Susceptibility to Online Misinformation: A Systematic Meta-Analysis of Demographic and Psychological Factors

Mubashir Sultan[1,2], Alan N. Tump[1,3], Nina Ehmann[1,4], Philipp Lorenz-Spreen[1],
Ralph Hertwig[1], Anton Gollwitzer[1,5], and Ralf H. J. M. Kurvers[1,3]

[1]Max Planck Institute for Human Development, Center for Adaptive Rationality, Berlin, 14195, Germany

[2]Department of Psychology, Humboldt University of Berlin, 12489, Berlin, Germany

[3]Technical University of Berlin, Exzellenzcluster Science of Intelligence, Berlin, 10587, Germany

[4]University of Konstanz, Department of Psychology, Konstanz, 78457, Germany

[5]BI Norwegian Business School, Department of Leadership and Organizational Behaviour, 0484, Oslo, Norway

Corresponding Author:

Mubashir Sultan, Max Planck Institute for Human Development, Center for Adaptive Rationality, Berlin, 14195, Germany, mubashirsultan@outlook.com.

Abstract

Despite a surge of research into misinformation, it is largely unclear who falls for misinformation and why. We conducted a systematic individual participant data meta-analysis covering 256,337 unique choices made by 11,561 participants across 31 experiments. Our meta-analysis reveals the impact of key demographic and psychological factors on online misinformation veracity judgments. We also disentangle the ability to discern between true and false news (discrimination ability) from the response bias, that is, a tendency to label news as either true (true-news bias) or false (false-news bias). We find that older age, higher analytic thinking skills, and identifying as a Democrat are associated with higher discrimination ability. Older age and higher analytical thinking skills are also associated with a false-news bias (caution). In contrast, ideological congruency, motivated reflection, and familiarity are associated with a true-news bias (naïvety). Our results touch upon ongoing debates in the literature and provide critical insights that can help design targeted interventions.

Keywords: misinformation, analytical thinking, partisan bias, illusory truth effect, signal detection theory

Susceptibility to Online Misinformation: A Systematic Meta-Analysis of Demographic and Psychological Factors

Belief in misinformation can have far-reaching consequences—for example, in politics (e.g., the U.S. Capitol riots; Prochaska et al., 2023), health (e.g., vaccine hesitancy; Pinna et al., 2022), and climate change (e.g., resistance to climate-friendly behaviour; Hornsey & Fielding, 2020). Despite the surge of research on misinformation, it is largely unclear who falls for it and why. Previous studies typically investigate single demographic (e.g., age, political identity) or psychological factors (e.g., analytical thinking, ideological congruency) in isolation, often leading to conflicting results. These studies typically employ the widely-used news headline paradigm, where participants evaluate the veracity of news headlines (i.e., a headline, potentially accompanied by a byline and/or an image). Pooling individual participant data from the news headline paradigm—totalling 256,337 unique choices made by 11,561 participants across 31 experiments—we conducted a systematic meta-analysis on how key demographic and psychological factors impact online misinformation veracity judgments. Understanding who falls for misinformation—and why—is crucial for identifying vulnerable groups, designing tailored interventions, and combatting its far-reaching consequences.

We examined four demographic factors that represent important population-level characteristics—age, gender, education, and political identity—and four psychological factors that are considered to play a pivotal role in judging misinformation—analytical thinking (Pennycook & Rand, 2021), ideological congruency (i.e., partisan bias; Van Bavel & Pereira, 2018), motivated reflection (Kahan et al., 2017), and familiarity (i.e., the illusory truth effect; e.g., Lewandowsky et al., 2012). We also explored meta-questions on the effects of the topic of news headlines (e.g., political, health), the study platform (e.g., MTurk), and the source being displayed alongside news headlines. We conducted the meta-analysis using individual participant data (i.e., raw trial-level data) and applied signal detection theory (SDT; Macmillan & Creelman, 2004) to delineate how the above demographic and psychological factors influence two frequently confounded decision-making mechanisms: *discrimination ability*, that is, the ability to distinguish between true and false news, and *response bias*, that is, a tendency to classify news as true (true-news bias) or false (false-news bias). Despite being conflated in much of the literature, these mechanisms can have very different ramifications for theory and intervention building (e.g., individualised interventions; Guay et al., 2023; Higham et al., 2023).

Given our approach, our meta-analysis uniquely allows us to touch upon key debates in the literature: Do older adults, despite their lower digital literacy and propensity to share false news, have more accurate veracity judgments compared to their younger counterparts? What is the effectiveness of formal education with its promises of critical thinking skills for navigating online information environments? Relatedly, what is the potential of analytical thinking skills to improve veracity judgments? What is the influence of political identity in shaping distinct perceptions of truth, including the role of ingroup vs. outgroup favouritism (ideological congruency)? And, finally, how much does familiarity breed belief in misinformation? We discuss each demographic and psychological factor next.

On one end of the age spectrum are "digital natives"—younger generations who were born into the digital age. On the other end are older adults, who entered the digital world later in life and are less digitally literate (Pew Research Center, 2021, 2024). This divide can manifest in the way people judge and share misinformation. The literature indicates that older adults visit more low-quality news websites and share more misinformation compared to younger adults (Grinberg et al., 2019; Guess et al., 2019; Guess et al., 2020b). However, studies suggest that older adults are generally better (e.g., Allcott & Gentzkow, 2017; Brashier & Schacter, 2020; Pennycook & Rand, 2019), or as good as their younger counterparts, at judging the veracity of news (e.g., Pehlivanoglu et al., 2020). Our meta-analysis aims to provide clearer insights into this apparent paradox.

There appear to be gender-based differences in news engagement patterns. Compared to male participants, female participants consume less news (Toff & Palmer, 2019), express lower interest in news (Newman et al., 2023), and are more likely to avoid news (Kalogeropoulos et al., 2020; Newman et al., 2023). Nevertheless, the literature has produced mixed findings on the impact of gender on veracity judgments. Wolverton and Stevens (2019) and Fadhila et al. (2021) report a null effect for gender, whereas Halpern et al. (2019) and Pennycook and Rand (2019; although in only one of their two studies) find that female participants are worse than male participants at judging the veracity of news headlines. Note, however, that Halpern et al. (2019) focus only on false news, making it challenging to distinguish between discrimination ability and response bias. We aim to overcome some of these methodological inconsistencies and address the effect of gender on misinformation veracity judgments.

Formal education teaches people critical thinking skills—challenging assumptions, scrutinising sources, and evaluating information. Multiple studies find that higher formal education correlates with better veracity judgments (e.g., Allcott & Gentzkow, 2017; Wolverton & Stevens, 2019). However, a parallel strand of research posits that traditional education in terms of critical thinking is insufficient for the digital age (e.g., Kozyreva et al., 2022b). Even people with high levels of education (e.g., professors) are swayed by factors such as objective-sounding language and official-looking websites (Breakstone et al., 2022; McGrew et al., 2018). The relationship between education and veracity judgments requires clarification.

Engagement with (mis)information online is heavily influenced by people's political identities, as it can occur in echo chambers (formed by either algorithmic or user-created processes) where people are surrounded by like-minded peers (Barberá et al., 2015; Robertson et al., 2023). Echo chambers and similar structures also play an important role in an ongoing debate over whether the information ecosystems of Democrats and Republicans diverge enough to drive polarisation and shape distinct realities. Various studies report that Democrats outperform Republicans in assessing news veracity (e.g., Bago et al., 2020; Dobbs et al., 2023; Garrett & Bond, 2021; Geers et al., 2024a; Roozenbeek et al., 2022). Our meta-analysis aims to delineate these potentially crucial asymmetries.

Analytical thinking is rooted in dual-process theory (Evans, 2003). In the context of misinformation, it holds that people fail to discriminate between true and false news because they rely on intuitive, albeit erroneous, responses instead of taking a moment to deliberate. Past research has generally found that higher analytical thinking skills predicts better discrimination ability, and in some cases, an increased tendency to treat news as false (i.e., false-news bias; Batailler et al., 2021; Pennycook & Rand, 2021; Sultan et al., 2022).

The concept of ideological congruency is based on ingroup versus outgroup favouritism outlined in social identity theory (Tajfel & Turner, 2004; Van Bavel & Pereira, 2018). According to this account, people treat news content that is congruent (incongruent) with their political identity favourably (unfavourably). The consequence being, people may treat false news as true because it is congruent with their political identity, or treat true news as false because it is incongruent with their political identity. Studies have found that both Democrats and Republicans are more likely to judge news headlines as true (false) when the headlines align (misalign) with their ideology (Batailler et al., 2021; Pennycook & Rand, 2021; Sultan et al., 2022).

Motivated reflection is an interplay between analytical thinking and ideological congruency (i.e., an interaction on the response bias). Individuals with higher analytical thinking skills, known to have higher discrimination ability and a more cautious approach, paradoxically exhibit greater susceptibility to ideological congruency (Kahan et al., 2017; Kunda, 1990). They use their higher analytical thinking skills to rationalise information that aligns (misaligns) to their ideological beliefs as true (false). The literature shows conflicting results for motivated reflection. For example, Sultan et al. (2023) find an effect for motivated reflection on veracity judgments, whereas Batailler et al. (2021) do not (see also Bago et al., 2020; Linden et al., 2018; Persson et al., 2021; Stagnaro et al., 2023).

Familiarity with a news headline is thought to increase the fluency of information processing, and as such, may act as a meta-cognitive cue for accuracy (Ecker et al., 2022; Unkelbach et al., 2019; Wang et al., 2016). Because familiar news is easier to process, it is more likely to be perceived as true. This is a robust effect (for reviews, see Ecker et al., 2022; Udry & Barber, 2024) across different levels of cognitive ability (De Keersmaecker et al., 2020) and prior accurate knowledge (Fazio, 2020; trivia statements), even months after exposure (Brown & Nix, 1996; trivia statements). Familiarity mostly affects response bias but has also been associated with reduced discrimination ability (Batailler et al., 2021; Sultan et al., 2022). All in all, within our meta-analysis, we seek to speak to the robustness of the findings related to the psychological factors and susceptibility to misinformation.

For all the factors above, the distinction between discrimination ability and response bias has been blurry. This is because most studies on misinformation focus on accuracy—summed across true and false news or for true and false news separately. Focusing on accuracy conflates participants' ability to distinguish between true and false news with their general tendency to judge news as true or false (e.g., due to being cautious or naïve, or due to experimental demands). This is highly relevant for designing interventions to combat misinformation. For example, Modirrousta-Galian and Higham (2023) found that an increase in participants' ability to spot misinformation after an intervention was in fact attributable to a higher false-news bias (i.e., a general tendency to judge news as false), not higher discrimination ability (see also; Batailler et al., 2021; Guay et al., 2023; Higham et al., 2023). Effective interventions also require an understanding of whether ideological congruency improves veracity judgments—that is, because people know more about ingroup issues—or whether it shifts people to a true-news bias. For these reasons, taking advantage of the framework of SDT is crucial as it allows us to distinguish between discrimination ability and response bias.

Based on past results, we hypothesised higher discrimination ability for older adults, individuals with higher formal education, and Democrats, and no effect of gender. Due to the conflation of discrimination ability and response bias in previous studies, we had no specific hypotheses for these effects. Turning to the psychological factors, we hypothesised higher discrimination ability for individuals with higher analytical thinking skills and for unfamiliar news headlines. We also hypothesised a higher tendency to classify news headlines as true (true-news bias) when they align with one's political identity (ideological congruency) and for familiarity.

To test our hypotheses, we conducted a preregistered systematic meta-analysis using individual participant data. All studies were U.S.-based and used the headline paradigm, including a non-confounded measure of veracity (i.e., no sharing decisions; for full eligibility criteria, see Methods). This allowed us to pool individual participant data spanning the eight demographic and psychological factors. An individual participant data meta-analysis allows for more detailed analysis compared to an effect-size based meta-analysis. Specifically, we were able to use trial-level raw data to conduct an SDT analysis using a single mixed-effects model across studies. This helped to reduce variation in the statistical methods across studies and enabled us to conduct subgroup analyses, consolidating the demographic and psychological factors into a unified model to assess their *relative* strengths. The mixed-effect model was also able to accommodate the hierarchical nature of the data, helping to account for risk of publication bias by better controlling participant-, study-, and headline-level variability. Overall, our approach provides a comprehensive explanation of how each factor is associated with misinformation veracity judgments.

Results

Study and Participant Characteristics

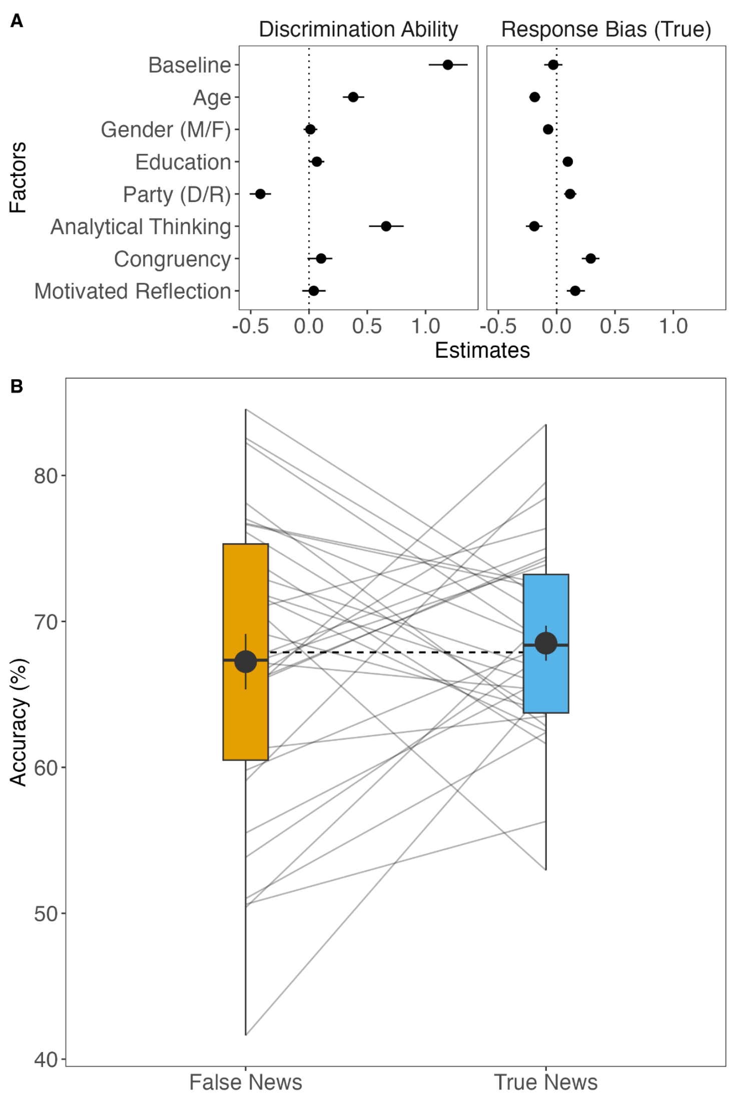
Our search for articles using relevant search terms on databases and via mailing lists returned 4,666 results. Of these, 21 articles encompassing 31 studies were included in the final analysis (for details on the screening process, see Methods). In total, 256,337 unique veracity judgments made by 11,561 participants (M_{age} = 41.29 years, *SD* = 15.68, range = 18–88) were included, comprising 53.91% identifying as female (46.09% male) and 41.85% identifying as Republican (58.15% Democrat).

SDT, Demographic Factors, and Psychological Factors

We conducted the SDT analysis using a Bayesian generalised linear mixed-effects model (for more, see Methods). Regarding interpretation, in the SDT model, higher (lower) discrimination ability is indicated by more positive (negative) values (Figure 1A). For response bias, a higher likelihood to judge headlines as true (false) is indicated by more positive (negative) values (Figure 1A). A true-news response bias refers to a tendency to judge news as true, which can lead to higher accuracy for true news headlines and lower for false news headlines. Likewise, a false-news bias refers to a tendency to judge news as false, which can lead to higher accuracy for false news but lower for true news.

Participants had a discrimination ability credibly higher than zero ($\beta_{\text{Discrimination Ability}}$ = 1.19, 95% credible intervals [CI] = 1.03–1.36; see baseline in Figure 1A). This effect was robust across all studies (Figure 1B; for separate estimates for all studies, see Supplementary Figure 1A). This effect is descriptively visualised in Figure 1B, showing that participants had better-than-chance overall accuracy (67.08%) across all studies and news headlines (dotted line). The baseline response bias was not credibly different from zero ($\beta_{\text{Response Bias}}$ = −0.03, CI = −0.11–0.05; see baseline in Figure 1A). Corroborating this, Figure 1B shows that participants were similarly accurate for true and false headlines, as evidenced by the mean and median values in the boxplots. Although there was no overall response-bias effect, single studies varied widely in their response bias (see individual study lines in Figure 1B), showing a true bias, a false bias, or no bias (for separate estimates for all studies, see Supplementary Figure 1B). In sum, averaged over all studies and factors, participants generally performed better-than-chance and did not have an overall response tendency to treat news as either true or false. We next turn to the results of the demographic and psychological factors, respectively. Note that the credible effects here either increase or decrease the effect of baseline discrimination ability and/or response bias.

Figure 1. Signal detection analysis alongside true and false news accuracy. *Panel A: Signal detection theory (SDT) model estimates. All results derive from a single SDT analysis using participants' responses (false news or true news) as the response variable but are shown in two panels for clarity. Left: Estimates for discrimination ability, with more positive (negative) values indicating higher (lower) discrimination ability. Right: Estimates for response bias, with more positive (negative) values indicating a higher likelihood to judge headlines as true (false). Baseline: Overall estimate of discrimination ability (left) and response bias (right). Gender (M/F) = coded Male to Female. Party (D/R) = political identity, coded Democrat to Republican. Congruency = ideological congruency. Dots represent the mean; error bars represent the 95% CI of the posterior distribution. All factors were mean centred. Panel B: Accuracy for true and false news headlines. Grey lines connect the mean accuracy scores for false and true news headlines within a study. Boxplots show the median and the interquartile range (IQR); whiskers indicate an additional 1.5 IQR. Large dots represent the aggregate mean with standard errors; dashed black line represents the aggregate mean across all studies.*



Age

Age had a positive effect on discrimination ability (β = 0.38, CI = 0.29–0.47; Figure 1A), indicating that older participants achieved higher overall accuracy (Figure 2A). Moreover, we also found a negative and credible effect of age on response bias (β = −0.19, CI = −0.24–−0.14]; Figure 1A), suggesting that with increasing age, individuals were generally more likely to judge a news headline as false. This resulted in older adults having higher accuracy for false news than true news (Figure 2A; see Supplementary Figure 2A, B for these effects per study). Being older was thus associated with better discrimination ability and a higher overall false-news bias, which explains the widening gap in true and false news accuracy with higher age (Figure 2A).

