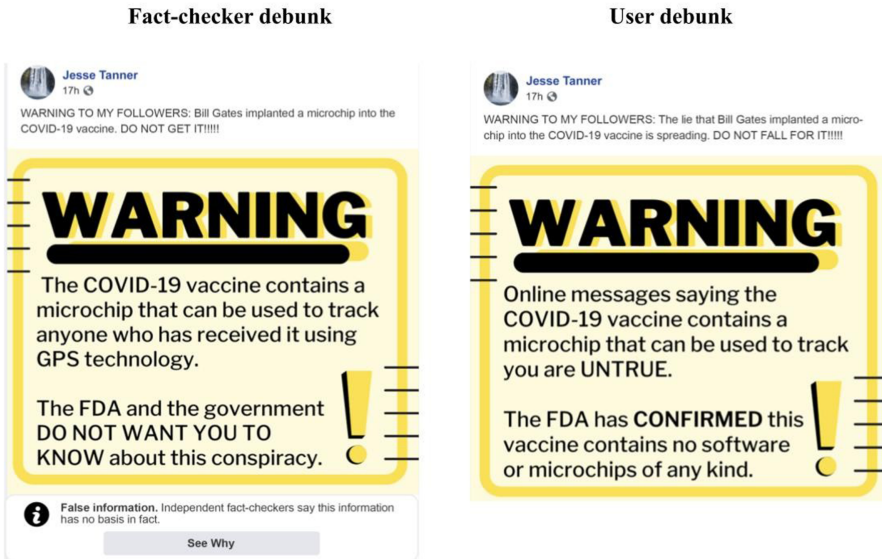


Figure 2. Sample of experimental manipulations

exposures. Order effects were controlled through four quasi-random stimulus orders and random assignment of participants to an order. To avoid the confounds of user characteristics, the fictional Facebook profile featured an androgynous name with a waterfall as the avatar. An *a priori* check ( $N = 181$ ) confirmed that the fact-checker debunk labels were noticed. A post-hoc independent samples *t*-test confirmed that participants perceived the rebuttal of falsehoods in a user debunk ( $M = 2.01$ ,  $SD = 0.75$ ) as more truthful than falsehoods debunked with fact-checker warning labels ( $M = 1.66$ ,  $SD = 0.73$ ),  $t(3,630) = 14.47$ ,  $p < 0.001$ .

**Measures.** Participants self-reported partisanship and other demographic information. Following Pennycook *et al.* (2020), we used an eight-item AOT-E battery. As an additive index (range = 8–48,  $M = 30.01$ ,  $SD = 6.35$ ), it achieved acceptable reliability ( $\alpha = 0.69$ ). ML was measured as a self-perceived skill and rated for societal importance, drawing from Vraga *et al.* (2015), Cheng and Chen (2020), and Koc and Barut (2016). The additive index (range = 5–25,  $M = 19.20$ ,  $SD = 3.79$ ) was reliable ( $\alpha = 0.82$ ). Trust in two public health agencies was also reliable (Spearman-Brown  $\rho = 0.92$ ) as an additive index (range = 0–20,  $M = 10.43$ ,  $SD = 6.35$ ). Perceived message veracity was measured with a three-point semantic differential response (range = 0–2) after exposure to each stimulus and subjected to score standardization. The resulting truth resistance index (range = 0–2) for each condition represented complete acceptance of truthful information coded as 0 and complete acceptance of untrue information as 2. Thus, for posts debunked with fact-check labels, truth resistance equaled the perceived veracity of the post; for user debunk posts, it equaled the reverse-coded veracity rating. Truth resistance served as the main dependent variable.

**Procedure.** After a set of demographic questions, each participant was exposed to eight stimuli and self-reported perceived message veracity after each stimulus. Trust in public health agencies was measured between the sixth and seventh stimulus, and AOT-E and ML after the last stimulus.

## Results

Moderation models were constructed to test H1 and find answers for three research questions. In addition, these models were robust when we subjected them to covariate inclusion, controlling for vaccine hesitancy, COVID-19 status and knowledge about vaccines. Those results are reported in the [supplementary materials \(Section 4\)](#).

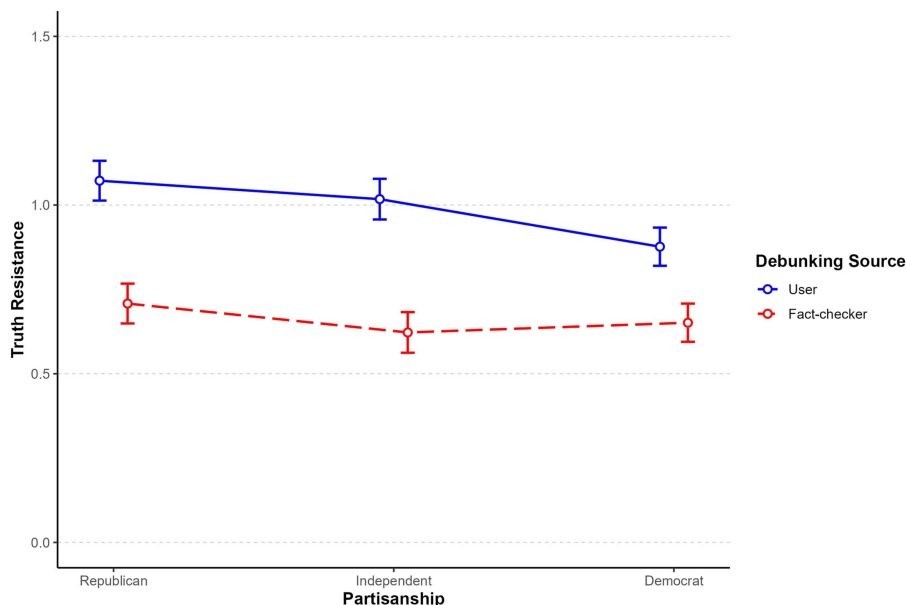
### *Debunk method as a moderator of partisan influence on truth resistance (H1)*

H1 proposed that the debunking source would moderate the association between partisanship and truth resistance, such that Republicans, Democrats and independents would exhibit (1) similar levels of truth resistance toward user-debunked messages and (2) higher truth resistance toward fact-checker-debunked messages compared to non-Republicans. A moderation model (PROCESS Model #1; Hayes, 2022) was tested with Republican partisanship as the reference group for indicator coding (Table 1). Overall, truth resistance was lower for fact-checker debunks and higher for user-generated debunks regardless of partisanship. There was a significant interaction between partisanship and debunking source on truth resistance. Conditional indirect effect analysis suggested that, counter to H1a, Republicans demonstrated higher truth resistance than Democrats for user-debunked messages ( $b = -0.196$ ,  $p < 0.001$ ). For fact-checker debunked messages, Republicans reported slightly higher truth resistance than independents ( $b = -0.086$ ,  $p = 0.0465$ ), but did not differ from Democrats ( $b = -0.057$ ,  $p = 0.173$ ). Therefore, H1b is partially supported. Two key findings are evident from Figure 3: (1) Republicans were more truth-resistant than Democrats to user-generated debunks of false content and (2) Republicans were more truth-resistant than independents to fact-checker debunked false content.