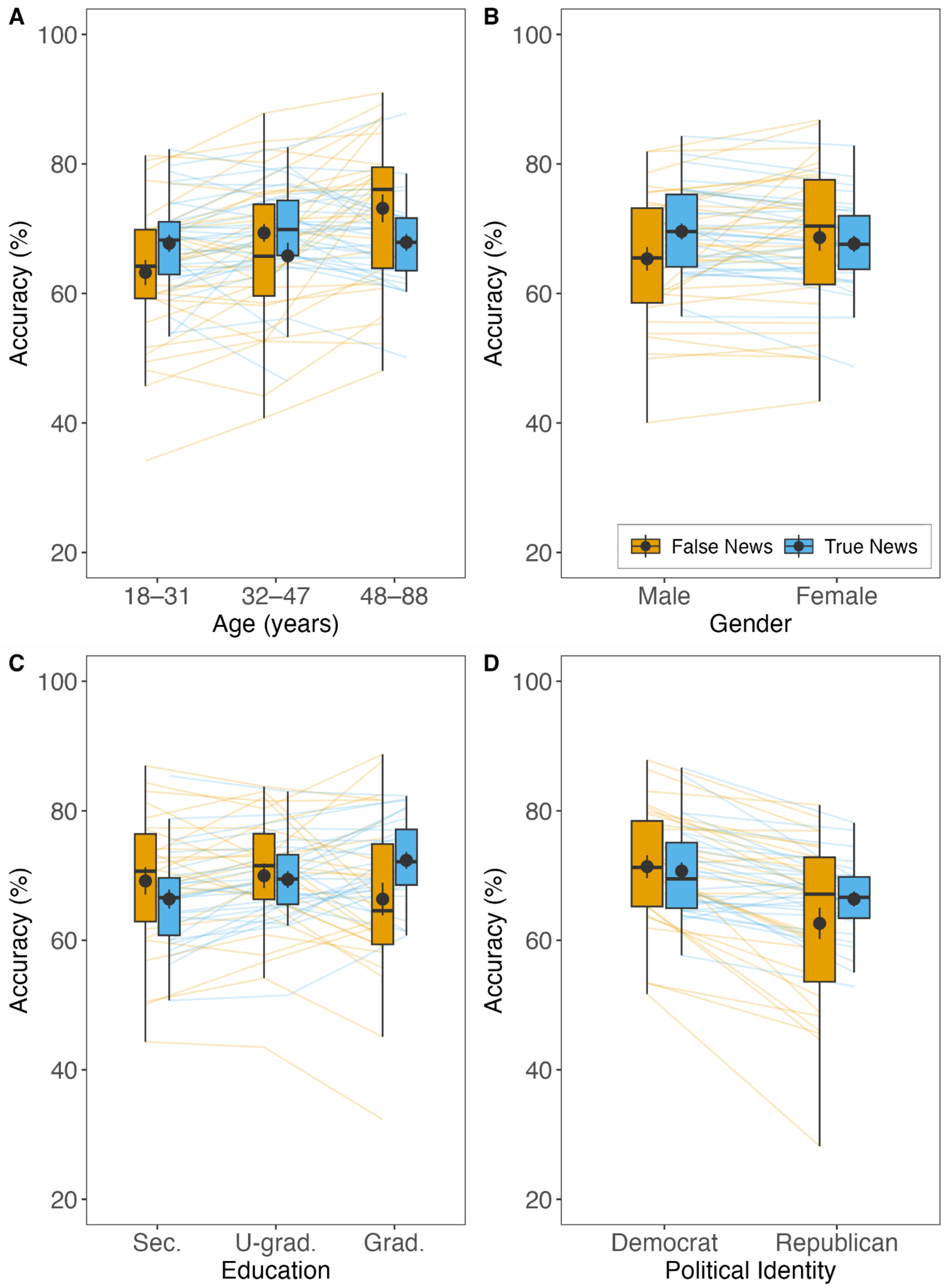
Gender

There was no credible effect of gender on discrimination ability (β = 0.01, CI = −0.05–0.07; Figure 1A). We did, however, find a small, negative, and credible association between gender and response bias (β = −0.07, CI = −0.12–−0.03; Figure 1A), with female participants being more likely than male participants to classify news headlines as false. Female participants had slightly higher accuracy for false news headlines than male participants, who in turn had slightly higher accuracy for true news (Figure 2B). Supplementary Figure 2C, D show these effects per study.

Education

Education did not have a credible effect on discrimination ability (β = 0.07, CI = 0.00–0.13; Figure 1A). We found a small, positive, and credible effect of education on response bias (β = 0.1, CI = 0.05–0.14; Figure 1A), with more educated individuals displaying a true-news bias that resulted in higher accuracy for true news and lower accuracy for false news (Figure 2C). Having more years of formal education was thus associated with an increased tendency to view news as true. Supplementary Figure 3A, B show these effects per study.

Figure 2. Demographic factors and true and false accuracy. *Accuracy for true and false news headlines across demographic factors. Age is split into tertiles for visualisation purposes. Coloured lines connect the accuracy scores within a study (e.g., false news accuracy across age tertiles). Boxplots show the median and the interquartile range (IQR); whiskers indicate an additional 1.5 IQR. Large dots represent the aggregate mean with standard errors. Sec = Secondary education. U-grad = undergraduate education. Grad = graduate education.*



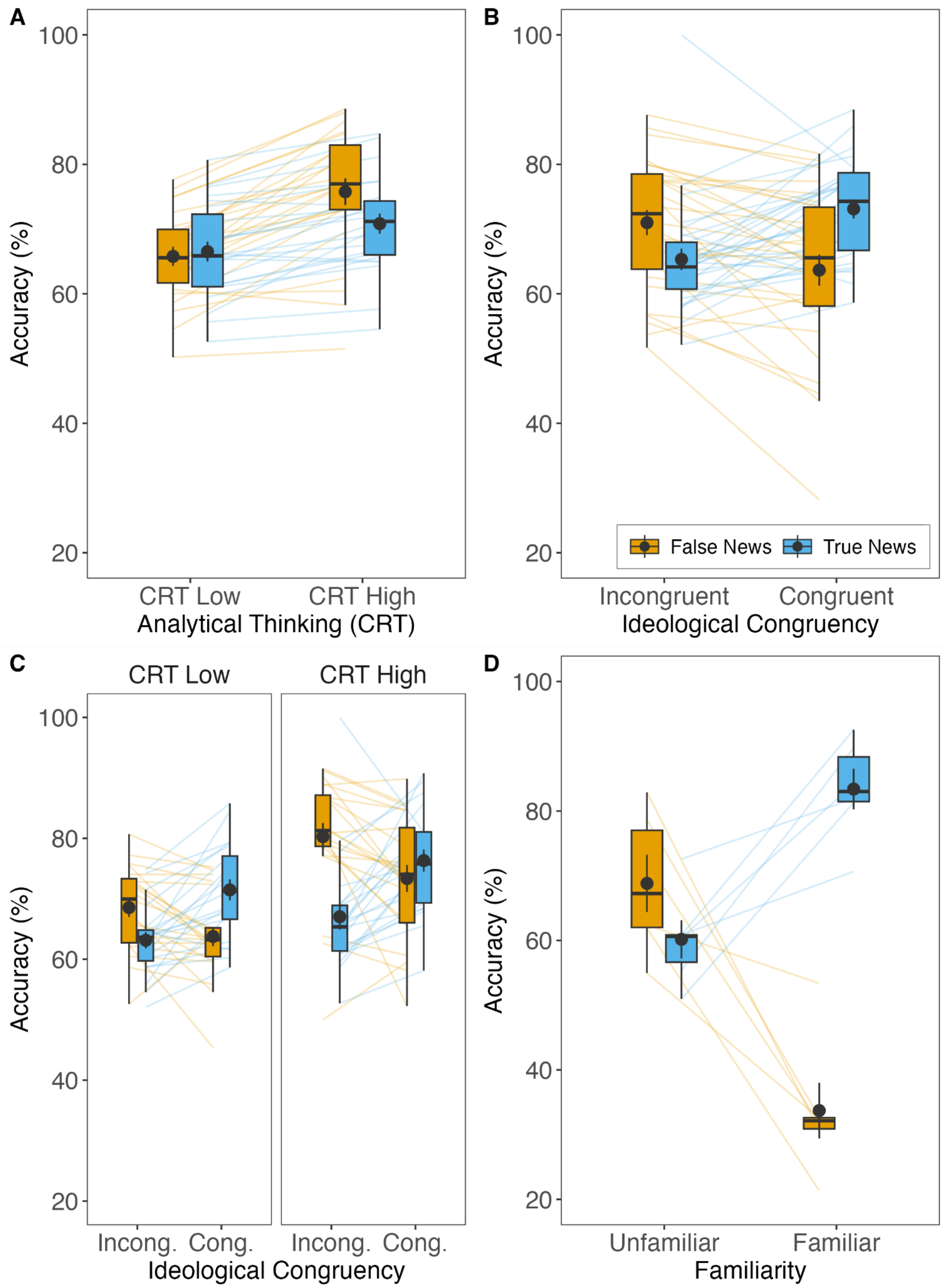
Political Identity

Political identity had a strong, credible, and negative effect on discrimination ability (β = −0.42, CI = −0.51–−0.32; Figure 1A), with Republican participants achieving lower overall accuracy compared to Democrat participants (Figure 2D). It also had a small, positive, and credible effect on response bias (β = 0.12, CI = 0.06–0.17; Figure 1A): Republicans were more likely to judge a news headline as true, resulting in them having slightly higher accuracy for true news (than false news), whereas Democrats had slightly higher accuracy for false news (than true news) (Figure 2D). Supplementary Figure 3C, D shows these effects per study. In sum, Republicans, compared to Democrats, showed reduced discrimination ability and a slightly more pronounced true-news bias (e.g., naïvety).

Analytical Thinking

Analytical thinking (represented by CRT score) had a strong, credible, and positive effect on discrimination ability (β = 0.66, CI = 0.52–0.81; Figure 1A), leading to higher overall accuracy for individuals with higher analytical thinking skills (Figure 3A). Analytical thinking was negatively associated with response bias (β = −0.19, CI = −0.26–−0.12; Figure 1A), meaning that individuals with higher analytical thinking skills were more inclined to judge a news headline as false, which resulted in greater accuracy for false news (Figure 3A). Supplementary Figure 4A, B show these effects per study. In sum, higher analytical thinking skills were linked to enhanced discrimination ability and a false-news bias (e.g., caution).

Figure 3. Psychological factors and true and false accuracy. *Accuracy for true and false news headlines across psychological factors. Analytical thinking measured via cognitive reflection test (CRT) scores, which are median split for visualisation purposes into CRT Low and CRT High. Coloured lines connect false and true news across categories (e.g., false news accuracy across CRT Low and High). Boxplots show the median and the interquartile range (IQR); whiskers indicate an additional 1.5 IQR. Large dots represent the aggregate mean with standard errors. Incong: Incongruent. Cong: Congruent.*



Ideological Congruency

There was no credible effect of ideological congruency on discrimination ability ($\beta = 0.1$, CI = −0.01–0.2; Figure 1A). We did find a strong, credible, and positive effect of ideological congruency on response bias ($\beta = 0.29$, CI = 0.22–0.37; Figure 1A), showing that participants were more inclined to judge a news headline as true (false) if it aligned (misaligned) with their ideological stance. This led to higher accuracy for congruent true headlines than for incongruent true headlines, and the reverse for false headlines: higher accuracy for incongruent false headlines than for congruent false headlines (Figure 3B). Supplementary Figure 4C, D show these effects per study. In sum, ideological congruency was associated with an increased tendency to believe news headlines (partisan bias) but had no effect on discrimination ability.

Supplementary Figure 5 shows the results of ideological congruency separated by political leaning: We found no credible differences.

Interaction Between Analytical Skills and Ideological Congruency

We found no credible effect of the interaction between analytical skills and ideological congruency on discrimination ability ($\beta = 0.04$, CI = −0.06–0.14; Figure 1A). As for motivated reflection, which is the interaction between analytical thinking and ideological congruency on the response bias, we found a credible and positive effect ($\beta = 0.16$, CI = 0.09–0.24; Figure 1A). The effect of ideological congruency on response bias was stronger for those with higher analytical skills (Figure 3C). Supplementary Figure 6A, B show these effects per study.

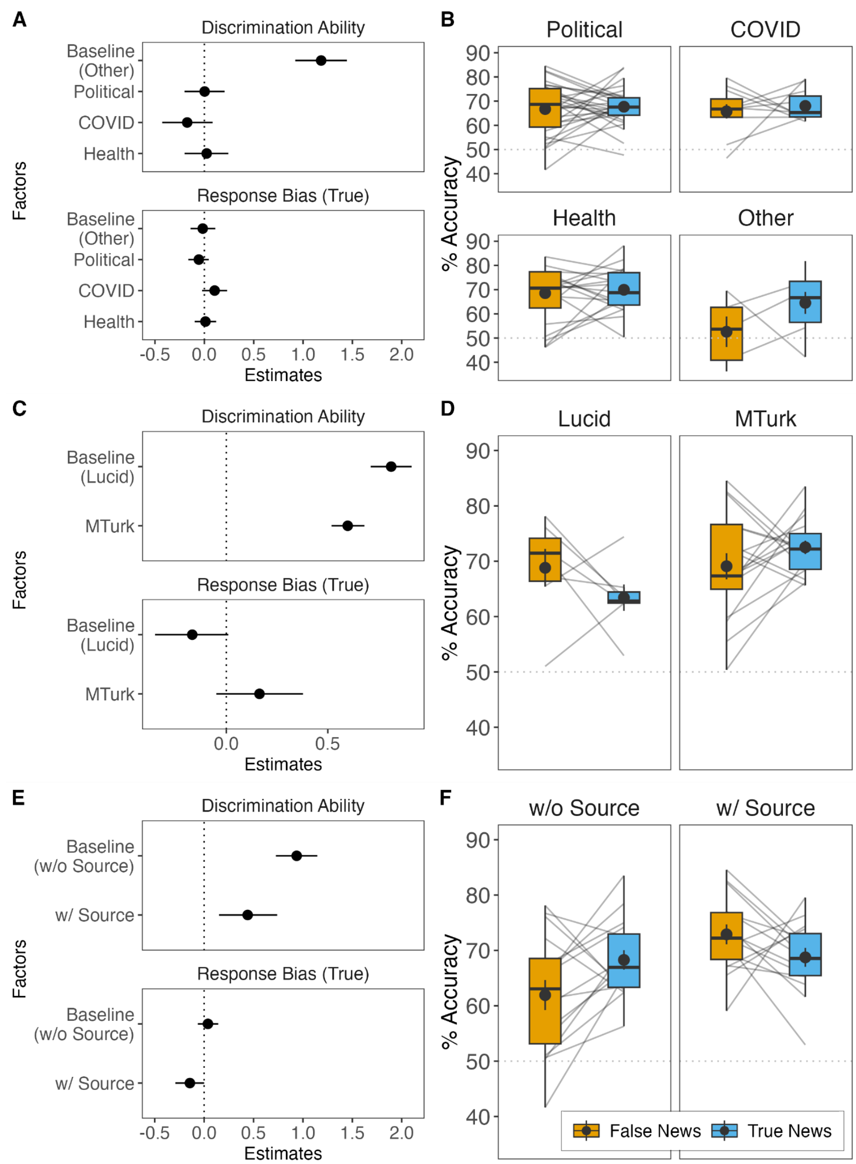
Familiarity

Only six studies (from five articles) included a measure of familiarity. Given the extent of missing data, we excluded familiarity from the main SDT model and instead ran a separate complete-case SDT model ($N_{\text{participants}} = 2{,}619$; $N_{\text{choices}} = 50{,}701$; $M_{age} = 42.13$ years, $SD = 16.23$, range = 18–88). Familiarity did not have an effect on discrimination ability ($\beta = 0.16$, CI = −0.03–0.3; Supplementary Figure 7A). However, it had a strong and positive effect on response bias ($\beta = 1.03$, CI = 0.67–1.34; Supplementary Figure 7A). Participants were much more likely to label news headlines as true if they were familiar, leading to higher accuracy for familiar true headlines and lower accuracy for familiar false headlines (Figure 3D). In sum, familiarity led to a strong true-news bias. Full regression results for familiarity can be found in Supplementary Figure 7A. Supplementary Figure 7B shows these effects per study.

Additional Analyses: Headline Topic, Platform, and Displayed Source

We also analysed the topic of the headlines (i.e., related to politics, COVID-19, or general health), the crowdsourcing platform (i.e., Lucid or MTurk), and whether the source was displayed (for detailed information on the analyses, see Methods and Figure 4). Noteworthy results include that the headline topic did not have an effect on discrimination ability, suggesting that the results for discrimination ability are consistent across topic types (Figure 4A, B). Data collection via MTurk, as opposed to Lucid, was associated with a strong, positive, and credible effect on discrimination ability (β = 0.6, CI = 0.52–0.68; Figure 4C). This led to MTurk participants having greater overall accuracy (Figure 4D). Finally, displaying the headline's source had a strong, positive, and credible impact on discrimination ability (β = 0.44, CI = 0.15–0.74; Figure 4E), resulting in higher overall accuracy when a source was displayed with the headline (Figure 4F).