**Table 1.** Moderation analyses predicting perceived message veracity

Model Predictors	Partisanship × debunk source ×			AOT-E			ML			Trust		
	b	SE	Bootstrapped 95%CI <sup>c</sup>	b	SE	Bootstrapped 95%CI <sup>c</sup>	b	SE	Bootstrapped 95%CI <sup>c</sup>	b	SE	Bootstrapped 95%CI <sup>c</sup>
(Intercept)	1.072***	0.030	[1.013–1.131]	1.271***	0.151	[0.929 – 1.601]	1.326***	0.156	[1.004 – 1.649]	1.426***	0.048	[1.335 – 1.513]
Partisanship <sup>a</sup>	–0.055	0.043	[–0.140 – 0.029]	0.200	0.209	[–0.245 – 0.649]	0.532*	0.221	[0.101 – 0.969]	–0.131	0.074	[–0.268 – 0.006]
Partisanship <sup>b</sup>	–0.196***	0.042	[–0.278 – –0.112]	–0.016	0.202	[–0.452 – 0.432]	0.303	0.218	[–0.123 – 0.734]	–0.196*	0.086	[–0.369 – –0.024]
Debunk Source	–0.364***	0.043	[–0.447 – –0.280]	–0.017	0.213	[–0.488 – 0.450]	–0.921***	0.221	[–1.380 – –0.470]	–0.699***	0.068	[–0.837 – –0.564]
Partisanship <sup>a</sup> × Debunk Source	–0.031	0.061	[–0.148 – 0.088]	0.235	0.296	[–0.370 – 0.848]	0.259	0.313	[–0.353 – 0.864]	0.037	0.105	[–0.161 – 0.238]
Partisanship <sup>b</sup> × Debunk Source	0.139*	0.059	[0.022–0.256]	0.893**	0.285	[0.286 – 1.495]	0.132	0.308	[–0.493 – 0.752]	0.406***	0.122	[0.156 – 0.659]
AOT-E/ML/Trust score				–0.007	0.005	[–0.018 – 0.005]	–0.013	0.008	[–0.031 – 0.004]	–0.041***	0.005	[–0.050 – –0.033]
Partisanship <sup>a</sup> × AOT-E/ML/Trust Score				–0.008	0.007	[–0.023 – 0.007]	–0.032**	0.012	[–0.055 – –0.009]	0.012	0.007	[–0.001 – 0.025]
Partisanship <sup>b</sup> × AOT-E/ML/Trust Score				–0.005	0.007	[–0.020 – 0.009]	–0.024*	0.011	[–0.046 – –0.002]	0.014*	0.007	[0.001 – 0.028]
Debunk Source × AOT-E/ML/Trust Score				–0.012	0.007	[–0.028 – 0.004]	0.029*	0.011	[0.005 – 0.053]	0.039***	0.006	[0.027 – 0.052]
Partisanship <sup>a</sup> × Debunk Source × AOT-E/ML/Trust Score				–0.008	0.010	[–0.027 – 0.012]	–0.015	0.016	[–0.047 – 0.018]	–0.011	0.010	[–0.030 – 0.008]
Partisanship <sup>b</sup> × Debunk Source × AOT-E/ML/Trust Score				–0.024*	0.010	[–0.044 – –0.004]	–0.001	0.016	[–0.034 – 0.031]	–0.034***	0.010	[–0.054 – –0.015]
F Statistics	40.915***			39.606***			26.008***			34.539***		
R <sup>2</sup> /Adj. R <sup>2</sup>	0.053/0.052			0.107/0.105			0.073/0.070			0.095/0.092		

**Note(s):** \* $p < 0.05$  \*\* $p < 0.01$  \*\*\* $p < 0.001$ , <sup>a</sup>Republican vs. Independent, <sup>b</sup>Republican vs. Democrats, <sup>c</sup>Bootstrap results are based on 10,000 bootstrap resamples



**Figure 3.** Effects of partisanship by debunking source on truth resistance

#### *Active open-minded thinking about evidence (AOT-E) as moderator (RQ1)*

Three moderated moderation models, using the PROCESS macro (Model #3) (Hayes, 2022), were used to find answers for RQ1-3, with Republican partisanship entered as the reference group (Table 1). RQ1 asked whether AOT-E would moderate the interaction between partisanship and debunking source in predicting truth resistance, with the expectation that higher AOT-E might reduce partisan differences.

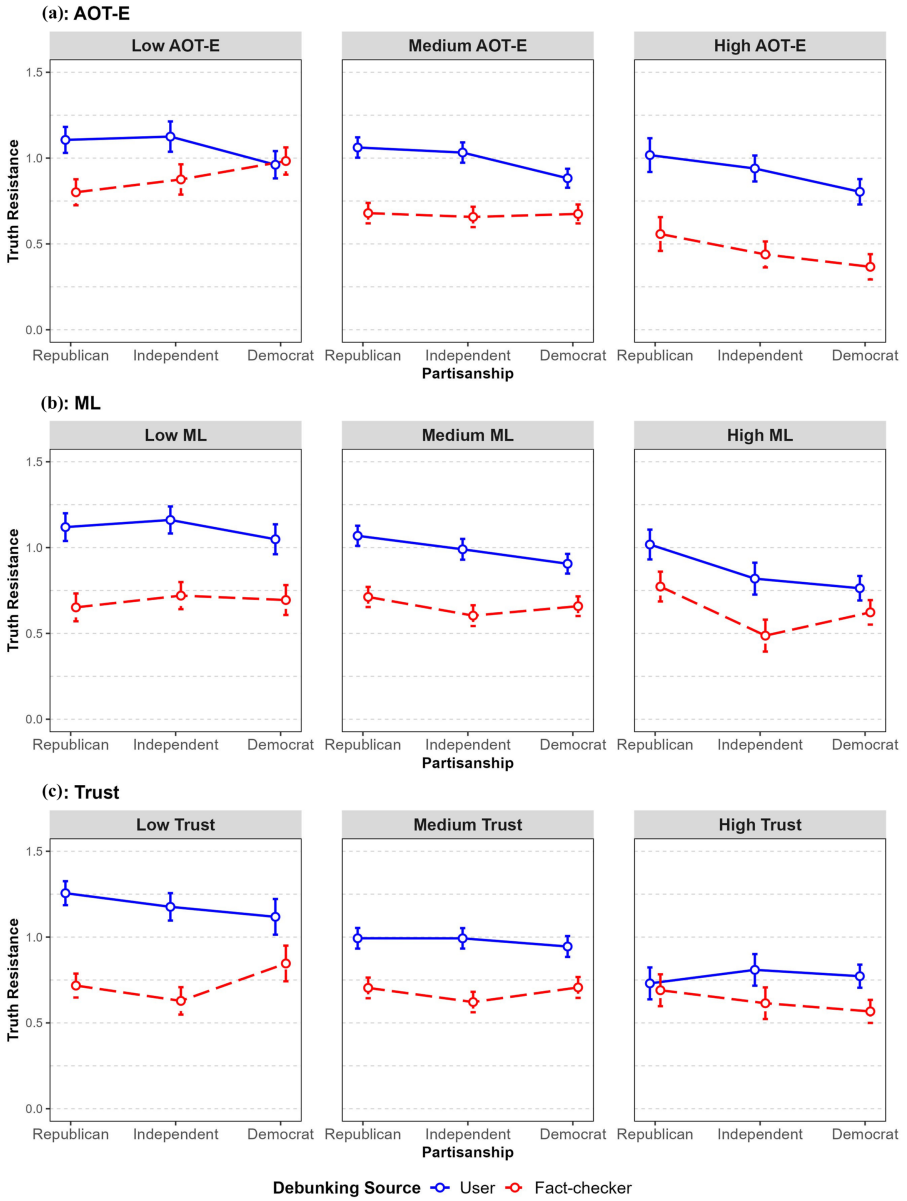
The three-way interaction between partisanship, debunking source and AOT-E was significant. However, the results indicate that higher AOT-E scores were associated with stronger, not weaker, partisan bias (Figure 4a). Conditional effect analysis suggested that, for the user debunk condition, Republicans had consistently higher truth resistance scores than Democrats, and this gap widened as AOT-E increased: low AOT-E ( $b = -0.145, p = 0.010$ ), medium ( $b = -0.179, p < 0.001$ ) and high ( $b = -0.214, p < 0.001$ ). For fact-checker debunks, Democrats had higher truth resistance than Republicans at low levels ( $b = 0.182, p = 0.001$ ), no difference at the mean ( $b = -0.005, p = 0.912$ ) and significantly lower truth resistance than Republicans at high AOT-E ( $b = -0.191, p = 0.002$ ). Thus, contrary to the direction of expectation underlying RQ1, high AOT-E does not reduce partisan divides for either debunking source.