Figure 4. Additional analyses for headline topic, study platform, and source display. *Panels A, C, E: Signal detection theory model estimates for additional analyses. The upper half of each panel shows estimates for discrimination ability, with more positive (negative) values indicating higher (lower) discrimination ability. The lower half shows the estimates for response bias, with more positive (negative) values indicating a higher likelihood to judge headlines as true (false). Dots represent the mean; error bars represent the 95% CI of the posterior distribution. The reference level for each model is presented in brackets after "Baseline": Baseline (Other): Estimate of discrimination ability and response bias when the headline topic is not related to politics, COVID-19, or general health. Panels B, D, F: Accuracy for true and false news headlines for the additional analyses. Lines connect accuracy scores within a study (e.g., false news accuracy across political headlines). Boxplots show the median and the interquartile range (IQR); whiskers indicate an additional 1.5 IQR. Large dots represent the aggregate mean with standard errors. w/o = without. w = with.*



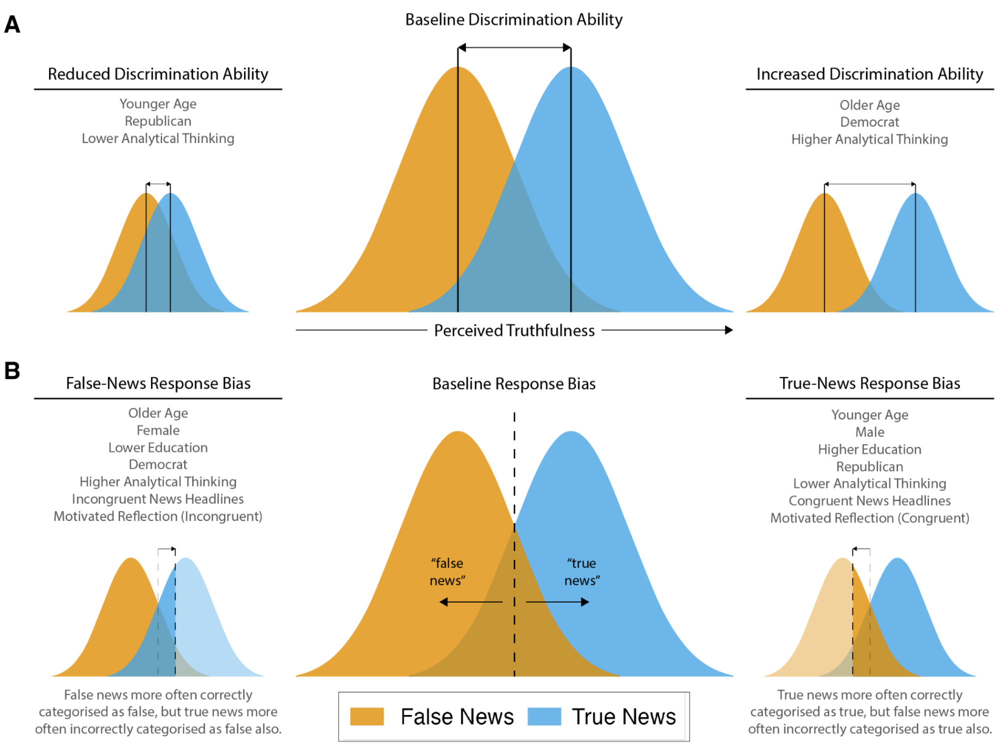
Discussion

Using 256,337 unique choices made by 11,561 participants across 31 studies, we investigated how four demographic factors—age, gender, education, and political identity—and four psychological factors—analytical thinking, ideological congruency, motivated reflection, and familiarity—impact online misinformation veracity judgments. We additionally analysed the topic of the headlines (i.e., related to politics, COVID-19, or general health), the platform (i.e., Lucid or MTurk), and whether the source was displayed. We used signal detection theory (SDT; Macmillan & Creelman, 2004) to analyse how these factors are associated with discrimination ability, the ability to distinguish between true and false news, and response bias, a tendency to classify news as true (true-news bias) or false (false-news bias).

Older age and higher analytical thinking skills were associated with better discrimination ability, whereas identifying as Republican (as opposed to Democrat) was associated with worse discrimination ability. Older age, identifying as female (as opposed to male), and higher analytical thinking skills were associated with a false-news response bias (caution), whereas higher education, identifying as Republican (as opposed to Democrat), ideological congruency, motivated reflection, and familiarity were associated with a true-news response bias (naïvety; Figure 5). For the additional analyses, headline topic was not associated with discrimination ability, whereas data collection via MTurk (as opposed to Lucid), and displaying the source of news headlines was associated with greater discrimination ability.

While older adults are often considered to be more digitally naïve (Pew Research Center, 2021, 2024), our findings show that this does not thwart their ability, possibly fostered in the offline world, to accurately discern between true and false news (see also Allcott & Gentzkow, 2017; Brashier & Schacter, 2020; Pennycook & Rand, 2019). Older adults were also more likely to classify news headlines as false, which can be interpreted as a cautious approach. By contrast, younger adults were less adept at discerning true and false news and were more naïve, suggesting that their higher digital literacy does not seem to help with their ability to judge online news. These findings raise alarms that younger adults are more vulnerable, not more resistant, to misinformation. Given these robust results (see also Supplementary Figure 2A), it is surprising that the effect of age on news veracity judgments has received relatively little attention and lacks a theoretical framework (cf. Pehlivanoglu et al., 2020).

Figure 5. Simplified visual summary of the main signal detection analysis. *Panel A: The middle shows a visual representation of baseline discrimination ability. The perceived truthfulness of a news headline is represented by an axis ranging from low truth to high truth, as represented via the two Gaussian distributions. The more the distributions overlap, the more similar the true and false news headlines are perceived (i.e., lower the discrimination ability); whereas the less they overlap, the more dissimilar the true and false headlines are perceived (ie., higher the discrimination ability). The left shows which factors were associated with reduced discrimination ability and right shows which factors were associated with increased discrimination ability. Panel B: The middle shows baseline response bias, which is determined by a decision criterion (i.e., vertical dashed line). The response for whether a news headline is true or false is dependent on where the headline falls relative to the criterion. If the criterion is placed higher up the perceived truthfulness dimension (left), more evidence is required to treat a news headline as true, hence a headline is treated as true less-often, resulting in a false-news response bias. The opposite holds for a true-news response bias (i.e., less evidence is required to render a news headline as true; right). The left shows which factors were associated with a false-news response bias and the right shows which factors were associated with a true-news response bias.*



Our results raise questions about which factors lead to older adults' higher discrimination ability. Is this the result of accumulated knowledge over time (e.g., crystallised intelligence; Cattell, 1963), enhanced vocabulary (e.g., enriched semantic memory; Cosgrove et al., 2023), or are there more tangible factors like news interest, news consumption patterns, and strategies and heuristics (e.g., source cues) at work that could be harnessed for future interventions? Insights into these questions will help clarify the paradox between older adults' veracity judgments and sharing judgments: what distracts adults—particularly older adults—from news veracity enough for them to end up sharing (the most) false news? Furthermore, sharing decisions likely tap into different cognitive and social processes, making it important to examine whether cognitive deficits (e.g., overreliance on fluency), social factors (e.g., community related motivations; Osmundsen et al., 2021), digital literacy, or the structure of social media itself (e.g., Epstein et al., 2023) are at the root of this paradox. As a cautionary note, it is important to consider that older adults that take part in online studies may not be representative of their age group, as they likely have higher digital literacy, which could influence the results.

Higher formal education levels were not associated with higher discrimination ability. Surprisingly, we found that more educated individuals had a slight tendency to judge news headlines as true (naïvety). This is striking, especially as education is associated with the development of critical thinking skills, including the ability to challenge assumptions, scrutinise sources, and weigh the pros and cons of information (although see Kozyreva et al., 2022b). These results, in combination with our findings on age, raise concerns about the adequacy of current educational frameworks, especially as other studies show that highly educated individuals are easily manipulated online by objective-sounding language and official-looking logos and domains (Breakstone et al., 2022; McGrew et al., 2018). Interventions such as lateral reading (McGrew et al., 2018; Panizza et al., 2022), media literacy (Guess et al., 2020a), and inoculation (prebunking; van der Linden & Roozenbeek, 2020), which are known to improve news veracity judgments and increase caution, may boost such competences once included into educational curricula.

Relatedly, higher analytical thinking skills were robustly associated with better discrimination ability and a tendency to categorise news as false (e.g., caution). This held across partisan lines, echoing previous findings, and cementing the important role of deliberation in enhancing news veracity judgments (Batailler et al., 2021; Pennycook & Rand, 2021; Sultan et al., 2022). Despite the null effect of critical thinking skills as learned via education, there is an opportunity here to harness analytical thinking skills to boost (Hertwig & Grüne-Yanoff, 2017) news veracity judgements. While thoughtful and deliberative processing of information has been shown to improve veracity judgments in the past (Bago et al., 2020), it is unclear what exactly the CRT is measuring (e.g., Attali & Bar-Hillel, 2020) and what mechanisms are activated during “deliberation”—are participants putting in more effort (e.g., time), overriding various biases (e.g., truth bias), using more discerning heuristics, or managing to retrieve cues with higher validity (e.g., source credibility) to arrive at correct judgments? If interventions are to more effectively improve veracity judgments, these mechanisms need to be better understood and nurtured.

We found that Democrats had a substantially higher discrimination ability compared to Republicans. This discrepancy represents a growing trend in the literature (Bago et al., 2020; Dobbs et al., 2023; Garrett & Bond, 2021), pointing to the considerable role of political identity in shaping people’s perceptions of truth. One potential explanation lies in the notion that different political groups inhabit distinct informational worlds, both offline and online—a situation that is likely exacerbated by online echo chambers, filter bubbles, and microtargeting (Barberá et al., 2015; Robertson et al., 2023; Simchon et al., 2024). For instance, studies show that Republican politicians are making fewer evidence-based statements and more belief-based statements (e.g., Lasser et al., 2023); that misinformation tends to favour more Republican (conservative) positions whereas true news tends to favour more Democratic (liberal) positions, thereby skewing perceptions of truth (Allcott & Gentzkow, 2017; Garrett & Bond, 2021); and that Republicans are also more exposed to—and share more—articles from unreliable websites (Grinberg et al., 2019; Guess et al., 2019; Guess et al., 2020b).

It is, nonetheless, important to emphasise that the main issue at hand is not whether Democrats or Republicans are better at identifying misinformation but that political identity can shape what people perceive to be true, which can in turn shape their behaviour. Because a functioning democracy is ultimately dependent on a shared reality (e.g., that a specific election outcome was legitimate and fair), further work is needed to understand and bridge divergent perceptions of truth. Research could delve deeper into the consequences of exposure to different base rates of partisan news and how that impacts veracity judgments and sharing judgments. Interventions could also aim to emphasise commonalities, shared civic values, and promote respectful dialogue across political boundaries to navigate these distinct realities (Voelkel et al., 2023).

The implications of how political identity can shape perceptions of truth also become much more apparent in light of ingroup vs. outgroup favouritism. Ideological congruency resulted in a strong true-news bias (see also Batailler et al., 2021; Pennycook & Rand, 2021), both for Democrats and Republicans. That is, news that was congruent (incongruent) with one's political identity was seen as true (false). One explanation for this lies with the identity-based model of political belief (Van Bavel et al., 2024; Van Bavel & Pereira, 2018), positing that people need to maintain a positive view of their ingroup (political party) in relation to their outgroup. This has led some to argue that partisan differences in vaccinations and public-health behaviours can be explained by such identity-driven motivations (Van Bavel et al., 2023). Note also that congruent news may also be more familiar news, resulting in the congruency effect partially, or fully, being a familiarity effect in disguise. This, in turn, could attenuate the role that identity-driven motivations play in the susceptibility to misinformation. How these two drivers relate deserves more research attention. Overall, as ideological congruency predominantly affects response bias, effective interventions should encourage individuals to objectively evaluate information, irrespective of its ideological alignment. We also encourage future research to investigate the cues that people treat as signals for ideological congruency in order to disentangle the mechanisms behind ideological congruency (e.g., see Van Bavel et al., 2024; Van Bavel & Pereira, 2018).

Finally, familiarity with a news headline strongly increased the likelihood that participants would believe it to be true, regardless of whether it was or not (Batailler et al., 2021; Sultan et al., 2022). Familiarity can easily be gamed online, where similar headlines can appear and reappear in endless news feeds, especially within echo chambers. Generally, it may be beneficial to trust familiar information: Doing so can reduce cognitive load and simplify decision-making processes. However, this adaptive process becomes maladaptive in an environment where familiarity can be manipulated—at an unprecedented speed and sophistication—and is not in itself indicative of veracity. Unfortunately, interventions to eliminate the effect of familiarity have largely been unsuccessful. Though interventions that ask participants to assess the familiarity or truthfulness of news headlines have shown some promise in reducing the effect (Udry & Barber, 2024). As the familiarity effect has been studied in other domains, such as with trivia statements (e.g., Henderson et al., 2021), this literature could be reviewed to identify interventions that have successfully reduced the familiarity effect. Note, however, that our findings on familiarity are based on only six studies due to our eligibility criteria, which excluded studies that experimentally manipulated prior exposure to news headlines. Given the robustness of the familiarity effect in prior studies, we are nonetheless confident in these results (for recent reviews, see Ecker et al., 2022; Udry & Barber, 2024).

In our additional analyses, we found that the results for discrimination ability hold across different types of news headline topics (i.e., related to politics, COVID-19, or general health). These results generally speak to the robustness of our findings across different headline topics. We also found that MTurk participants exhibited higher discrimination ability than Lucid users, highlighting that veracity judgments can be systematically biased depending on the recruitment platform. This could be due to various reasons, including task engagement of participants (e.g., quickly clicking through), experience with news related tasks, and how quality of participants is maintained on the platforms. Conducting studies across multiple platforms—where possible—may mitigate this effect.

People were better at judging the veracity of a headline if a source was also displayed alongside the news headline. Studies looking into source credibility (e.g., that manipulate the size of sources or the partisan alignment of sources) have found mixed results (Dias et al., 2020; Traberg & van der Linden, 2022). Our results, however, clearly suggest that the presence of a source is used as a cue for veracity. Naturally, relying on source credibility as a heuristic is adaptive, especially given established trustworthiness of sources. This, however, can easily be gamed in the age of artificial intelligence, reinforcing the need to maintain trust in institutions. Beyond this, at a methodological level, omitting source display can lead to issues that conflate source credibility with content veracity (e.g., isolating the direct effects of content characteristics on judgment veracity). It can also, for example, make it challenging to assess efficacy uniformly across intervention studies that vary in their display of source, or studies that mix source display altogether.

We find no evidence for an overall response bias across studies. However, there is substantial variation in response bias across studies, with studies exhibiting a true-news bias, a false-news bias, or no bias (Supplementary Figure 1B). This variability may be indicative of broader study-level features, such as news headline selection. To illustrate, while false news headlines are primarily sourced from fact-checking organisations (e.g., Snopes), true news headlines are largely derived from mainstream sources and may exhibit systematic differences. The reverse scenario may also hold, where fact-checked headlines differ systematically from true-news headlines. Additionally, participants' responses may be influenced by experimental and contextual demands (e.g., true–false base rates, participants' awareness of misinformation in the study). To probe the influence of experiment demands, studies could employ experience-sampling methods or social media simulators (e.g., Butler et al., 2023; Epstein & Lin, 2022).

Several limitations need to be considered. First, the included news headlines do not fully represent the spectrum of news encountered online. Second, almost all studies use an equal distribution of true and false news, which does not reflect base rates encountered online (Altay et al., 2023; Orchinik et al., 2024). Third, not all studies collected data on every demographic and psychological factor, leading to some data being imputed (see Methods). Finally, our analysis focused on a U.S.-only sample and simplified gender and political identity into binary categories, which does not capture their full complexity.

Given the multitude of demographic and psychological factors shaping misinformation veracity judgments, misinformation research should adopt a multifaceted and multipronged approach (Geers et al., 2024b; Kozyreva et al., 2022a). The insights provided by our meta-analysis establish the robustness (or lack thereof) of demographic and psychological factors, touch upon key debates in the literature, pave the way for future research (see Box 1), and highlight the need for tailored interventions that take these factors into account. Beyond this, we call for more representative and ecologically valid research methodologies to enhance the generalisability of findings in this rapidly evolving domain.

Box 1. *Key recommendations for future research.*

Include relevant variables

Important variables like age and political identity are often collected but not adequately analysed. Future research could systematically include and measure these variables to avoid omitting meaningful variance.

Go beyond overall accuracy

Using analytical frameworks like SDT that can assess discrimination ability and response bias simultaneously is crucial. This approach provides a more nuanced understanding of how people discern truth in news that can be used to develop more targeted interventions.

Incorporate reaction time data

Including reaction time data in future studies could offer deeper insights into the cognitive mechanisms behind news veracity judgments (e.g., via a drift–diffusion modelling approach; Alvarez-Zuzek et al., 2024; Gollwitzer, Tump et al., 2024).

Select diverse headlines
The number and types of headlines used in studies should be carefully considered. A representative sample of headlines, reflective of the broader online ecosystem, is essential for more accurate insights. More ecologically valid studies are critical for generalisability.

Understand veracity as part of a larger puzzle
Veracity is just one aspect of misinformation. A view that considers other elements of news consumption and sharing is necessary.

Dive deeper into headline characteristics
The headlines used in studies are complex stimuli. A deeper analysis of headline characteristics (e.g., text complexity, emotion) could help shed light on veracity judgments and sharing judgments.

Expand beyond headlines
While our focus has been on news headlines, there is a larger context of news consumption that should be explored, including other forms of news media consumption (e.g., full-length articles; offline consumption, instant messaging platforms).

Practise open science
Most of the necessary data for our meta-analysis were accessible. Researchers should maintain transparency, making data readily available for replication and further analysis.

Improve data coding and documentation
Improvements in how data are coded and documented would greatly aid future research. Providing a detailed codebook, for instance, would facilitate data comprehension and reusability.

Move beyond WEIRD samples (Henrich et al., 2010)

This study focused on U.S.-based sample populations, as much of the current research is conducted there. We encourage research that covers other geographical areas, in particular countries that are not based on a two-party political system.

Methods

This meta-analysis was preregistered and conducted in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA; Moher et al., 2009) guidelines (Figure 6). The preregistration, along with all the materials to reproduce our analyses, are available on OSF. Deviations from the preregistration can also be found in "Protocol Deviations" on OSF. Our study did not require ethical approval as we re-analysed pre-existing data. Data that were not publicly available were requested from the authors.

Eligibility Criteria

Table 1 presents the study eligibility criteria.

Search Strategy

We used two general databases, Web of Science and Scopus, and one subject-specific database, PsycINFO, to search for relevant literature. This search strategy was developed with—and tested against—a preselected list of six articles that matched the eligibility criteria (Bago et al., 2020; Brashier et al., 2021; Martel et al., 2020; Pennycook et al., 2020, 2021; Roozenbeek et al., 2022). The search string was developed to include terms from four main categories of interest: news AND misinformation AND veracity AND human. Exact search terms for Web of Science, Scopus, and PsycINFO are listed in Supplementary Table 1. The search was limited to English-language articles, conference papers, and early access papers from 2004 until November 2022. To identify unpublished and shelved data, we searched PsyArXiv Preprints and sent email requests to the Society for Judgment and Decision Making and the Fake News Sci mailing list. For a detailed breakdown of the search strategy, including exact search terms for each platform and dates, see "Search Syntax" on OSF.

Screening

We selected 21 articles (encompassing 31 studies) for data extraction (for an overview, see Figure 6). EndNote (https://endnote.com/) was used to track and manage search results (N = 4,666), including de-duplication. The lead author (MS) first conducted a title-only screening of articles (N = 3,314) in order to exclude articles that were clearly irrelevant (e.g., articles that used machine learning to detect misinformation, non-U.S.-based samples; $N_{remaining}$= 1,261). Next, two coders (MS and NE) screened 100 titles and abstracts against the eligibility criteria to clarify questions and assess inter-rater agreement. This served as a training round to ensure that both raters' subjective ratings agreed with the eligibility criteria, with disagreements overcome through discussion. We had preregistered to repeat this step if the inter-rater agreement was below 90%. In our case, an inter-rater agreement of 87% was reached in each of two consecutive attempts (n = 200), so we decided to proceed to the next step: We used Abstrackr (Wallace et al., 2012) to make title- and abstract-based decisions with reference to the eligibility criteria (N_{total} = 1,257). Article titles and abstracts (along with other fields, e.g., journal and authors), were displayed on Abstrackr. The articles were randomised for both MS and NE. MS and NE then screened the full texts of all remaining articles for eligibility (N = 267); when uncertain, they consulted two co-authors (RHJMK and ANT). An inter-rater agreement of 87% and an inter-rater reliability of 0.64 (Cohen's *Kappa*, p < .001; using the irr R package; Gamer et al., 2012) was reached. Four further articles were identified via mailing lists but did not meet the inclusion criteria. Table 2 presents the final list of articles, including the number of studies per article, each study's sample size, and number of headlines. See also "Search Results" on OSF with the full—and final—lists of studies, including notes on disagreements.

Figure 6. *PRISMA flow diagram summarising the study screening process.*

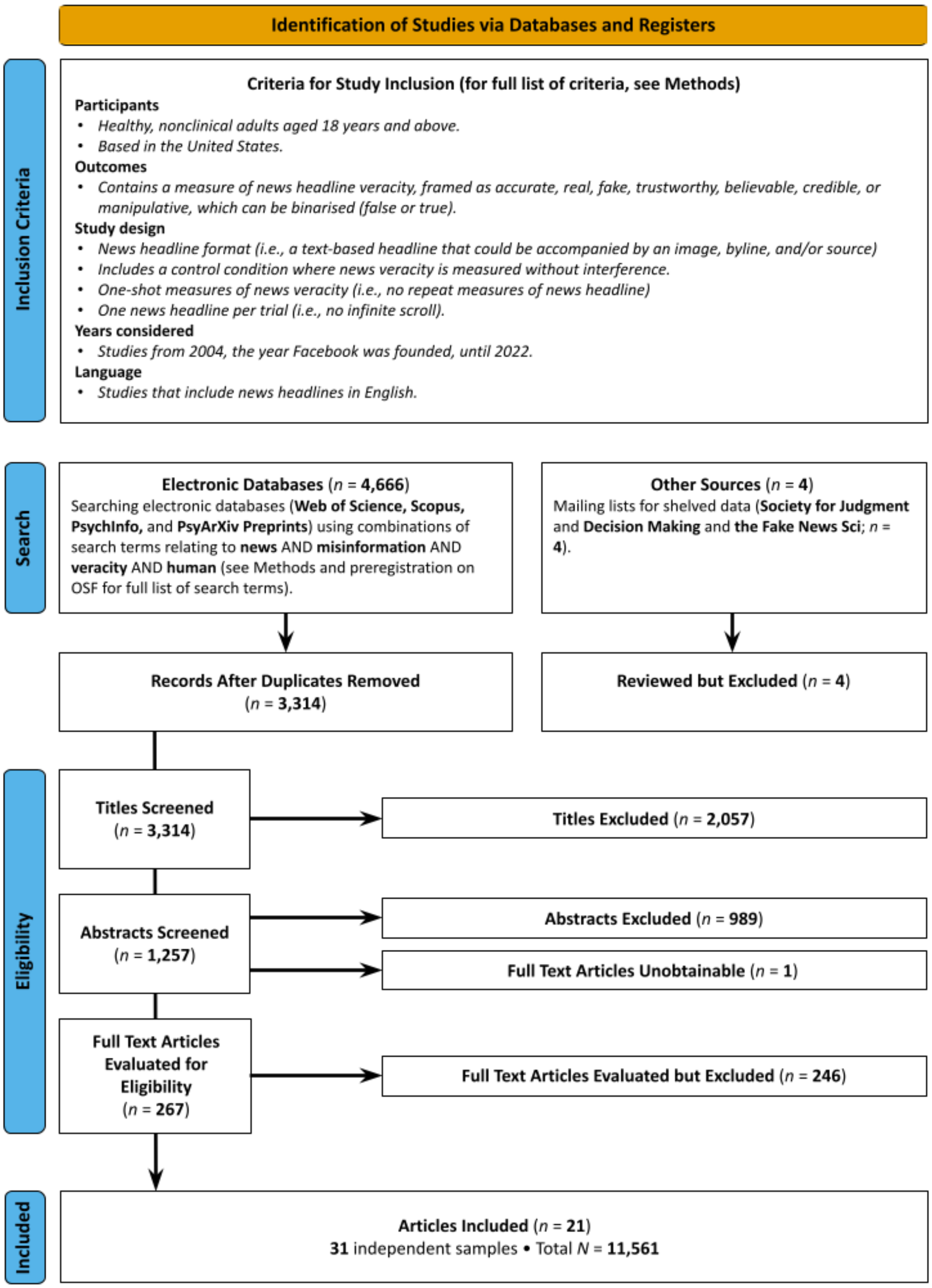


Table 1. *Eligibility criteria*

Criteria	Description
Participants	Only healthy (i.e., nonclinical) adults who were 18 years of age or older and based in the United States were included.
Study design	Participants judged the veracity of a series of news headlines. Headlines (e.g., the Misinformation Susceptibility Test; Maertens et al., 2023) could be accompanied by an image, a byline, and/or a source. Studies all included both true and false news headlines that were viewed individually (e.g., no infinite scroll). Studies that provided immediate feedback were excluded.
Outcome measure	Eligible veracity question framing included whether the news headline was accurate, real, fake, trustworthy, believable, credible, or manipulative. Acceptable responses were either binary (true/false) or derived from even-numbered Likert scales (odd-numbered Likert scales cannot be binarised, as the middle point cannot be assigned to either side of the scale). Note that previous research has shown little difference in veracity judgments across different question framings and response formats (Roozenbeek et al., 2022).
Control data	Data from control treatments in studies featuring both an intervention and a control were included if they satisfied all other inclusion criteria. Data were excluded if participants had to assess the news headlines on aspects other than its veracity (e.g., sharing intentions). Exceptions were made for evaluations of news headlines' familiarity and participants' confidence in their veracity judgments. Familiarity was considered because we were interested in its effect. Since confidence levels can be viewed as meta-judgments to veracity judgments, we determined that this aspect would not alter participants' responses (see also Roozenbeek et al., 2022).
Model input	Data that could not be converted to the desired model input were excluded (see Statistical Analysis below). For example, data were excluded if the political identity of participants was measured using an odd-numbered Likert scale, which cannot be binarised.
Time frame	Studies where participants were asked to re-rate headlines they had already been shown were excluded (i.e., one-time ratings of news headlines only).

Study year	Studies conducted before 2004, the year Facebook was founded, were excluded due to our interest in news headlines as seen on social media (i.e., image and headline).
Language	Only studies that used English-language headlines were considered.
Publication status	Published and unpublished (e.g., preprints; shelved data) original research was considered, excluding reviews and meta-analyses.
Data availability	Studies were excluded if their raw data were not accessible online and remained unavailable after multiple unanswered data requests. A list of studies excluded due to inaccessible data can be found in "Search Results" on OSF.

Table 2. *Overview of studies*

Reference	Study ID (Platform)	$N_{participants}$	$N_{headlines}$ (Total)*
Arechar et al. (2023). Understanding and combatting misinformation across 16 countries on six continents. *Nat. Hum. Behav.*	1 (Lucid)	352	20
Bago, B., Rand, D. G., & Pennycook, G. (2020). Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. *J. Exp. Psychol. Gen.*	2 (MTurk)	134	10
	3 (MTurk)	355	16 (24)
Brashier, N. M., Pennycook, G., Berinsky, A. J., & Rand, D. G. (2021). Timing matters when correcting fake news. *PNAS.*	4 (MTurk)	284	36
	5 (MTurk)	279	36
Bronstein, M. V., Pennycook, G., Bear, A., Rand, D. G., & Cannon, T. D. (2019). Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking. *J. Appl. Res. Mem. Cogn.*	6 (MTurk)	629	24
Bronstein, M. V., Pennycook, G., Buonomano, L., & Cannon, T. D. (2021). Belief in fake news, responsiveness to cognitive conflict, and analytic reasoning engagement. *Think. Reason.*	7 (Lab)	126	24
Calvillo, D. P., & Smelter, T. J. (2020). An initial accuracy focus reduces the effect of prior exposure on perceived accuracy of news headlines. *Cogn. Res. Princ. Implic.*	8 (MTurk)	75	16 (32)
	9 (MTurk)	159	16
Calvillo, D. P., Ross, B. J., Garcia, R. J. B., Smelter, T. J., & Rutchick, A. M. (2020). Political	10 (MTurk)	409	16

ideology predicts perceptions of the threat of COVID-19 (and susceptibility to fake news about it). *Soc. Psychol. Pers. Sci.*	11 (MTurk)	354	16
Calvillo, D. P., Garcia, R. J. B., Bertrand, K., & Mayers, T. A. (2021). Personality factors and self-reported political news consumption predict susceptibility to political fake news. *Pers. Individ. Dif.*	12 (MTurk)	266	24
Calvillo, D. P., Rutchick, A. M., & Garcia, R. J. B. (2021). Individual differences in belief in fake news about election fraud after the 2020 U.S. election. *Behav. Sci.*	13 (MTurk)**	291	15
Carnahan, D., Bergan, D. E., & Lee, S. (2021). Do corrective effects last? Results from a longitudinal experiment on beliefs toward immigration in the U.S. *Polit. Behav.*	14 (RN/SSI)**	247	6
Epstein, Z., Sirlin, N., Arechar, A., Pennycook, G., & Rand, D. (2023). The social media context interferes with truth discernment. *Sci. Adv.*	15 (Lucid)**	156	25
	16 (Lucid)	451	24 (60)
Garrett, R. K., & Bond, R. M. (2021). Conservatives' susceptibility to political misperceptions. *Sci. Adv.*	17 (YouGov)	852	20[†]
Longoni, C., Fradkin, A., Cian, L., & Pennycook, G. (2022). News from generative artificial intelligence is believed less. *2022 ACM Conf. Fairness, Accountability, and Transparency.*	18 (Lucid)	1047	42
Martel, C., Pennycook, G., & Rand, D. G. (2020). Reliance on emotion promotes belief in fake news. *Cogn. Res. Princ. Implic.*	19 (MTurk)	409	20 (32)
	20 (MTurk)	244	12
	21 (MTurk)	234	10
	22 (Lucid)	269	12
Newton, C., Feeney, J., & Pennycook, G. (2023). On the disposition to think analytically: Four distinct	23 (YouGov)	729	10 (20)

intuitive-analytic thinking styles. *Pers. Soc. Psychol. Bull.*			
Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychol. Sci.*	24 (Lucid)	414	30
Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*.	25 (MTurk)	497	36
Roozenbeek, J., Maertens, R., Herzog, S. M., Geers, M., Kurvers, R., Sultan, M., & van der Linden, S. (2022). Susceptibility to misinformation is consistent across question framings and response modes and better explained by myside bias and partisanship than analytical thinking. *Judgm. Decis. Mak.*	26 (Prolific)	650	20
Ross, R. M., Rand, D. G., & Pennycook, G. (2021). Beyond 'fake news': Analytic thinking and the detection of false and hyperpartisan news headlines. *Judgm. Decis. Mak.*	27 (MTurk)	505	20
	28 (Lucid)	490	20
Smelter, T. J., & Calvillo, D. P. (2020). Pictures and repeated exposure increase perceived accuracy of news headlines. *Appl. Cogn. Psychol.*	29 (Lab)	209	28
	30 (MTurk)	64	20
Sultan, M., Tump, A. N., Geers, M., Lorenz-Spreen, P., Herzog, S. M., & Kurvers, R. H. J. M. (2022). Time pressure reduces misinformation discrimination ability but does not alter response bias. *Sci. Rep*.	31 (Prolific)	381	32 (64)
Total	31	11,561	774[§]

Note. * = Some studies sampled headlines from a larger pool, with pool sizes in brackets. ** = Unequal base rates of true and false news headlines. † = Study included 12 waves of headlines, of which we used only the first. § = Total number of headlines across the 31 studies. However, some studies used the same headlines, resulting in a total of 562 unique headlines. RN/SSI = Research Now/SSI. Study numbers correspond to study numbers listed in the Supplementary Materials.

Data Extraction

Data from all studies were recorded and extracted by MS and NE. The data comprise five categories: study and participants, news headlines (e.g., veracity, political leaning), demographic factors, psychological factors, and veracity judgments. Supplementary Table 2 outlines each extracted variable. Supplementary Table 3 provides basic descriptive statistics for the sample, including age, gender, education, and political identity. Further descriptive statistics for age (Supplementary Figure 8), gender and political identity (Supplementary Figure 9), education (Supplementary Figure 10), and analytical thinking (Supplementary Figure 11) are also provided. Correlations between the factors are provided in Supplementary Figure 12.

Preprocessing and Missing Data

We had to standardise data across studies due to variation in data collection methods. Notably, this included converting data points from various scale response modes (e.g., a 6-point scale) into a binary format. Because education was measured inconsistently across studies, we converted it into three levels: secondary (including those who did not complete secondary education), undergraduate, and graduate, and turned it into a ridit score to better account for the ordinal nature of the data (for more, see Donaldson, 1988). Due to the broad scope of this meta-analysis, not all of the 31 studies included the demographic and psychological factors of interest. This resulted in missing data for education (10.71% overall; 4/31 studies), political identity (6.52% overall; 3/31 studies), analytical thinking (31.3% overall; 8/31 studies), and familiarity (80.23% overall; 6/31 studies). These data were mean-imputed. Due to the large amount of missing data for familiarity, however, we excluded it from the main analysis and ran a complete-case analysis instead (see Complete-Case Familiarity SDT Model below).

To determine ideological congruency, we required the political leaning of a given news headline, which was missing in 16/31 studies. We therefore used GPT4 to categorise the political leanings of news headlines (see also Rathje et al., 2023). We first assessed the quality of GPT4 ratings by asking GPT4 to code the political leanings of headlines we had already categorised; this resulted in high inter-rater agreement (88%) and inter-rater reliability of 0.78 (Cohen's *Kappa,* $p < .001$; using the irr R package; Gamer et al., 2012). Given the high reliability, we used GPT4 to categorise the political leanings of all headlines into the following categories: strongly Republican, moderately Republican, lean Republican, neutral, lean Democratic, moderately

Democratic, strongly Democratic (for more on the procedure, see Supplementary Figure 13, Supplementary Figure 14, and Supplementary Section 1).

Statistical Analysis: SDT Using Bayesian Generalised Linear Mixed-Effects Modelling

Main SDT Analysis

We conducted an SDT analysis using a mixed-effects model. We used a Bayesian generalised linear mixed-effects model (GLMM) with the R package brms (Bürkner, 2017) assuming a Bernoulli-distributed response with a probit link function. This implementation of a mixed-effects signal detection model allowed us to differentiate between discrimination ability and response bias (for a detailed overview, see DeCarlo, 1998; Rouder & Lu, 2005; Vuorre, 2017). In the GLMM, we used participants' response to the veracity question (i.e., veracity_response_binary; false, true) as the outcome variable. The predictor variables were item_veracity (whether the headline is false or true); part_age; part_gender (male, female); part_education (i.e., secondary, undergraduate, and graduate turned into ridit scores); part_political_identity_binary (Democrat, Republican); CRT (proportion of correct responses on the CRT; range 0–1); ideological_congruency (strongly incongruent, moderately incongruent, lean incongruent, neutral, lean congruent, moderately congruent, strongly congruent); and motivated_reflection (CRT × ideological_congruency). Finally, for random effects, we accounted for variability at multiple levels: We accounted for individual differences in discrimination ability and response biases among participants by including participant (part_id) as a random intercept and item_veracity as a random slope, and we included study-specific intercept and slope variations by adding study (study_id) as a random intercept and adding all of the fixed effects as random slopes into the model. We also accounted for headline-specific variations by adding news headline (item_id) as random intercepts (for full model specification, see Supplementary Section 2).

The intercept in this regression model reflects the response bias (i.e., the overall likelihood to classify a given headline as true) and the predictors' coefficients reflect their influence on this response bias. The only exception to this is when the coefficients include item_veracity, which indicates whether the headline is actually true or false. A positive estimate of headline veracity indicates increased ability to identify true news as true and false news as false (i.e., discrimination ability). The influence of the predictors on discrimination ability is, thus, inferred via the estimates of their interactions with headline veracity. To aid with model interpretation, all predictors were mean centred (i.e., value − mean). Part_age, part_education, CRT, and ideological_congruency were also divided by two standard deviations after mean centering (Gelman, 2008). For this model and those detailed below, the parameter estimates were generated by simulating four Markov chain Monte Carlo (MCMC) chains with 10,000 iterations each, discarding the first 5,000 as burn-in. We used the Gelman–Ruben statistic (Rhat) and visually inspected the Markov chains to ensure that all chains had converged. We report the mean of the posterior distribution and the 95% credible iIntervals (CI).