#### *Media literacy (ML) as moderator (RQ2)*

RQ2 posited that ML would moderate the indirect influence of the debunking source on the relationship between partisanship and truth resistance. However, the three-way interaction between partisanship, debunking methods and ML was not significant (Table 1), despite a trend (Figure 4b) and some significant paired results reported in Table S3 that appear to suggest that media literacy is associated with larger partisan differences for fact-checker than user debunked posts. There is no support for ML as a moderator of the moderated relationship.

#### *Trust in public health agencies as moderator (RQ3)*

RQ3 suggested that trust in public health agencies would moderate the indirect influence of the debunking source on the relationship between partisanship and truth resistance. The three-way



**Figure 4.** Three-way interaction effects of partisanship, debunking source and three cognitive predispositions on truth resistance

interaction between partisanship, debunking source and trust was significant (Table 1). Conditional indirect effect analysis revealed that, for the user debunk condition, Republicans showed higher truth resistance than Democrats at low levels of trust ( $b = -0.138, p = 0.031$ ), but this difference diminished at medium ( $b = -0.048, p = 0.272$ ) and high trust levels ( $b = 0.042, p = 0.474$ ). For the fact-checker-debunk condition, Democrats showed higher truth resistance than Republicans at low trust levels ( $b = 0.129, p = 0.044$ ), but the difference

disappeared at medium trust ( $b = 0.003, p = 0.952$ ), and reversed at high trust ( $b = -0.123, p = 0.035$ ). Our data support the idea that high levels of trust mitigate partisan differences in trust resistance for user but not fact-checker debunked posts.

## Discussion

We investigated the potential of two debunking sources (user and fact-checker) and three cognitive predispositions (AOT-E, ML and trust in public health agencies) to buffer partisan divides in COVID-19 debunk receptivity. A demographically proportionate sample of US adults was exposed to eight Facebook posts that mentioned COVID-related falsehoods, half with a fact-checker debunk warning label and the other half with scientifically rooted, text-based debunking of these falsehoods by a Facebook user. Surprisingly, we found that AOT-E exacerbated partisan differences in responding to corrections. ML did not bridge partisan divides, and trust in public health agencies was the only buffer with qualified effectiveness. In this section, we begin by discussing each substantive finding, followed by discussions of limitations and future work. We conclude by highlighting the general theoretical and practical implications of our findings.

First, our results provide evidence that Republicans are more truth-resistant than Democrats, especially in the user debunk condition across all levels of AOT-E and ML, and among low-trust individuals. Moreover, results from the fact-checker debunk condition offer indirect evidence for the same conclusion. Specifically, the truth resistance of high AOT-E Republicans declined marginally compared to a significant drop for high-AOT-E Democrats. Furthermore, low- and high-trust Republicans showed no discernible difference in truth resistance, again a clear contrast to Democrats, whose truth resistance declined significantly across three levels of trust. Finally, none of the cognitive predispositions we tested reliably reduced the truth resistance of Republicans, indicating the entrenchment of such resistance. Yet, the fact-checker debunk condition also showed that for low AOT-E and low trust groups, Democrats can be more truth-resistant than Republicans and at mean levels, partisan differences tend to disappear.

The finding that AOT-E increased partisan divides deserves consideration. High AOT-E participants, regardless of partisanship, were less truth-resistant than low AOT-E partisans (see [Figure 4a](#)). However, the drop was significantly smaller among Republicans than Democrats, particularly in the fact-checker debunk condition, resulting in unequal reductions of truth resistance. As a result, higher AOT-E was associated with greater partisan divide rather than convergence. It is possible that conservatism compels Republicans to ground beliefs in interpersonal hierarchies (social groups) rather than intrapersonal mental processes ([Jost et al., 2003](#)). In addition, the impact of AOT-E varies across belief domains and polarization levels. Republicans who are more open-minded may nevertheless be closed-minded for topic domains they view as exceptional ([Baron, 2017](#)), especially one as politicized and as central to their political identity and beliefs as the COVID-19 pandemic ([Kahan, 2013](#)). It will require more research to tease out the impact of AOT-E on truth resistance across debunk sources and issues that vary in political polarization.

Media literacy was not a significant moderator of partisanship's association with truth resistance for either debunking source. This is not to say that ML does not matter as a defense against misinformation – much work has established that high ML lowers susceptibility to conspiracy theories (e.g. [Craft et al., 2017](#)) and increases one's ability to tell truthful from false news content ([Guess et al., 2020](#)). Nevertheless, our findings suggest that ML does not reduce gaps in how Democrats and Republicans process fact-checking messages. Those seeking to bridge partisan gaps in misinformation processing, in addition to reducing misinformation susceptibility, may find ML to have limited effectiveness in contexts that are highly politicized.

In summary, high AOT-E and ML predispositions were not associated with decreases in partisan differences in truth resistance for either debunking source. Thus, health science

communicators seeking to bridge partisan divides on vaccine information should not count on AOT-E or ML as mechanisms for accomplishing that goal. Yet, unlike AOT-E and ML, trust in public health agencies effectively buffered partisan divides on truth resistance for user-debunked posts. This finding supports health communication researchers who see trust in public health institutions as an antidote to health misinformation (Swire-Thompson and Lazer, 2020; Zhang *et al.*, 2024). Yet, advancing trust in public health after the polarizing impact of COVID-19 is challenging. As Goldberg (2012) argued, public health policy-making is inherently political. Confidence in health agencies relies on two critical factors: the effective communication of scientific findings that inform policy (Blendon and Benson, 2022) and the consistency of messaging from political figures, whose comments can significantly influence public health beliefs (Jones-Jang and Noland, 2022).

It is also noteworthy that trust's buffering of partisan divides can be attributed to the sharp decline in Republican truth resistance for the user debunk condition as trust levels increase. Figure 4c offers a visualization of this trend. Republican truth resistance to user-debunked information is approximately a half point lower at high than low trust levels, whereas the fact-checker condition is associated with similar levels of truth resistance for Republicans across trust levels. In fact, at high trust levels, Republican truth resistance is at near equal levels for user and fact-checker conditions. These findings suggest indirect evidence of the robustness of Republican resistance to fact-checker debunks compared to Democrats, for whom truth resistance significantly declined as trust levels increased.

The partisan gap in truth resistance among low-trust participants also deserves scrutiny. Similar to AOT-E findings, low-trust Democrats displayed greater truth resistance to fact-checker corrections compared to Republicans and independents (see Figure 4a and c). This trend reverses at high levels of trust (and AOT-E), where Democrats show the least truth resistance. It is possible that the comparatively high truth resistance among low-trust Democrats reflects historical associations between vaccine skepticism and certain left-leaning groups (e.g. outspoken celebrities, Whole Foods Bubbles, Democrat leaning communities) before the COVID-19 pandemic (Motta, 2021). Indeed, pre-2020 anti-vaccine sentiment, such as resistance to the measles vaccine, was not exclusively a partisan issue and was sometimes more prevalent in liberal metropolitan areas (Austin, TX; Seattle, WA). This residual sentiment may still influence low-trust Democrats, particularly in a fact-checker condition, which they may perceive as being tied to health institutions.

Our study's data set did not produce evidence to support either of our predictions regarding user and fact-checker debunks as moderators. Though the results are surprising, we hesitate to conclude that user debunks do, and more importantly, that fact-checker debunks do not bridge partisan gaps. Our ability to generalize from this finding is limited by the debunking source manipulation, which was not fully crossed in the design. Put differently, our stimuli debunked falsehoods but not reliable information. The results show that debunking sources have unique moderating effects on partisan truth resistance, but contrary to the expectation that Republicans would be less receptive to fact-checkers than user debunks, given the partisan animosity toward fact-checking practices. Yet, as Figure 3 shows, partisans were equally responsive to fact-checker debunking but varied significantly on user debunking. What we did not test is the outcome of scientifically sound COVID-19 information debunked with fact-checker-endorsed content labels or user rebuttals. Our decision not to pursue a fully crossed design rests on two motivations. First, there is the internal validity consideration of exhausting participants with exposure to double the number of posts (and responses to them) required by an additional scientific veracity factor. We favored message repetition per experimental condition and message multiplicity (eight different posts) over a fully crossed design. Second, we considered the ethical implications of exposing participants to stimuli that debunk scientifically supported information during a pandemic. A thorough debrief after participation might not have neutralized exposure effects, potentially influencing vaccination decisions with life-threatening outcomes. Given that a non-negligible number of participants would

likely belong to demographic groups vulnerable to false information, we decided against a fully crossed design.

Two other shortcomings offer opportunities for future work. First, we used newly built stimuli to resemble – but not represent directly sourced online content. We decided on stimulus construction to maximize internal validity and to avoid pre-exposure and confounds from user profile features. Yet, trading control for the ecological validity of stimuli invariably raises the question of whether stimuli adequately represent debunking in the wild. Future studies may do well utilizing LLMs to generate falsehoods and debunking messages, using large samples of scraped social media content for training data and prompts. Relatedly, testing human versus AI-generated veracity labels might have theoretical and applied value. Second, we excluded political and scientific knowledge measures in our theoretical model, which [Miller et al. \(2016\)](#) demonstrated are critical in understanding vulnerability to ideologically motivated conspiracy theories. Among US conservatives but not liberals, high political knowledge combined with a distrustful disposition amplifies susceptibility to political conspiracies. Testing the knowledge-trust interaction in the health realm with scientific knowledge and public health trust could offer important insights into partisan responses to falsehoods and debunking methods.

Despite these limitations, our findings advance existing insight into how partisans process fact-checks and offer practical implications for public health communicators. Though much prior work has separately investigated the effects of partisanship (e.g. [Jennings and Stroud, 2023](#)) and cognitive predispositions (e.g. [Lee et al., 2023a, b](#); [Guess et al., 2020](#)) on responses to falsehoods and fact-checking, few prior studies have tested how they interact with each other. Our results show that open-minded thinking and media literacy do not reduce partisan effects on debunking resistance. In fact, these normatively desirable predispositions may invoke rather than buffer partisan-based motivated reasoning. In contrast, trust in public health authorities shows promise in mitigating partisan differences, suggesting that bolstering institutional trust reduces partisanship-based resistance to fact-checking.

Importantly, our findings also highlight that reduced truth resistance does not necessarily translate into reduced partisan polarization. Although higher levels of AOT-E are associated with lower truth resistance among partisan groups, partisan differences widen. This suggests that while some strategies or interventions may reduce truth resistance overall, reducing polarization in truth resistance may require more tailored, group-specific strategies that directly address identity-based motivated reasoning. These findings align with prior observations that the correction of health misinformation is a formidable challenge, especially when the subject matter is politically polarized (e.g. [Chan and Albarracín, 2023](#)). However, we identified trust in public health institutions as a promising buffer for partisan responses to corrective information, highlighting the importance of public health professionals and political elites to strengthen public trust in health authorities. At the same time, evidence of AOT-E and ML as ineffective partisan buffers should caution health communicators and intervention designers. In fact, the intervention practice of module-based ML training may be least effective for Republicans, the group most susceptible to conspiratorial health beliefs, pointing to the need for alternative, trust-centered approaches.

### Acknowledgments

The authors acknowledge Jennifer Hwang, Kevin Mudavadi, Jimmy Ochieng, and Gillian Paxton for contributing to early versions of the manuscript. They would also like to acknowledge Mike Schmierbach and Jay Hmielowski for their helpful feedback on an earlier version of the manuscript, as well as Andreas Veglis and anonymous reviewers for their helpful comments. The data can be made available upon reasonable request to the corresponding author.

### Supplementary material

The supplementary material for this article can be found online.