Complete-Case Familiarity SDT Model. As only six studies (from five articles) included a measurement of familiarity, we ran a separate complete-case SDT model for familiarity ($N_{\text{participants}}$ = 2,619; N_{choices} = 50,701; M_{age} = 42.13 years, *SD* = 16.23, range = 18–88), comprising 51.93% female (48.07% male) and 42.04% Republican (57.96% Democrat). While familiarity was complete-case, we had missing data for education (24%; 1/6 studies) and analytical thinking (36.52%; 2/6 studies). As with previous data, these values were mean imputed. The model was identical to the main model described above, apart from the following exceptions. This model included headline familiarity—familiarity_binary (unfamiliar, familiar)—as an additional predictor variable. This was mean centred. To account for study-specific intercept and slope variations, we used study (study_id) as a random intercept and only familiarity_binary as a random slope in the model. This was done to avoid convergence issues given the limited number of studies in the model.

Additional Analyses

We used the same method as in the main SDT analysis to study the impact of headline topic (i.e., related to politics, COVID-19, and general health), the platform (i.e., Lucid or MTurk), and whether the source was displayed on discrimination ability and response bias. See “Main SDT Analysis” above for guidelines on model interpretation.

Headline Topic. We used participants’ response to the veracity question (i.e., veracity_response_binary; false, true) as the outcome variable. The predictor variables were item_veracity (whether the headline is false or true), topic_political (whether the news headline was political; 0, 1), topic_covid (whether the headline was COVID-19 related on not; 0, 1), and topic_health (whether the headline was related to general health or not; 0, 1). Finally, for random effects, at the participant level, we modelled random intercepts for participants (part_id) and random slopes for item_veracity; we also modelled study (study_id), and news headlines (item_id) as random intercepts.

We used GPT 4 to generate the topic categories of the headlines. For each headline, the following prompts were used: (1) “Is the headline related to political issues or discussions? Answer with 0 for no and 1 for yes”; (2) “Is the headline related to COVID-19? Answer with 0 for no and 1 for yes”; and (3) “Is the headline related to health or medical information? Answer with 0 for no and 1 for yes.”

Platform. All aspects of the model remained the same apart from the predictor variables, which were item_veracity (whether the headline is false or true) and study_platform (Lucid, MTurk).

Source. All aspects of the model remained the same apart from the predictor variables, which were item_veracity (whether the headline is false or true) and item_source (not present, present).

References

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*(2), 211–236. https://doi.org/10.1257/jep.31.2.211

Altay, S., Berriche, M., & Acerbi, A. (2023). Misinformation on misinformation: Conceptual and methodological challenges. *Social Media + Society*, *9*(1). https://doi.org/10.1177/20563051221150412

Alvarez-Zuzek, L. G., Grujic, J., & Gallotti, R. (2024). Drift Diffusion Model to understand (mis)information sharing dynamic in complex networks. *arXiv preprint*. https://doi.org/10.48550/arXiv.2401.17846

Attali, Y., & Bar-Hillel, M. (2020). The false allure of fast lures. *Judgment and Decision Making*, *15*(1), 93-111. https://doi.org/10.1017/s1930297500006938

Bago, B., Rand, D. G., & Pennycook, G. (2020). Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. *Journal of Experimental Psychology. General*, *149*(8), 1608–1613. https://doi.org/10.1037/xge0000729

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting from left to right: is online political communication more than an echo chamber? *Psychological Science*, *26*(10), 1531–1542. https://doi.org/10.1177/0956797615594620

Batailler, C., Brannon, S. M., Teas, P. E., & Gawronski, B. (2021). A signal detection approach to understanding the identification of fake news. *Perspectives on Psychological Science*, *17*(1), 78–98. https://doi.org/10.1177/1745691620986135

Brashier, N. M., Pennycook, G., Berinsky, A. J., & Rand, D. G. (2021). Timing matters when correcting fake news. *Proceedings of the National Academy of Sciences*, *118*(5). https://doi.org/10.1073/pnas.2020043118

Brashier, N. M., & Schacter, D. L. (2020). Aging in an era of fake news. *Current Directions in Psychological Science*, *29*(3), 316–323. https://doi.org/10.1177/0963721420915872

Breakstone, J., Smith, M., Ziv, N., & Wineburg, S. (2022). Civic preparation for the digital age: How college students evaluate online sources about social and political issues. *The Journal of Higher Education*, *93*(7), 963–988. https://doi.org/10.1080/00221546.2022.2082783

Brown, A. S., & Nix, L. A. (1996). Turning lies into truths: Referential validation of falsehoods. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1088–

1100. https://doi.org/10.1037/0278-7393.22.5.1088

Bürkner, P.-C. (2017). brms: An r package for bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), Article 1. https://doi.org/10.18637/jss.v080.i01

Butler, L. H., Lamont, P., Wan, D. L. Y., Prike, T., Nasim, M., Walker, B., Fay, N., & Ecker, U. K. H. (2023). The (Mis)Information Game: A social media simulator. *Behavior Research Methods*, *56*(3), 2376–2397. https://doi.org/10.3758/s13428-023-02153-x

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*(1), 1–22. https://doi.org/10.1037/h0046743

Cosgrove, A. L., Beaty, R. E., Diaz, M. T., & Kenett, Y. N. (2023). Age differences in semantic network structure: Acquiring knowledge shapes semantic memory. *Psychology and Aging*, *38*(2), 87–102. https://doi.org/10.1037/pag0000721

De Keersmaecker, J., Dunning, D., Pennycook, G., Rand, D. G., Sanchez, C., Unkelbach, C., & Roets, A. (2020). Investigating the robustness of the illusory truth effect across individual differences in cognitive ability, need for cognitive closure, and cognitive style. *Personality & Social Psychology Bulletin*, *46*(2), 204–215. https://doi.org/10.1177/0146167219853844

DeCarlo, L. T. (1998). Signal detection theory and generalized linear models. *Psychological Methods*, *3*(2), 186–205. https://doi.org/10.1037/1082-989X.3.2.186

Dias, N., Pennycook, G., & Rand, D. G. (2020). Emphasizing publishers does not effectively reduce susceptibility to misinformation on social media. *Harvard Kennedy School Misinformation Review*. https://doi.org/10.37016/mr-2020-001

Dobbs, M., DeGutis, J., Morales, J., Joseph, K., & Swire-Thompson, B. (2023). Democrats are better than Republicans at discerning true and false news but do not have better metacognitive awareness. *Communications Psychology*, *1*(1). https://doi.org/10.1038/s44271-023-00040-x

Donaldson, G. (1998). Ridit scores for analysis and interpretation of ordinal pain data. *European Journal of Pain*, *2*(3), 221–227. https://doi.org/10.1016/s1090-3801(98)90018-0

Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E. K., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, *1*(1), 13–29. https://doi.org/10.1038/s44159-021-00006-y

Epstein, Z., & Lin, H. (2022). Yourfeed: Towards open science and interoperable systems for social media. *arXiv preprint*. https://doi.org/10.48550/arXiv.2207.07478

Epstein, Z., Sirlin, N., Arechar, A., Pennycook, G., & Rand, D. (2023). The social media context interferes with truth discernment. *Science Advances*, *9*(9). https://doi.org/10.1126/sciadv.abo6169

Evans, J. St. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7*(10), 454–459. https://doi.org/10.1016/j.tics.2003.08.012

Fadhila, S., Nisa, Y. F., Nihayah, Z., Hayat, B., Syani, P. A., & Adelina, R. (2021). Perceived accuracy of fake news on social media. *2021 9th International Conference on Cyber and IT Service Management (CITSM)*, 1–7. https://doi.org/10.1109/CITSM52892.2021.9588841

Fazio, L. K. (2020). Repetition increases perceived truth even for known falsehoods. *Collabra: Psychology*, *6*(1). https://doi.org/10.1525/collabra.347

Fortunati, L., Deuze, M., & de Luca, F. (2013). The new about news: How print, online, free, and mobile coconstruct new audiences in Italy, France, Spain, the UK, and Germany. *Journal of Computer-Mediated Communication*, *19*(2), 121–140. https://doi.org/10.1111/jcc4.12017

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, *19*(4), 25–42. https://doi.org/10.1257/089533005775196732

Garrett, R. K., & Bond, R. M. (2021). Conservatives' susceptibility to political misperceptions. *Science Advances*, *7*(23). https://doi.org/10.1126/sciadv.abf1234

Geers, M., Fischer, H., Lewandowsky, S., & Herzog, S. (2024a). The political (a)symmetry of metacognitive insight into detecting misinformation. *PsyArXiv*. https://doi.org/10.31234/osf.io/5mc8s

Geers, M., Swire-Thompson, B., Lorenz-Spreen, P., Herzog, S. M., Kozyreva, A., & Hertwig, R. (2024b). The Online Misinformation Engagement Framework. *Current Opinion in Psychology*, *55*. https://doi.org/10.1016/j.copsyc.2023.101739

Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine*, *27*(15), 2865–2873. https://doi.org/10.1002/sim.3107

Gollwitzer, A., Tump, A. N., Martel, C., Sultan, M., Kurvers, R., & Hertwig, R. (2024). A new framework for understanding and intervening on false news sharing. *PsyArXiv*.

https://doi.org/10.31234/osf.io/pxn29

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, *363*(6425), 374–378. https://doi.org/10.1126/science.aau2706

Guay, B., Berinsky, A. J., Pennycook, G., & Rand, D. (2023). How to think about whether misinformation interventions work. *Nature Human Behaviour*, *7*(8), 1231–1233. https://doi.org/10.1038/s41562-023-01667-w

Guess, A. M., Lerner, M., Lyons, B., Montgomery, J. M., Nyhan, B., Reifler, J., & Sircar, N. (2020a). 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*(27), 15536–15545. https://doi.org/10.1073/pnas.1920498117

Guess, A. M., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, *5*(1). https://doi.org/10.1126/sciadv.aau4586

Guess, A. M., Nyhan, B., & Reifler, J. (2020b). Exposure to untrustworthy websites in the 2016 U.S. election. *Nature Human Behaviour*, *4*(5), 472–480. https://doi.org/10.1038/s41562-020-0833-x

Halpern, D., Valenzuela, S., Katz, J., & Miranda, J. P. (2019). From belief in conspiracy theories to trust in others: Which factors influence exposure, believing and sharing fake news. In G. Meiselwitz (Ed.), *Social Computing and Social Media. Design, Human Behavior and Analytics* (pp. 217–232). Springer International Publishing. https://doi.org/10.1007/978-3-030-21902-4_16

Henderson, E. L., Westwood, S. J., & Simons, D. J. (2021). A reproducible systematic map of research on the illusory truth effect. *Psychonomic Bulletin and Review*, *29*(3), 1065–1088. https://doi.org/10.3758/s13423-021-01995-w

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hertwig, R., & Grüne-Yanoff, T. (2017). Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science*, *12*(6), 973–986. https://doi.org/10.1177/1745691617702496

Higham, P. A., Modirrousta-Galian, A., & Seabrooke, T. (2023). Mean rating difference scores are poor measures of discernment: The role of response criteria. *Current Opinion in Psychology*, *56*. https://doi.org/10.1016/j.copsyc.2023.101785

Holcombe, A. O., Kovacs, M., Aust, F., & Aczel, B. (2020). Documenting contributions to scholarly articles using CRediT and tenzing. *PLoS One*, *15*(12). https://doi.org/10.1371/journal.pone.0244611

Hornsey, M. J., & Fielding, K. S. (2020). Understanding (and reducing) inaction on climate change. *Social Issues and Policy Review*, *14*(1), 3–35. https://doi.org/10.1111/sipr.12058

Kahan, D. M., Peters, E., Dawson, E. C., & Slovic, P. (2017). Motivated numeracy and enlightened self-government. *Behavioural Public Policy*, *1*(1), 54–86. https://doi.org/10.1017/bpp.2016.2

Kalogeropoulos, A., Fletcher, R., & Nielsen, R. K. (2020). *Initial surge in news use around coronavirus in the UK has been followed by significant increase in news avoidance* (UK COVID-19 News and Information Project). Reuters Institute for the Study of Journalism.

Kozyreva, A., Lorenz-Spreen, P., Herzog, S. M., Ecker, U. K. H., Lewandowsky, S., Hertwig, R., … Wineburg, S. (2022a). Toolbox of interventions against online misinformation. *PsyArXiv*. https://doi.org/10.31234/osf.io/x8ejt

Kozyreva, A., Wineburg, S., Lewandowsky, S., & Hertwig, R. (2022b). Critical ignoring as a core competence for digital citizens. *Current Directions in Psychological Science*, *32*(1), 81–88. https://doi.org/10.1177/09637214221121570

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–498. https://doi.org/10.1037/0033-2909.108.3.480

Lasser, J., Aroyehun, S. T., Carrella, F., Simchon, A., Garcia, D., & Lewandowsky, S. (2023). From alternative conceptions of honesty to alternative facts in communications by US politicians. *Nature Human Behaviour*, *7*(12), 2140–2151. https://doi.org/10.1038/s41562-023-01691-w

Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: continued influence and successful debiasing. *Psychological Science in the Public Interest*, *13*(3), 106–131. https://doi.org/10.1177/1529100612451018

Linden, S. van der, Leiserowitz, A., & Maibach, E. (2017). Scientific agreement can neutralize politicization of facts. *Nature Human Behaviour*, *2*(1), 2–3. https://doi.org/10.1038/s41562-017-0259-2

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology Press.

Maertens, R., Götz, F. M., Golino, H. F., Roozenbeek, J., Schneider, C. R., Kyrychenko, Y., Kerr, J. R., Stieger, S., McClanahan, W. P., Drabot, K., He, J., & van der Linden, S. (2023). The Misinformation Susceptibility Test (MIST): A psychometrically validated measure of news veracity discernment. *Behavior Research Methods*, *56*(3), 1863–1899. https://doi.org/10.3758/s13428-023-02124-2

Martel, C., Pennycook, G., & Rand, D. G. (2020). Reliance on emotion promotes belief in fake news. *Cognitive Research: Principles and Implications*, *5*(1). https://doi.org/10.1186/s41235-020-00252-3

McGrew, S., Breakstone, J., Ortega, T., Smith, M., & Wineburg, S. (2018). Can students evaluate online sources? Learning from assessments of civic online reasoning. *Theory & Research in Social Education*, *46*(2), 165–193. https://doi.org/10.1080/00933104.2017.1416320

Modirrousta-Galian, A., & Higham, P. A. (2023). Gamified inoculation interventions do not improve discrimination between true and fake news: Reanalyzing existing research with receiver operating characteristic analysis. *Journal of Experimental Psychology: General*, *152*(9), 2411–2437. https://doi.org/10.1037/xge0001395

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, *151*(4), 264–269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Newman, N., Fletcher, R., Eddy, K., Robertson, C. T., & Nielsen, R. K. (2023). Digital news report 2023. *Reuters Institute*. https://policycommons.net/artifacts/4164711/digital_news_report_2023/4973510/

Orchinik, R., Martel, C., Rand, D. G., & Bhui, R. (2023). Uncommon errors: adaptive intuitions in high-quality media environments increase susceptibility to misinformation. *PsyArXiv*. https://doi.org/10.31234/osf.io/q7r58

Osmundsen, M., Bor, A., Vahlstrup, P. B., Bechmann, A., & Petersen, M. B. (2021). Partisan Polarization Is the Primary Psychological Motivation behind Political Fake News Sharing on Twitter. *The American Political Science Review*, *115*(3), 999–1015. https://doi.org/10.1017/s0003055421000290

Panizza, F., Ronzani, P., Martini, C., Mattavelli, S., Morisseau, T., & Motterlini, M. (2022). Lateral reading and monetary incentives to spot disinformation about science. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-09168-y

Pehlivanoglu, D., Lighthall, N. R., Lin, T., Chi, K. J., Polk, R., Perez, E., Cahill, B. S., & Ebner, N. C. (2022). Aging in an "infodemic": The role of analytical reasoning, affect, and news consumption frequency on news veracity detection. *Journal of Experimental Psychology: Applied*, *28*(3), 468–485. https://doi.org/10.1037/xap0000426

Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*, *592*(7855), 590–595. https://doi.org/10.1038/s41586-021-03344-2

Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science*, *31*(7), 770–780. https://doi.org/10.1177/0956797620939054

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, *188*, 39–50. https://doi.org/10.1016/j.cognition.2018.06.011

Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, *25*(5), 388–402. https://doi.org/10.1016/j.tics.2021.02.007

Persson, E., Andersson, D., Koppel, L., Västfjäll, D., & Tinghög, G. (2021). A preregistered replication of motivated numeracy. *Cognition*, *214*. https://doi.org/10.1016/j.cognition.2021.104768

Pew Research Center. (2021, February 4). *7% of Americans don't use the internet. Who are they?* Pew Research Center. https://www.pewresearch.org/short-reads/2021/04/02/7-of-americans-dont-use-the-internet-who-are-they/

Pew Research Center. (2024, January 31). *Social media fact sheet*. Pew Research Center. https://www.pewresearch.org/internet/fact-sheet/social-media/

Pinna, M., Picard, L., & Goessmann, C. (2022). Cable news and COVID-19 vaccine uptake. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-20350-0

Prochaska, S., Duskin, K., Kharazian, Z., Minow, C., Blucker, S., Venuto, S., West, J. D., & Starbird, K. (2023). Mobilizing manufactured reality: How participatory disinformation shaped deep stories to catalyze action during the 2020 U.S. presidential election. *Proceedings of the ACM on Human-Computer Interaction*, *7*(CSCW1), 1–39. https://doi.org/10.1145/3579616

Rathje, S., Mirea, D.-M., Sucholutsky, I., Marjieh, R., Robertson, C., & Van Bavel, J. J. (2023). GPT is an effective tool for multilingual psychological text analysis. *PsyArXiv*. https://doi.org/10.31234/osf.io/sekf5