**References**

- Agley, J. and Xiao, Y. (2021), "Misinformation about COVID-19: evidence for differential latent profiles and a strong association with trust in science", *BMC Public Health*, Vol. 21 No. 1, p. 89, doi: [10.1186/s12889-020-10103-x](https://doi.org/10.1186/s12889-020-10103-x).
- Albarracín, D. (2022), "Processes of persuasion and social influence in conspiracy beliefs", *Current Opinion in Psychology*, Vol. 48, 101463, doi: [10.1016/j.copsyc.2022.101463](https://doi.org/10.1016/j.copsyc.2022.101463).
- Altay, S., De Angelis, A. and Hoes, E. (2024), "Media literacy tips promoting reliable news improve discernment and enhance trust in traditional media", *Communications Psychology*, Vol. 2 No. 1, p. 74, doi: [10.1038/s44271-024-00121-5](https://doi.org/10.1038/s44271-024-00121-5).
- Baron, J. (2017), "Comment on Kahan and Corbin: can polarization increase with actively open-minded thinking?", *Research and Politics*, Vol. 4 No. 1, 2053168016688122, doi: [10.1177/2053168016688122](https://doi.org/10.1177/2053168016688122).
- Baron, J. (2019), "Actively open-minded thinking in politics", *Cognition*, Vol. 188, pp. 8-18, doi: [10.1016/j.cognition.2018.10.004](https://doi.org/10.1016/j.cognition.2018.10.004).
- Blendon, R.J. and Benson, J.M. (2022), "Trust in medicine, the health system and public health", *Dædalus*, Vol. 151 No. 4, pp. 67-82, doi: [10.1162/daed\\_a\\_01944](https://doi.org/10.1162/daed_a_01944).
- Bulger, M. and Davison, P. (2018), "The promises, challenges, and futures of media literacy", *Journal of Media Literacy Education*, Vol. 10 No. 1, pp. 1-21, doi: [10.23860/JMLE-2018-10-1-1](https://doi.org/10.23860/JMLE-2018-10-1-1).
- Cardenal, A.S., Aguilar-Paredes, C., Galais, C. and Pérez-Montoro, M. (2019), "Digital technologies and selective exposure: how choice and filter Bubbles shape news media exposure", *The International Journal of Press/Politics*, Vol. 24 No. 4, pp. 465-486, doi: [10.1177/1940161219862988](https://doi.org/10.1177/1940161219862988).
- Chan, M. (2024), "News literacy, fake news recognition, and authentication behaviors after exposure to fake news on social media", *New Media and Society*, Vol. 26 No. 8, pp. 4669-4688, doi: [10.1177/14614448221127675](https://doi.org/10.1177/14614448221127675).
- Chan, M.P.S. and Albarracín, D. (2023), "A meta-analysis of correction effects in science-relevant misinformation", *Nature Human Behaviour*, Vol. 7 No. 9, pp. 1514-1525, doi: [10.1038/s41562-023-01623-8](https://doi.org/10.1038/s41562-023-01623-8).
- Chan, M.S., Jones, C.R., Hall Jamieson, K. and Albarracín, D. (2017), "Debunking: a meta-analysis of the psychological efficacy of messages countering misinformation", *Psychological Science*, Vol. 28 No. 11, pp. 1531-1546, doi: [10.1177/0956797617714579](https://doi.org/10.1177/0956797617714579).
- Cheng, Y. and Chen, Z.F. (2020), "The influence of presumed fake news influence: examining public support for corporate corrective response, media literacy interventions, and governmental regulation", *Mass Communication and Society*, Vol. 23 No. 5, pp. 705-729, doi: [10.1080/15205436.2020.1750656](https://doi.org/10.1080/15205436.2020.1750656).
- Conway, L.G., Chan, L. and Woodard, S.R. (2020), "Socio-ecological influences on political ideology", *Current Opinion in Psychology*, Vol. 32, pp. 76-80, doi: [10.1016/j.copsyc.2019.06.034](https://doi.org/10.1016/j.copsyc.2019.06.034).
- Craft, S., Ashley, S. and Maksl, A. (2017), "News media literacy and conspiracy theory endorsement", *Communication and the Public*, Vol. 2 No. 4, pp. 388-401, doi: [10.1177/2057047317725539](https://doi.org/10.1177/2057047317725539).
- Ecker, U.K.H., O'Reilly, Z., Reid, J.S. and Chang, E.P. (2020), "The effectiveness of short-format refutational fact-checks", *British Journal of Psychology*, Vol. 111 No. 1, pp. 36-54, doi: [10.1111/bjop.12383](https://doi.org/10.1111/bjop.12383).
- Eichmeier, A. and Stenhouse, N. (2019), "Differences that don't make much difference: party asymmetry in open-minded cognitive styles has little relationship to information processing behavior", *Research and Politics*, Vol. 6 No. 3, 2053168019872045, doi: [10.1177/2053168019872045](https://doi.org/10.1177/2053168019872045).
- Erlich, A., Garner, C., Pennycook, G. and Rand, D.G. (2023), "Does analytic thinking insulate against pro-kremlin disinformation? Evidence from Ukraine", *Political Psychology*, Vol. 44 No. 1, pp. 79-94, doi: [10.1111/pops.12819](https://doi.org/10.1111/pops.12819).
- Festinger, L. (1957), *A Theory of Cognitive Dissonance*, Stanford University Press, Stanford, CA.