Robertson, R. E., Green, J., Ruck, D. J., Ognyanova, K., Wilson, C., & Lazer, D. (2023). Users choose to engage with more partisan news than they are exposed to on Google Search. *Nature*, *618*(7964), 342–348. https://doi.org/10.1038/s41586-023-06078-5

Roozenbeek, J., Maertens, R., Herzog, S. M., Geers, M., Kurvers, R., Sultan, M., & Linden, S. van der. (2022). 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*(3), 547–573. https://doi.org/10.1017/S1930297500003570

Rouder, J. N., & Lu, J. (2005). An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic Bulletin & Review*, *12*(4), 573–604. https://doi.org/10.3758/BF03196750

Simchon, A., Edwards, M., & Lewandowsky, S. (2024). The persuasive effects of political microtargeting in the age of generative artificial intelligence. *PNAS Nexus*, *3*(2). https://doi.org/10.1093/pnasnexus/pgae035

Stagnaro, M. N., Tappin, B. M., & Rand, D. G. (2023). No association between numerical ability and politically motivated reasoning in a large US probability sample. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(32). https://doi.org/10.1073/pnas.2301491120

Sultan, M., Tump, A. N., Geers, M., Lorenz-Spreen, P., Herzog, S. M., & Kurvers, R. H. J. M. (2022). Time pressure reduces misinformation discrimination ability but does not alter response bias. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-26209-8

Tajfel, H., & Turner, J. (2004). *The Social Identity Theory of Intergroup Behavior*. In Psychology Press eBooks (pp. 276–293). https://doi.org/10.4324/9780203505984-16

Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making*, *11*(1), 99–113. https://doi.org/10.1017/s1930297500007622

Toff, B., & Palmer, R. A. (2019). Explaining the gender gap in news avoidance: "News-Is-for-Men" perceptions and the burdens of caretaking. *Journalism Studies*, *20*(11), 1563–1579. https://doi.org/10.1080/1461670X.2018.1528882

Traberg, C. S., & van der Linden, S. (2022). Birds of a feather are persuaded together: Perceived source credibility mediates the effect of political bias on misinformation susceptibility. *Personality and Individual Differences*, *185*. https://doi.org/10.1016/j.paid.2021.111269

Udry, J., & Barber, S. J. (2024). The illusory truth effect: A review of how repetition increases belief in misinformation. *Current Opinion in Psychology*, *56*. https://doi.org/10.1016/j.copsyc.2023.101736

Unkelbach, C., Koch, A., Silva, R. R., & Garcia-Marques, T. (2019). Truth by repetition: Explanations and implications. *Current Directions in Psychological Science*, *28*(3), 247–253. https://doi.org/10.1177/0963721419827854

Van Bavel, J. J., & Pereira, A. (2018). The partisan brain: An identity-based model of political belief. *Trends in Cognitive Sciences*, *22*(3), 213–224. https://doi.org/10.1016/j.tics.2018.01.004

Van Bavel, J. J., Pretus, C., Rathje, S., Pärnamets, P., Vlasceanu, M., & Knowles, E. D. (2023). The costs of polarizing a pandemic: Antecedents, consequences, and lessons. *Perspectives on Psychological Science*. https://doi.org/10.1177/17456916231190395

Van Bavel, J. J., Rathje, S., Vlasceanu, M., & Pretus, C. (2024). Updating the identity-based model of belief: From false belief to the spread of misinformation. *Current Opinion in Psychology*, *56*. https://doi.org/10.1016/j.copsyc.2023.101787

van der Linden, S., & Roozenbeek, J. (2020). Psychological inoculation against fake news. In R. Greifeneder, M. E. Jaffé, E. J. Newman, & N. Schwarz (Eds.), *The Psychology of Fake News* (1st ed., pp. 147–169). Routledge. https://doi.org/10.4324/9780429295379-11

Voelkel, J. G., Stagnaro, M., Chu, J., Pink, S. L., Mernyk, J. S., Redekopp, C., Ghezae, I., Cashman, M., Adjodah, D., Allen, L., Allis, V., Baleria, G., Ballantyne, N., Van Bavel, J. J., Blunden, H., Braley, A., Bryan, C., Celniker, J., Cikara, M., … Willer, R. (2023). Megastudy identifying effective interventions to strengthen Americans' democratic attitudes. *Open Science Framework Preprint*. https://doi.org/10.31219/osf.io/y79u5

Vuorre, M. (2017). *Bayesian Estimation of Signal Detection Models*. https://mvuorre.github.io/posts/2017-10-09-bayesian-estimation-of-signal-detection-theory-models/

Wang, W.-C., Brashier, N. M., Wing, E. A., Marsh, E. J., & Cabeza, R. (2016). On known unknowns: Fluency and the neural mechanisms of illusory truth. *Journal of Cognitive Neuroscience*, *28*(5), 739–746. https://doi.org/10.1162/jocn_a_00923

Wolverton, C., & Stevens, D. (2019). The impact of personality in recognizing disinformation. *Online Information Review*, *44*(1), 181–191. https://doi.org/10.1108/OIR-04-2019-0115

Acknowledgements

R.H.J.M.K. and A.N.T. acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2002/1 'Science of Intelligence' - project number 390523135. P.L-S. and R.H. acknowledge funding from the Volkswagen Foundation (grant 'Reclaiming individual autonomy and democratic discourse online: How to rebalance human and algorithmic decision making'). We thank Deborah Ain very much for editing the manuscript.

Author Contributions

Conceptualisation: M.S., A.N.T., and R.H.J.M.K.; Methodology: M.S., A.N.T., and R.H.J.M.K.; Software: M.S. and N.E.; Validation: M.S.; Formal analysis: M.S., A.N.T., and R.H.J.M.K.; Investigation: M.S., A.N.T., and R.H.J.M.K.; Data curation: M.S. and N.E.; Writing—original draft: M.S., A.N.T., and R.H.J.M.K.; Writing—review and editing: M.S., A.N.T., N.E., P.L.-S., R.H., A.G., and R.H.J.M.K.; Supervision: A.N.T. and R.H.J.M.K.; Project administration: M.S., A.N.T., and R.H.J.M.K.; Funding acquisition: R.H.J.M.K., A.N.T., P.L-S., and R.H. (for more on the CRediT taxonomy, see; Holcombe et al., 2020).

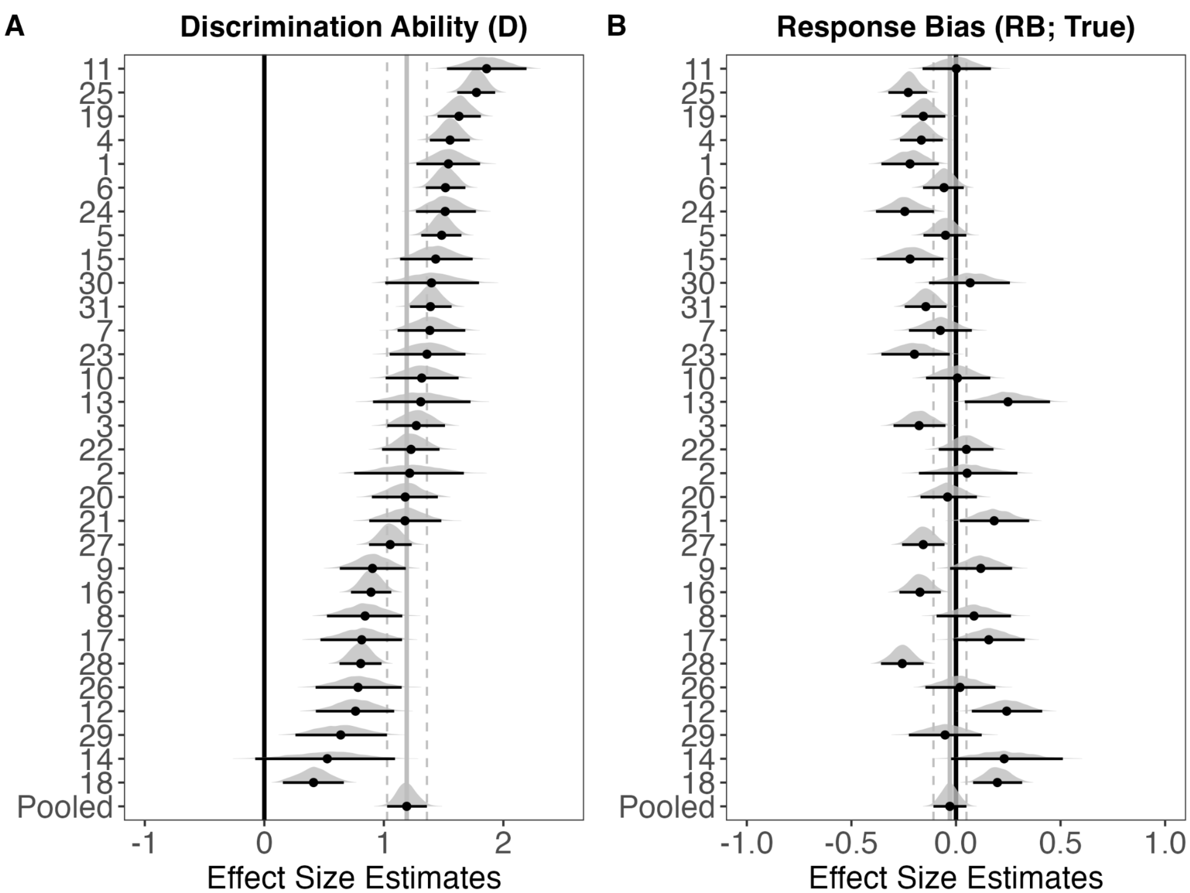
Competing Interests

The authors declare no competing interests.

Supplementary Materials

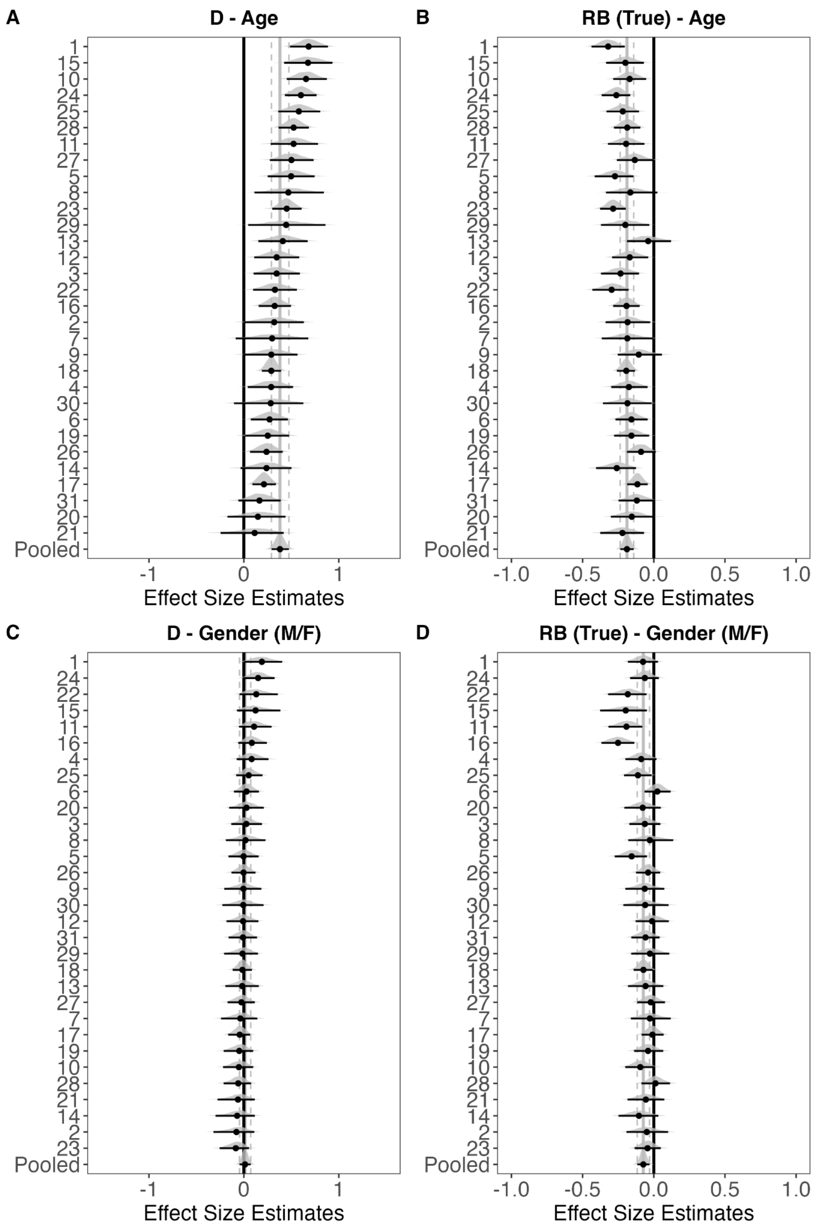
Supplementary Figure 1.

Study-level SDT model estimates for discrimination ability and response bias. All estimates derive from the SDT model in the main text (Figure 1A). Panel A shows the estimates for discrimination ability, with more positive (negative) values indicating higher (lower) discrimination ability. Panel B shows the estimates for response bias, with more positive (negative) values indicating a higher likelihood to judge headlines as true (false). Pooled: Overall estimate of discrimination ability (A) and response bias (B). Dots represent the mean and the error bars the 95% CI of the posterior distribution. All factors were mean centred. Studies are ordered from highest to lowest discrimination ability, and study numbers correspond to study numbers in Table 2.



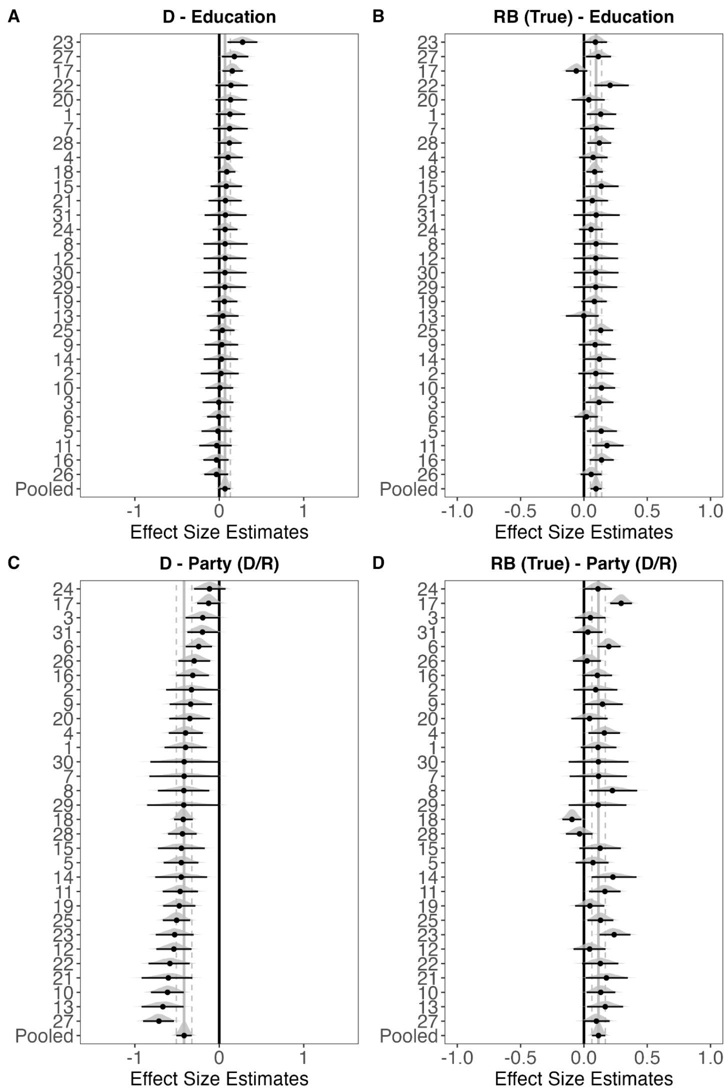
Supplementary Figure 2.

Study-level SDT estimates for discrimination ability and response bias for age and gender. All estimates derive from the SDT model in the main text (Figure 1A). Left panels (A, C) show the estimates for discrimination ability, with more positive (negative) values indicating higher (lower) discrimination ability. Right panels (B, D) show the estimates for response bias, with more positive (negative) values indicating a higher likelihood to judge headlines as true (false). D = discrimination ability. RB = response bias. Pooled: Overall estimate of discrimination ability (A, C) and response bias (B, D). Dots represent the mean and the error bars the 95% CI of the posterior distribution. All factors were mean centred. Studies in the upper (lower) panels are ordered from highest to lowest effect of age (gender) on discrimination ability, and study numbers correspond to study numbers in Table 2.



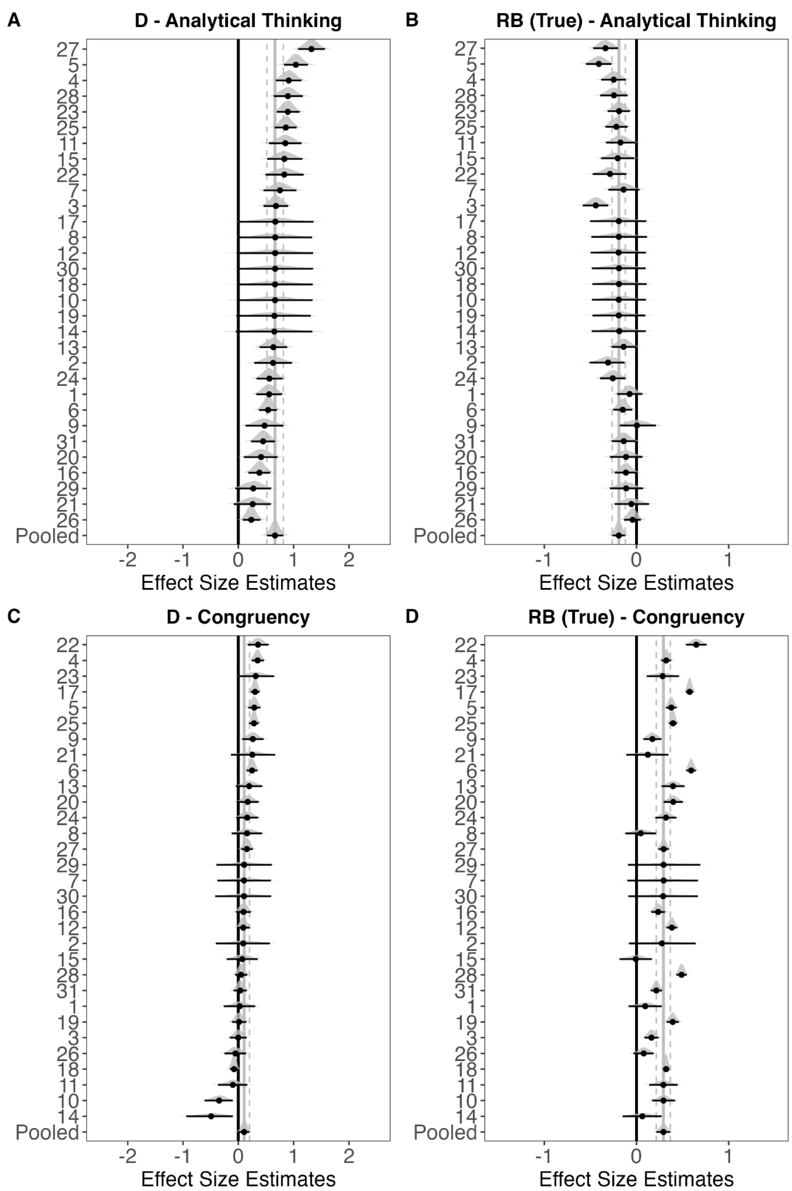
Supplementary Figure 3.

Study-level SDT estimates for discrimination ability and response bias for education and political identity. All estimates derive from the SDT model in the main text (Figure 1A). Left panels (A, C) show the estimates for discrimination ability, with more positive (negative) values indicating higher (lower) discrimination ability. Right panels (B, D) show the estimates for response bias, with more positive (negative) values indicating a higher likelihood to judge headlines as true (false). D = discrimination ability. RB = response bias. Pooled: Overall estimate of discrimination ability (A, C) and response bias (B, D). Dots represent the mean and the error bars the 95% CI of the posterior distribution. All factors were mean centred. Studies in the upper (lower) panels are ordered from highest to lowest effect of education (political identity) on discrimination ability, and study numbers correspond to study numbers in Table 2.



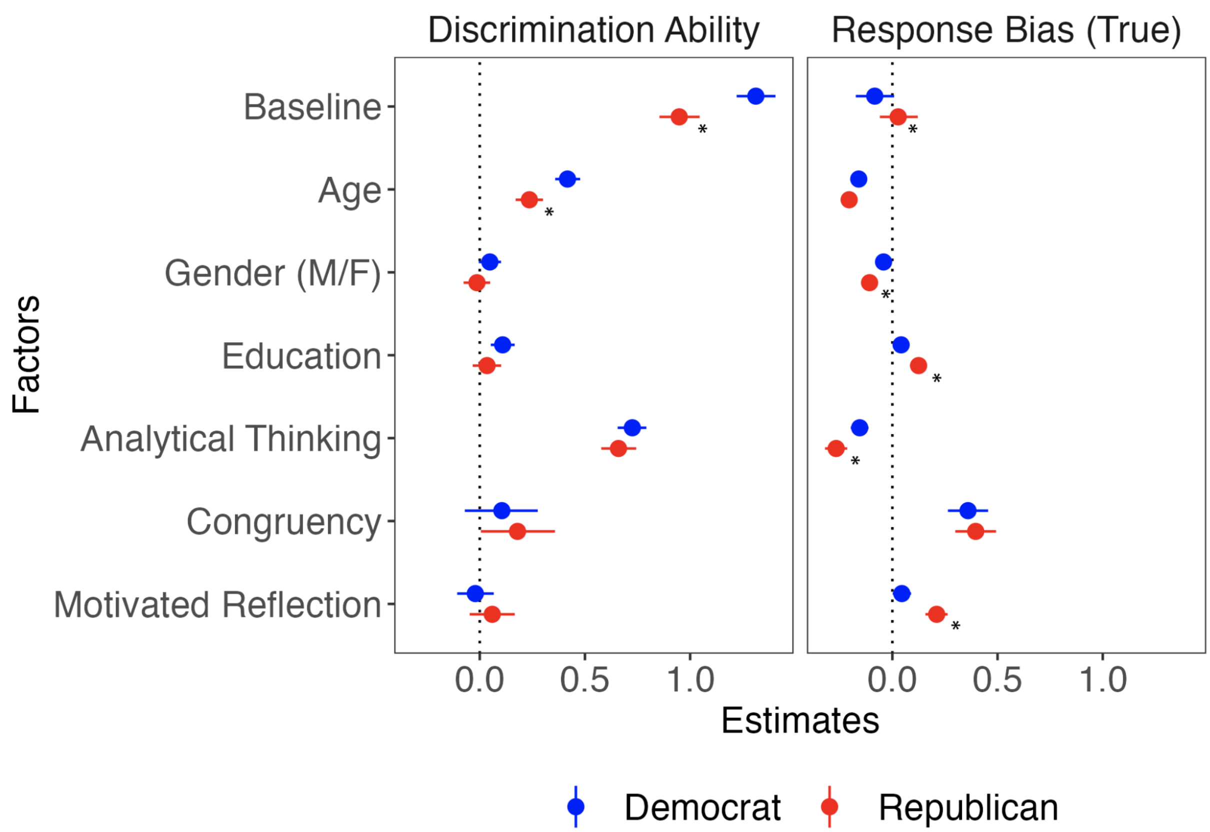
Supplementary Figure 4.

Study-level SDT estimates for discrimination ability and response bias for analytical thinking and ideological congruency. All estimates derive from the SDT model in the main text (Figure 1A). Left panels (A, C) show the estimates for discrimination ability, with more positive (negative) values indicating higher discrimination ability. Right panels (B, D) show the estimates for response bias, with more positive (negative) values indicating a higher (lower) likelihood to judge headlines as true (false). D = discrimination ability. RB = response bias. Pooled: Overall estimate of discrimination ability (A, C) and response bias (B, D). Dots represent the mean and the error bars the 95% CI of the posterior distribution. All factors were mean centred. Studies in the upper (lower) panels are ordered from highest to lowest effect of analytical thinking (congruency) on discrimination ability, and study numbers correspond to study numbers in Table 2.



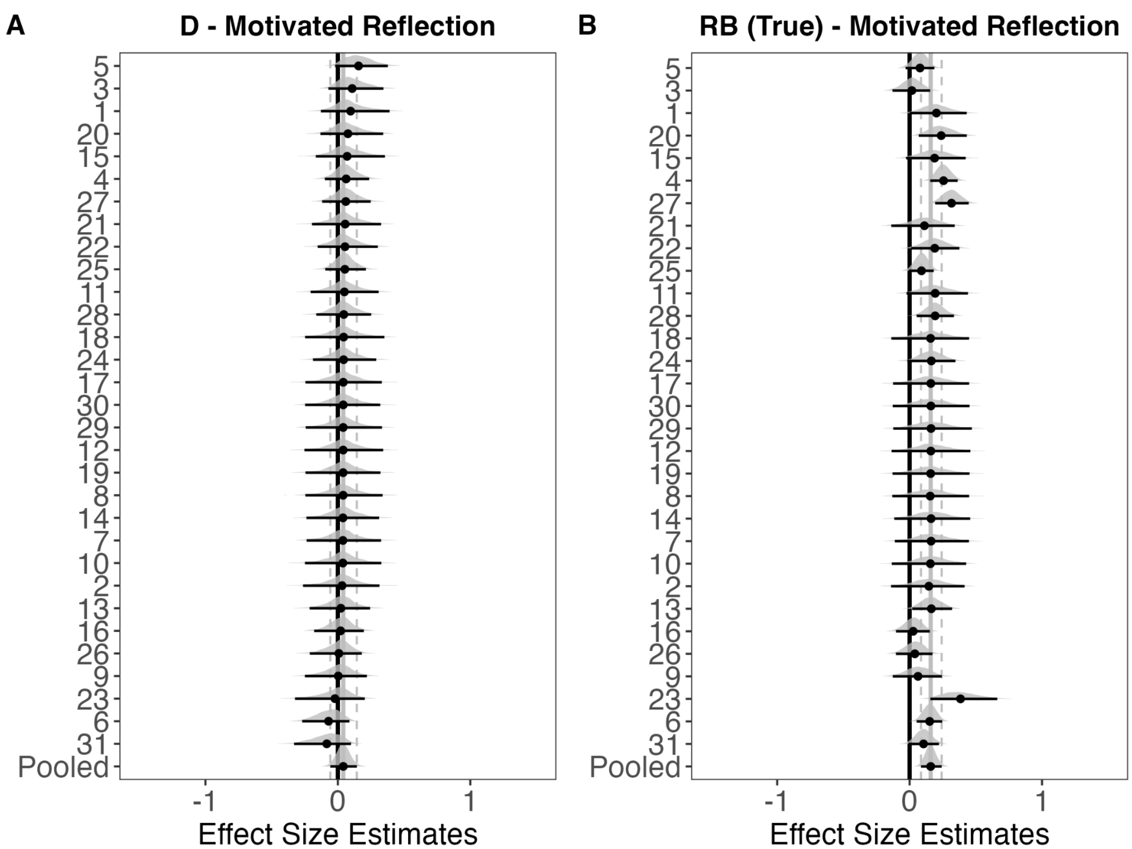
Supplementary Figure 5.

*SDT model estimates separated by Democrat (blue) and Republican (red) participants. The left panel shows the estimates for discrimination ability, with more positive (negative) values indicating higher (lower) discrimination ability. The right panel shows the estimates for response bias, with more positive (negative) values indicating a higher likelihood to judge headlines as true (false). Baseline: Overall estimate of discrimination ability (left panel) and response bias (right panel), separated for Democrats and Republicans. Gender (M/F) = coded Male to Female. Congruency = ideological congruency. Dots represent the mean and the error bars the 95% CI of the posterior distribution. Note that all factors were mean centred. * = credibly different effects between Democrat and Republican participants.*



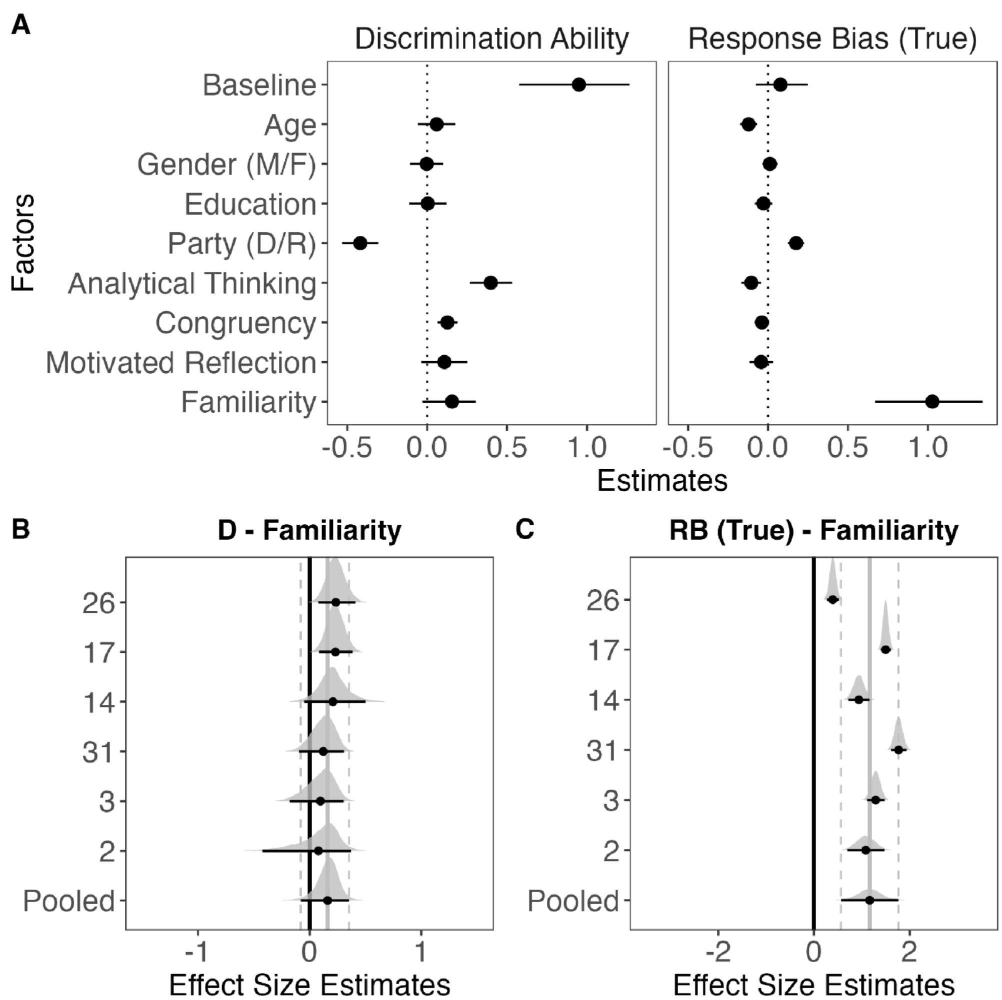
Supplementary Figure 6.

Study-level SDT estimates for discrimination ability and response bias for motivated reflection. All estimates derive from the SDT model in the main text (Figure 1A). Left panel (A) shows the estimates for discrimination ability, with more positive (negative) values indicating higher (lower) discrimination ability. Right panel (B) shows the estimates for response bias, with more positive (negative) values indicating a higher likelihood to judge headlines as true (false). D = discrimination ability. RB = response bias. Pooled: Overall estimate of discrimination ability (A) and response bias (B). Dots represent the mean and the error bars the 95% CI of the posterior distribution. All factors were mean centred. Studies are ordered from highest to lowest effect of motivation reflection on discrimination ability, and study numbers correspond to study numbers in Table 2.



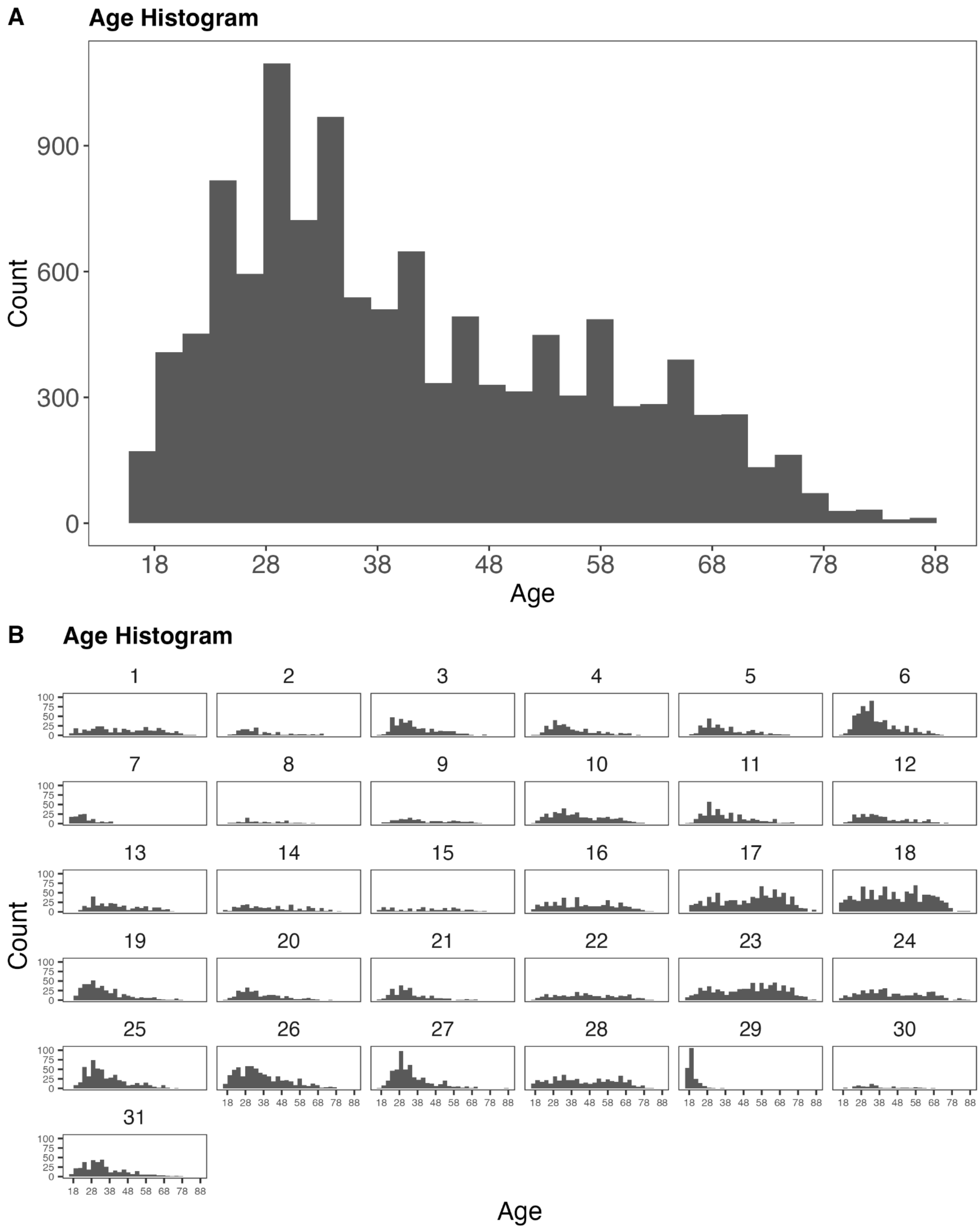
Supplementary Figure 7.