- Fletcher, R. and Park, S. (2017), "The impact of trust in the news media on online news consumption and participation", *Digital Journalism*, Vol. 5 No. 10, pp. 1281-1299, doi: [10.1080/21670811.2017.1279979](https://doi.org/10.1080/21670811.2017.1279979).
- Gadarian, S.K., Goodman, S.W. and Pepinsky, T.B. (2021), "Partisanship, health behavior, and policy attitudes in the early stages of the COVID-19 pandemic", *PLoS One*, Vol. 16 No. 4, p. e0249596, doi: [10.1371/journal.pone.0249596](https://doi.org/10.1371/journal.pone.0249596).
- Garrett, R.K. and Bond, R.M. (2021), "Conservatives' susceptibility to political misperceptions", *Science Advances*, Vol. 7 No. 23, eabf1234, doi: [10.1126/sciadv.abf1234](https://doi.org/10.1126/sciadv.abf1234).
- Gauchat, G. (2012), "Politicization of science in the public sphere: a study of public trust in the United States, 1974 to 2010", *American Sociological Review*, Vol. 77 No. 2, pp. 167-187, doi: [10.1177/0003122412438225](https://doi.org/10.1177/0003122412438225).
- Giudice, K.D. (2010), "Crowdsourcing credibility: the impact of audience feedback on Web page credibility: crowdsourcing credibility: the impact of audience feedback on Web page credibility", *Proceedings of the American Society for Information Science and Technology*, Vol. 47 No. 1, pp. 1-9, doi: [10.1002/meet.14504701099](https://doi.org/10.1002/meet.14504701099).
- Gligorić, V., Da Silva, M.M., Eker, S., Van Hoek, N., Nieuwenhuijzen, E., Popova, U. and Zeighami, G. (2021), "The usual suspects: how psychological motives and thinking styles predict the endorsement of well-known and COVID -19 conspiracy beliefs", *Applied Cognitive Psychology*, Vol. 35 No. 5, pp. 1171-1181, doi: [10.1002/acp.3844](https://doi.org/10.1002/acp.3844).
- Goldberg, D.S. (2012), "Against the very idea of the politicization of public health policy", *American Journal of Public Health*, Vol. 102 No. 1, pp. 44-49, doi: [10.2105/AJPH.2011.300325](https://doi.org/10.2105/AJPH.2011.300325).
- Goldstein, D.A.N. and Wiedemann, J. (2022), "Who do you trust? The consequences of partisanship and trust for public responsiveness to COVID-19 orders", *Perspectives on Politics*, Vol. 20 No. 2, pp. 412-438, doi: [10.1017/S1537592721000049](https://doi.org/10.1017/S1537592721000049).
- Gratz, K.L., Richmond, J.R., Woods, S.E., Dixon-Gordon, K.L., Scamaldo, K.M., Rose, J.P. and Tull, M.T. (2021), "Adherence to social distancing guidelines throughout the COVID-19 pandemic: the roles of pseudoscientific beliefs, trust, political party affiliation, and risk perceptions", *Annals of Behavioral Medicine*, Vol. 55 No. 5, pp. 399-412, doi: [10.1093/abm/kaab024](https://doi.org/10.1093/abm/kaab024).
- Graves, L. (2016), *Deciding What's True: The Rise of Political Fact-Checking in American Journalism*, Columbia University Press, New York.
- Gruzd, A. and Mai, P. (2020), "Going viral: how a single tweet spawned a COVID-19 conspiracy theory on Twitter", *Big Data and Society*, Vol. 7 No. 2, 2053951720938405, doi: [10.1177/2053951720938405](https://doi.org/10.1177/2053951720938405).
- Guess, A.M., Lerner, M., Lyons, B., Montgomery, J.M., Nyhan, B., Reifler, J. and Sircar, N. (2020), "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*, Vol. 117 No. 27, pp. 15536-15545, doi: [10.1073/pnas.1920498117](https://doi.org/10.1073/pnas.1920498117).
- Guillory, J.J. and Geraci, L. (2013), "Correcting erroneous inferences in memory: the role of source credibility", *Journal of Applied Research in Memory and Cognition*, Vol. 2 No. 4, pp. 201-209, doi: [10.1016/j.jarmac.2013.10.001](https://doi.org/10.1016/j.jarmac.2013.10.001).
- Hameleers, M. (2022), "Separating truth from lies: comparing the effects of news media literacy interventions and fact-checkers in response to political misinformation in the US and Netherlands", *Information, Communication and Society*, Vol. 25 No. 1, pp. 110-126, doi: [10.1080/1369118X.2020.1764603](https://doi.org/10.1080/1369118X.2020.1764603).
- Hameleers, M. (2024), "Why do social media users accept, doubt or resist corrective information? A qualitative analysis of comments in response to corrective information on social media", *Journalism Studies*, Vol. 25 No. 7, pp. 776-793, doi: [10.1080/1461670X.2024.2340591](https://doi.org/10.1080/1461670X.2024.2340591).
- Haran, U., Ritov, I. and Mellers, B.A. (2013), "The role of actively open-minded thinking in information acquisition, accuracy, and calibration", *Judgment and Decision Making*, Vol. 8 No. 3, pp. 188-201, doi: [10.1017/S1930297500005921](https://doi.org/10.1017/S1930297500005921).
- Hayes, A.F. (2022), *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*, 3rd ed., The Guilford Press, New York.

- Howard, M.C. (2022), "Are face masks a partisan issue during the COVID -19 pandemic? Differentiating political ideology and political party affiliation", *International Journal of Psychology*, Vol. 57 No. 1, pp. 153-160, doi: [10.1002/ijop.12809](https://doi.org/10.1002/ijop.12809).
- Jennings, J. and Stroud, N.J. (2023), "Asymmetric adjustment: partisanship and correcting misinformation on Facebook", *New Media and Society*, Vol. 25 No. 7, pp. 1501-1521, doi: [10.1177/14614448211021720](https://doi.org/10.1177/14614448211021720).
- Jennings, W., Stoker, G., Bunting, H., Valgarðsson, V.O., Gaskell, J., Devine, D., McKay, L. and Mills, M.C. (2021), "Lack of trust, conspiracy beliefs, and social media use predict COVID-19 vaccine hesitancy", *Vaccines*, Vol. 9 No. 6, p. 593, doi: [10.3390/vaccines9060593](https://doi.org/10.3390/vaccines9060593).
- Jia, C., Boltz, A., Zhang, A., Chen, A. and Lee, M.K. (2022), "Understanding effects of algorithmic vs. Community label on perceived accuracy of hyper-partisan misinformation", *Proceedings of the ACM on Human-Computer Interaction*, Vol. 6 CSCW2, pp. 1-27, doi: [10.1145/3555096](https://doi.org/10.1145/3555096).
- Jolls, T. (2015), "The new curricula: propelling the growth of media literacy education", *Journal of Media Literacy Education*, Vol. 7 No. 1, pp. 65-71, doi: [10.23860/jmle-7-1-7](https://doi.org/10.23860/jmle-7-1-7).
- Jones-Jang, S.M. and Noland, C. (2022), "The politicization of health and science: role of political cues in shaping the beliefs of the vaccine-autism link", *Health Communication*, Vol. 37 No. 5, pp. 608-616, doi: [10.1080/10410236.2020.1859723](https://doi.org/10.1080/10410236.2020.1859723).
- Jost, J.T., Glaser, J., Kruglanski, A.W. and Sulloway, F.J. (2003), "Political conservatism as motivated social cognition", *Psychological Bulletin*, Vol. 129 No. 3, pp. 339-375, doi: [10.1037/0033-2909.129.3.339](https://doi.org/10.1037/0033-2909.129.3.339).
- Kahan, D.M. (2013), "Ideology, motivated reasoning, and cognitive reflection", *Judgment and Decision Making*, Vol. 8 No. 4, pp. 407-424, doi: [10.1017/S1930297500005271](https://doi.org/10.1017/S1930297500005271).
- Kahan, D.M. and Corbin, J.C. (2016), "A note on the perverse effects of actively open-minded thinking on climate-change polarization", *Research and Politics*, Vol. 3 No. 4, 2053168016676705, doi: [10.1177/2053168016676705](https://doi.org/10.1177/2053168016676705).
- Kim, S. and Kim, S. (2020), "The crisis of public health and infodemic: analyzing belief structure of fake news about COVID-19 pandemic", *Sustainability*, Vol. 12 No. 23, 9904, doi: [10.3390/su12239904](https://doi.org/10.3390/su12239904).
- Knight Foundation and Gallup (2020), "Free expression, harmful speech and censorship in a digital world", available at: [https://knightfoundation.org/wp-content/uploads/2020/06/KnightFoundation\\_Panel6-Techlash2\\_rprt\\_061220-v2\\_es-1.pdf](https://knightfoundation.org/wp-content/uploads/2020/06/KnightFoundation_Panel6-Techlash2_rprt_061220-v2_es-1.pdf) (accessed 27 June 2025).
- Koc, M. and Barut, E. (2016), "Development and validation of new media literacy scale (NMLS) for university students", *Computers in Human Behavior*, Vol. 63, pp. 834-843, doi: [10.1016/j.chb.2016.06.035](https://doi.org/10.1016/j.chb.2016.06.035).
- Krause, N.M., Freiling, I., Beets, B. and Brossard, D. (2020), "Fact-checking as risk communication: the multi-layered risk of misinformation in times of COVID-19", *Journal of Risk Research*, Vol. 23 Nos 7-8, pp. 1052-1059, doi: [10.1080/13669877.2020.1756385](https://doi.org/10.1080/13669877.2020.1756385).
- Kunda, Z. (1990), "The case for motivated reasoning", *Psychological Bulletin*, Vol. 108 No. 3, pp. 480-498, doi: [10.1037/0033-2909.108.3.480](https://doi.org/10.1037/0033-2909.108.3.480).
- Lee, J.J. (2021), "Party polarization and trust in science: what about Democrats?", *Socius: Sociological Research for a Dynamic World*, Vol. 7, 23780231211010101, doi: [10.1177/23780231211010101](https://doi.org/10.1177/23780231211010101).
- Lee, D.K.L. and Ramazan, O. (2021), "Fact-checking of health information: the effect of media literacy, metacognition and health information exposure", *Journal of Health Communication*, Vol. 26 No. 7, pp. 491-500, doi: [10.1080/10810730.2021.1955312](https://doi.org/10.1080/10810730.2021.1955312).
- Lee, S.J., Lee, C.-J. and Hwang, H. (2023a), "The role of deliberative cognitive styles in preventing belief in politicized COVID-19 misinformation", *Health Communication*, Vol. 38 No. 13, pp. 2904-2914, doi: [10.1080/10410236.2022.2125119](https://doi.org/10.1080/10410236.2022.2125119).
- Lee, S.J., Lee, C.-J. and Hwang, H. (2023b), "The impact of COVID-19 misinformation and trust in institutions on preventive behaviors", *Health Education Research*, Vol. 38 No. 1, pp. 95-105, doi: [10.1093/her/cyac038](https://doi.org/10.1093/her/cyac038).

- Lu, S. and Zhong, L. (2022), "From believing to sharing: examining the effects of partisan media's correction of COVID-19 vaccine misinformation", *International Journal of Communication*, Vol. 16 No. 0, p. 22.
- Martel, C., Mosleh, M. and Rand, D.G. (2021), "You're definitely wrong, maybe: correction style has minimal effect on corrections of misinformation online", *Media and Communication*, Vol. 9 No. 1, pp. 120-133, doi: [10.17645/mac.v9i1.3519](https://doi.org/10.17645/mac.v9i1.3519).
- Meta (2021), "Code of practice on disinformation – meta baseline report", available at: <https://disinfocode.eu/reports/download/38> (accessed 25 August 2025).
- Meyer, M., Alfano, M. and De Bruin, B. (2024), "Epistemic vice predicts acceptance of Covid-19 misinformation", *Episteme*, Vol. 21 No. 1, pp. 207-228, doi: [10.1017/epi.2021.18](https://doi.org/10.1017/epi.2021.18).
- Miller, J.M., Saunders, K.L. and Farhart, C.E. (2016), "Conspiracy endorsement as motivated reasoning: the moderating roles of political knowledge and trust", *American Journal of Political Science*, Vol. 60 No. 4, pp. 824-844, doi: [10.1111/ajps.12234](https://doi.org/10.1111/ajps.12234).
- Moon, W.-K., Chung, M. and Jones-Jang, S.Mo. (2023), "How can we fight partisan biases in the COVID-19 pandemic? AI source labels on fact-checking messages reduce motivated reasoning", *Mass Communication and Society*, Vol. 26 No. 4, pp. 646-670, doi: [10.1080/15205436.2022.2097926](https://doi.org/10.1080/15205436.2022.2097926).
- Morrow, G., Swire-Thompson, B., Polny, J.M., Kopec, M. and Wihbey, J.P. (2022), "The emerging science of content labeling: contextualizing social media content moderation", *Journal of the Association for Information Science and Technology*, Vol. 73 No. 10, pp. 1365-1386, doi: [10.1002/asi.24637](https://doi.org/10.1002/asi.24637).
- Mosleh, M., Yang, Q., Zaman, T., Pennycook, G. and Rand, D.G. (2024), "Differences in misinformation sharing can lead to politically asymmetric sanctions", *Nature*, Vol. 634 No. 8034, pp. 609-616, doi: [10.1038/s41586-024-07942-8](https://doi.org/10.1038/s41586-024-07942-8).
- Mosseri, A. (2016), "Addressing hoaxes and fake news", *Metalworking News*, 15 December, available at: <https://about.fb.com/news/2016/12/news-feed-fyi-addressing-hoaxes-and-fake-news/> (accessed 12 July 2025).
- Motta, M. (2021), "Republicans, not Democrats, are more likely to endorse anti-vaccine misinformation", *American Politics Research*, Vol. 49 No. 5, pp. 428-438, doi: [10.1177/1532673X211022639](https://doi.org/10.1177/1532673X211022639).
- Neely, S.R., Eldredge, C., Ersing, R. and Remington, C. (2022), "Vaccine hesitancy and exposure to misinformation: a survey analysis", *Journal of General Internal Medicine*, Vol. 37 No. 1, pp. 179-187, doi: [10.1007/s11606-021-07171-z](https://doi.org/10.1007/s11606-021-07171-z).
- Nyhan, B. and Reifler, J. (2015), "Estimating fact-checking's effects: evidence from a long-term experiment during campaign 2014", available at: <https://search.issuelab-dev.org/resource/estimating-fact-checking-s-effects-evidence-from-a-long-term-experiment-during-campaign-2014> (accessed 25 August 2025).
- Pennycook, G. and Rand, D.G. (2019), "Lazy, not biased: susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning", *Cognition*, Vol. 188, pp. 39-50, doi: [10.1016/j.cognition.2018.06.011](https://doi.org/10.1016/j.cognition.2018.06.011).
- Pennycook, G., Cheyne, J.A., Koehler, D.J. and Fugelsang, J.A. (2020), "On the belief that beliefs should change according to evidence: implications for conspiratorial, moral, paranormal, political, religious, and science beliefs", *Judgment and Decision Making*, Vol. 15 No. 4, pp. 476-498, doi: [10.1017/S1930297500007439](https://doi.org/10.1017/S1930297500007439).
- Pennycook, G., Bago, B. and McPhetres, J. (2023), "Science beliefs, political ideology, and cognitive sophistication", *Journal of Experimental Psychology: General*, Vol. 152 No. 1, pp. 80-97, doi: [10.1037/xge0001267](https://doi.org/10.1037/xge0001267).
- Robertson, C.T., Mourão, R.R. and Thorson, E. (2020), "Who uses fact-checking sites? The impact of demographics, political antecedents, and media use on fact-checking site awareness, attitudes, and behavior", *The International Journal of Press/Politics*, Vol. 25 No. 2, pp. 217-237, doi: [10.1177/1940161219898055](https://doi.org/10.1177/1940161219898055).
- Robertson, R.E., Green, J., Ruck, D.J., Ognyanova, K., Wilson, C. and Lazer, D. (2023), "Users choose to engage with more partisan news than they are exposed to on Google Search", *Nature*, Vol. 618 No. 7964, pp. 342-348, doi: [10.1038/s41586-023-06078-5](https://doi.org/10.1038/s41586-023-06078-5).

- Roozenbeek, J., Schneider, C.R., Dryhurst, S., Kerr, J., Freeman, A.L.J., Recchia, G., Van Der Bles, A.M. and van der Linden, S. (2020), "Susceptibility to misinformation about COVID-19 around the world", *Royal Society Open Science*, Vol. 7 No. 10, 201199, doi: [10.1098/rsos.201199](https://doi.org/10.1098/rsos.201199).
- Ruiz, J.B. and Bell, R.A. (2021), "Predictors of intention to vaccinate against COVID-19: results of a nationwide survey", *Vaccine*, Vol. 39 No. 7, pp. 1080-1086, doi: [10.1016/j.vaccine.2021.01.010](https://doi.org/10.1016/j.vaccine.2021.01.010).
- Saltz, E., Leibowicz, C.R. and Wardle, C. (2021), "Encounters with visual misinformation and labels across platforms: an interview and diary study to inform ecosystem approaches to misinformation interventions", In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21)*, Association for Computing Machinery, New York, NY, USA, doi: [10.1145/3411763.3451807](https://doi.org/10.1145/3411763.3451807).
- Samore, T., Fessler, D.M.T., Sparks, A.M. and Holbrook, C. (2021), "Of pathogens and party lines: social conservatism positively associates with COVID-19 precautions among U.S. Democrats but not Republicans", edited by Capraro, V, *PLoS One*, Vol. 16 No. 6, e0253326, doi: [10.1371/journal.pone.0253326](https://doi.org/10.1371/journal.pone.0253326).
- Schalbe, M.C., Joseff, K., Woolley, S. and Cohen, G.L. (2024), "When politics trumps truth: political concordance versus veracity as a determinant of believing, sharing, and recalling the news", *Journal of Experimental Psychology: General*, Vol. 153 No. 10, pp. 2524-2551, doi: [10.1037/xge0001650](https://doi.org/10.1037/xge0001650).
- Stanovich, K.E. and Toplak, M.E. (2019), "The need for intellectual diversity in psychological science: our own studies of actively open-minded thinking as a case study", *Cognition*, Vol. 187, pp. 156-166, doi: [10.1016/j.cognition.2019.03.006](https://doi.org/10.1016/j.cognition.2019.03.006).
- Su, Y., Borah, P. and Xiao, X. (2022), "Understanding the 'infodemic': social media news use, homogeneous online discussion, self-perceived media literacy and misperceptions about COVID-19", *Online Information Review*, Vol. 46 No. 7, pp. 1353-1372, doi: [10.1108/OIR-06-2021-0305](https://doi.org/10.1108/OIR-06-2021-0305).
- Swire-Thompson, B. and Lazer, D. (2020), "Public health and online misinformation: challenges and recommendations", *Annual Review of Public Health*, Vol. 41 No. 1, pp. 433-451, doi: [10.1146/annurev-publhealth-040119-094127](https://doi.org/10.1146/annurev-publhealth-040119-094127).
- Szewach, P., Reifler, J. and Oscarsson, H. (2022), "Is resistance futile? Citizen knowledge, motivated reasoning, and fact-checking", in Strömbäck, J., Wikfors, Å., Glüer, K., Lindholm, T. and Oscarsson, H. (Eds), *Knowledge Resistance in High-Choice Information Environments*, Routledge, London.
- Tyson, B. and Kennedy, A. (2023), "How Americans view future harms from climate change in their community and around the U.S.", *Pew Research Center*, available at: <https://www.pewresearch.org/science/2023/10/25/how-americans-view-future-harms-from-climate-change-in-their-community-and-around-the-u-s/> (accessed 12 February 2024).
- Van Bavel, J.J. and Pereira, A. (2018), "The partisan brain: an identity-based model of political belief", *Trends in Cognitive Sciences*, Vol. 22 No. 3, pp. 213-224, doi: [10.1016/j.tics.2018.01.004](https://doi.org/10.1016/j.tics.2018.01.004).
- Vraga, E., Tully, M., Kotcher, J., Smithson, A.-B. and Broeckelman-Post, M. (2015), "A multi-dimensional approach to measuring news media literacy", *Journal of Media Literacy Education*, Vol. 7 No. 3, pp. 41-53, doi: [10.23860/jmle-7-3-4](https://doi.org/10.23860/jmle-7-3-4).
- Vogels, E.A. (2022), "Support for more regulation of tech companies has declined in U.S., especially among Republicans", *Pew Research Center*, 13 May, available at: <https://www.pewresearch.org/short-reads/2022/05/13/support-for-more-regulation-of-tech-companies-has-declined-in-u-s-especially-among-republicans/> (accessed 27 June 2025).
- Walter, N., Cohen, J., Holbert, R.L. and Morag, Y. (2020), "Fact-checking: a meta-analysis of what works and for whom", *Political Communication*, Vol. 37 No. 3, pp. 350-375, doi: [10.1080/10584609.2019.1668894](https://doi.org/10.1080/10584609.2019.1668894).
- Ye, X. (2023), "Exploring the relationship between political partisanship and COVID-19 vaccination rate", *Journal of Public Health*, Vol. 45 No. 1, pp. 91-98, doi: [10.1093/pubmed/fdab364](https://doi.org/10.1093/pubmed/fdab364).

- Zhang, J., Featherstone, J.D., Calabrese, C. and Wojcieszak, M. (2021), "Effects of fact-checking social media vaccine misinformation on attitudes toward vaccines", *Preventive Medicine*, Vol. 145, 106408, doi: [10.1016/j.ypmed.2020.106408](https://doi.org/10.1016/j.ypmed.2020.106408).
- Zhang, S., Zhou, H. and Zhu, Y. (2024), "Have we found a solution for health misinformation? A ten-year systematic review of health misinformation literature 2013-2022", *International Journal of Medical Informatics*, Vol. 188, 105478, doi: [10.1016/j.ijmedinf.2024.105478](https://doi.org/10.1016/j.ijmedinf.2024.105478).
- Zhao, S., Hu, S., Zhou, X., Song, S., Wang, Q., Zheng, H., Zhang, Y. and Hou, Z. (2023), "The prevalence, features, influencing factors, and solutions for COVID-19 vaccine misinformation: systematic review", *JMIR Public Health and Surveillance*, Vol. 9, e40201, doi: [10.2196/40201](https://doi.org/10.2196/40201).

**Corresponding author**

Wei Wang can be contacted at: [wei.wang@asc.upenn.edu](mailto:wei.wang@asc.upenn.edu)