Panel A: Complete-case SDT model estimates for familiarity. All results derive from a single SDT analysis using participants' responses (false news or true news) as the response variable but are shown in two panels to ease interpretation. Panel B and Panel C: Study-level SDT estimates for discrimination ability and response bias for familiarity. All estimates derive from a complete-case SDT model for familiarity (A). The left panels show the estimates for discrimination ability, with more positive (negative) values indicating higher (lower) discrimination ability. The right panels show the estimates for response bias, with more positive (negative) values indicating a higher (lower) likelihood to judge headlines as true (false). Dots represent the mean and the error bars the 95% CI of the posterior distribution. Note that all factors were mean centred. In Panel B and Panel C, studies are ordered from highest to lowest effect of familiarity on discrimination ability, and study numbers correspond to study numbers in Table 2. Baseline: Overall estimate of discrimination ability (A, left panel) and response bias (A, right panel). Party = political identity. Congruency = ideological congruency. D = discrimination ability. RB = response bias. Pooled: Overall estimate of discrimination ability (B) and response bias (C).



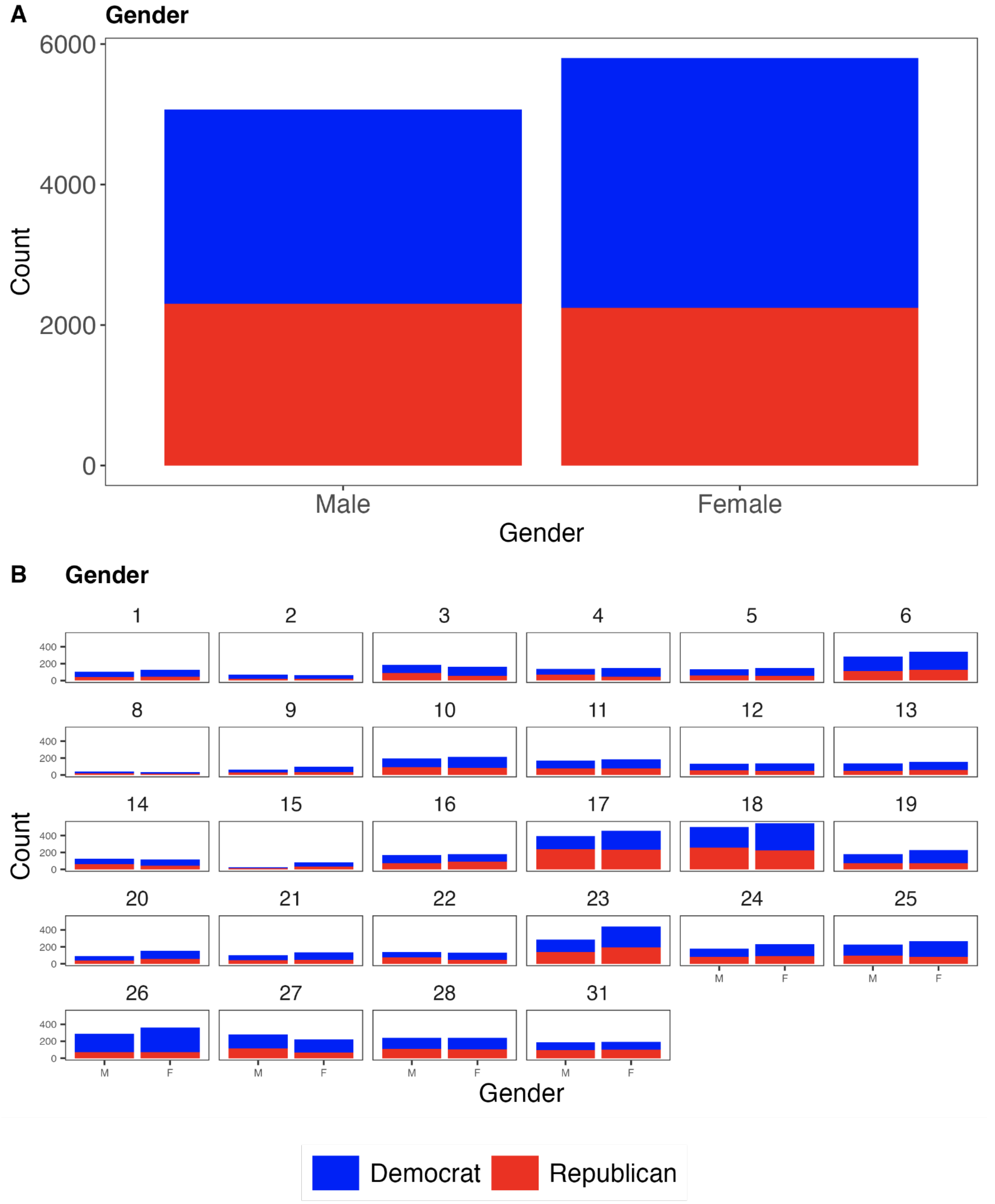
Supplementary Figure 8.

Histogram for age across all studies (A) and separated by study (B). The study numbers correspond to study numbers in Table 2.



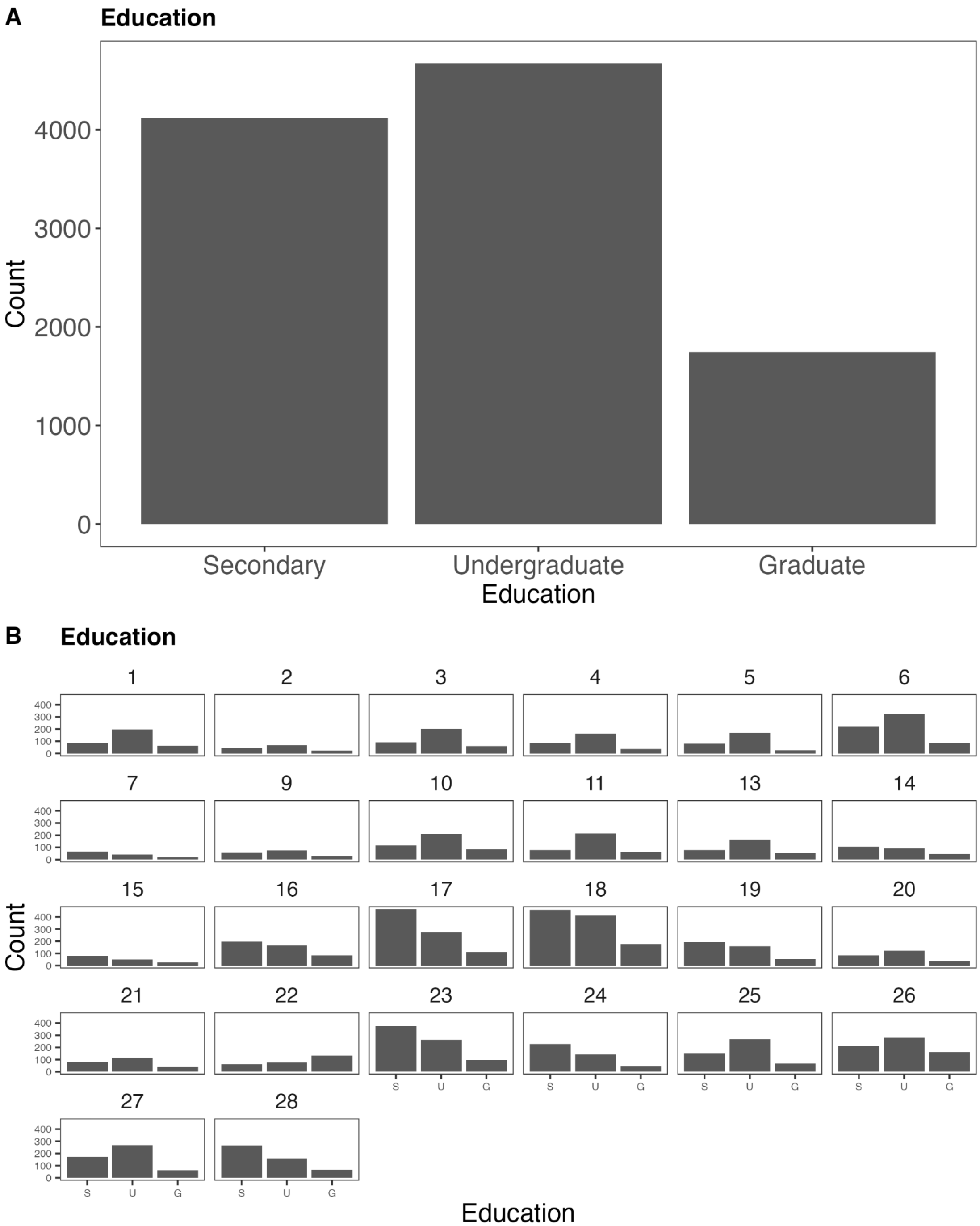
Supplementary Figure 9.

Histogram for gender grouped by political identity across all studies (A) and separated by study (B). The study numbers correspond to study numbers in Table 2. M: Male. F: Female.



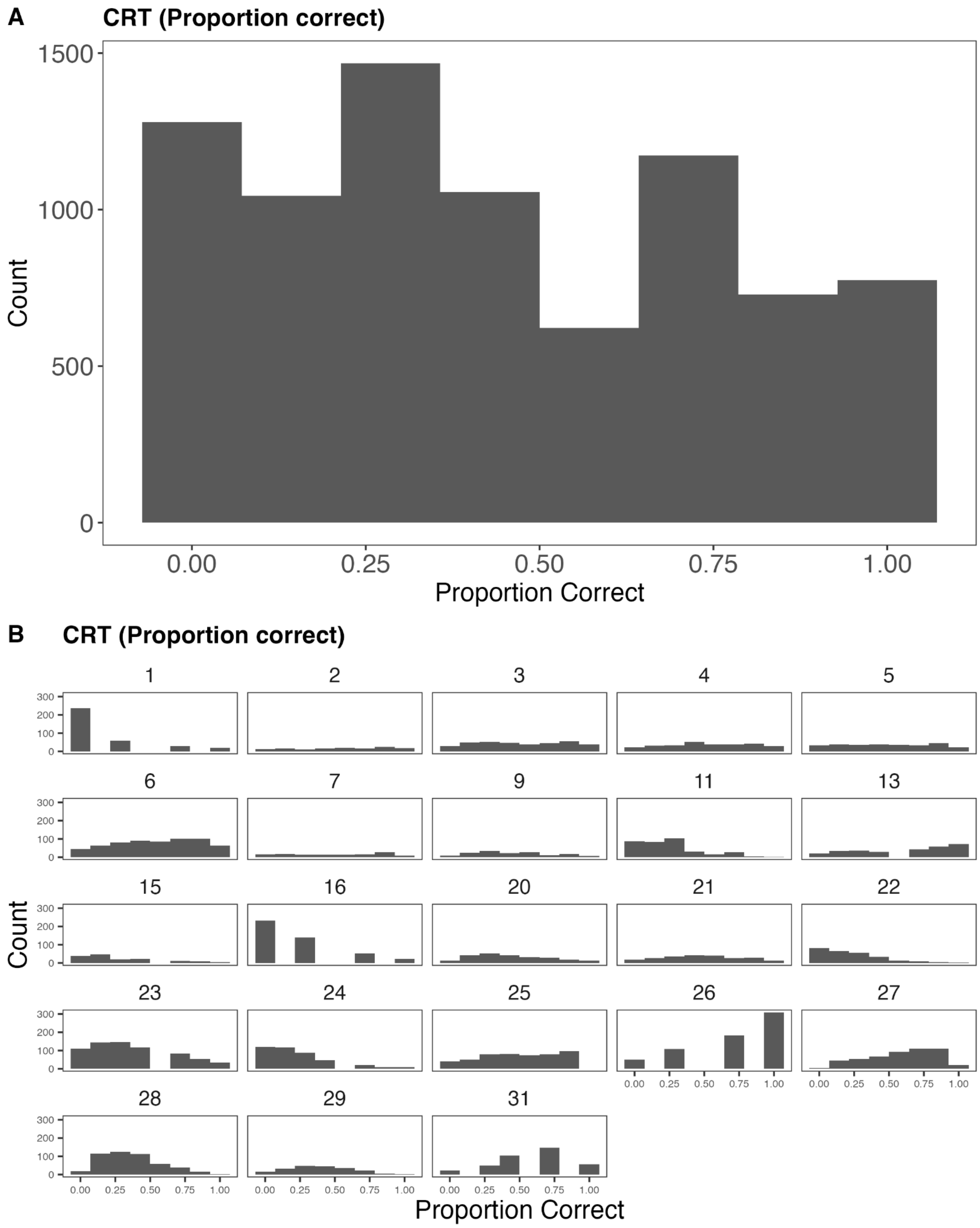
Supplementary Figure 10.

Histogram for education across all studies (A) and separated by study (B). The study numbers correspond to study numbers in Table 2. S: Secondary. U: Undergraduate. G: Graduate.



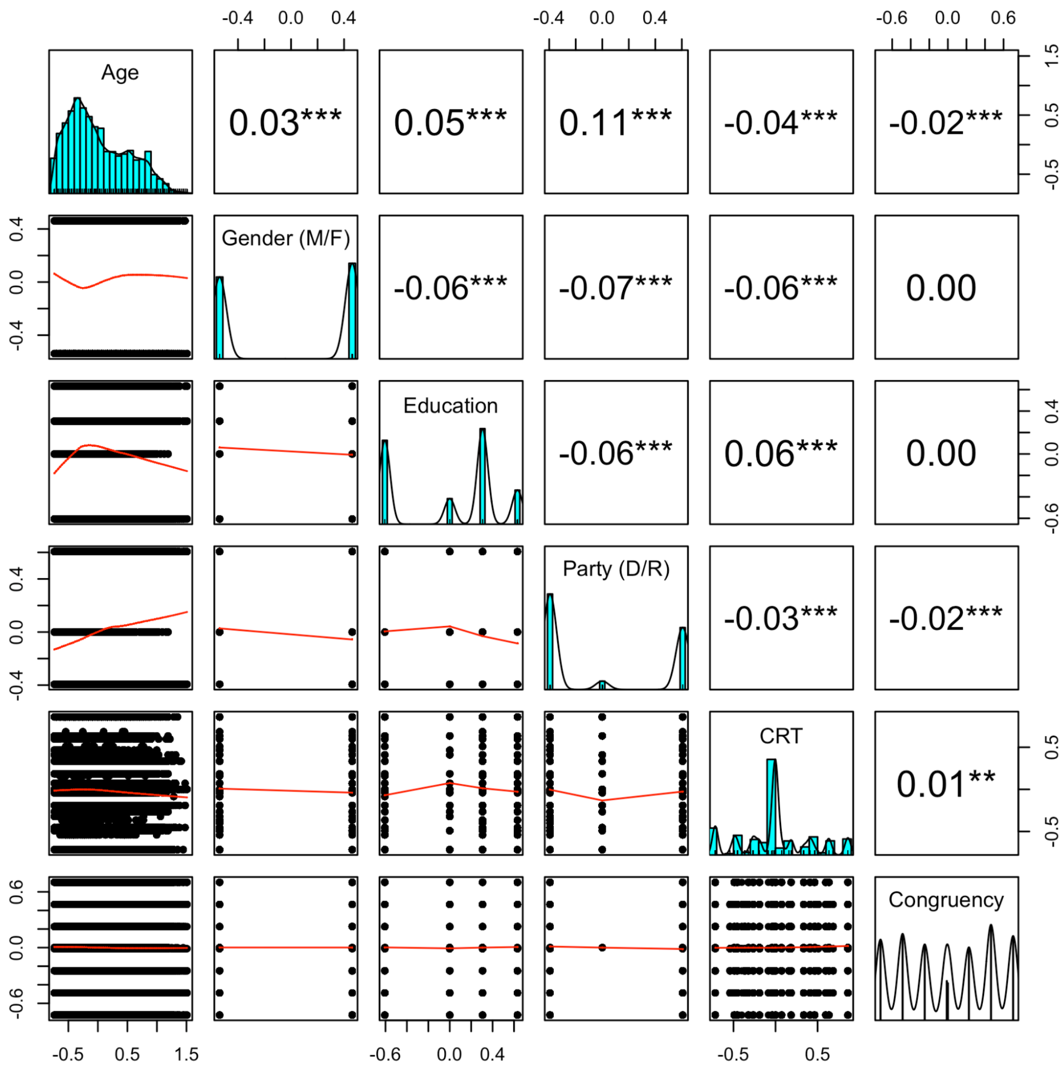
Supplementary Figure 11.

Histogram for CRT proportion scores across all studies (A) and separated by study (B). The study numbers correspond to study numbers in Table 2.



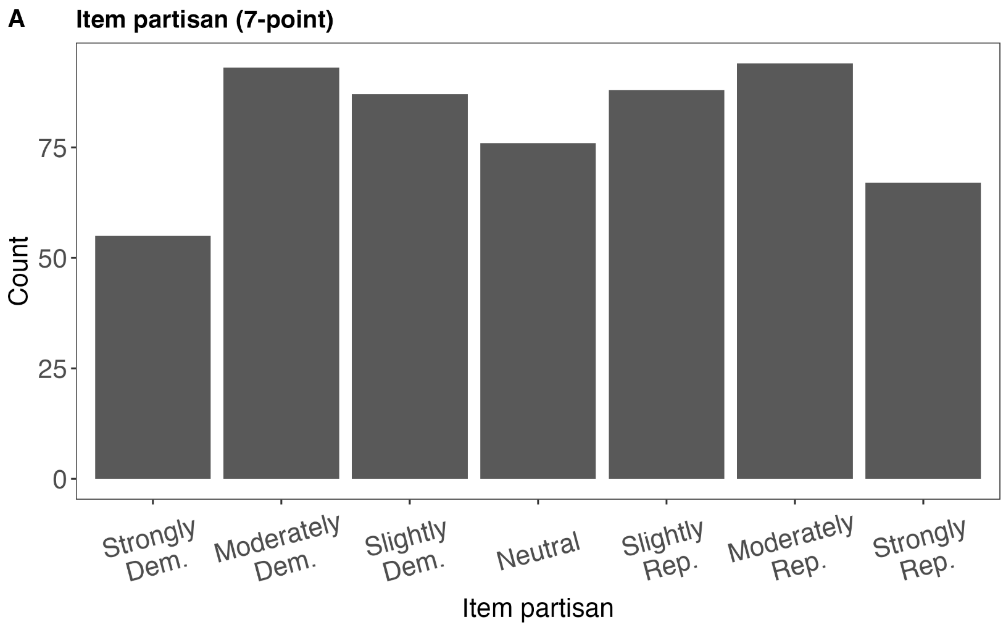
Supplementary Figure 12.

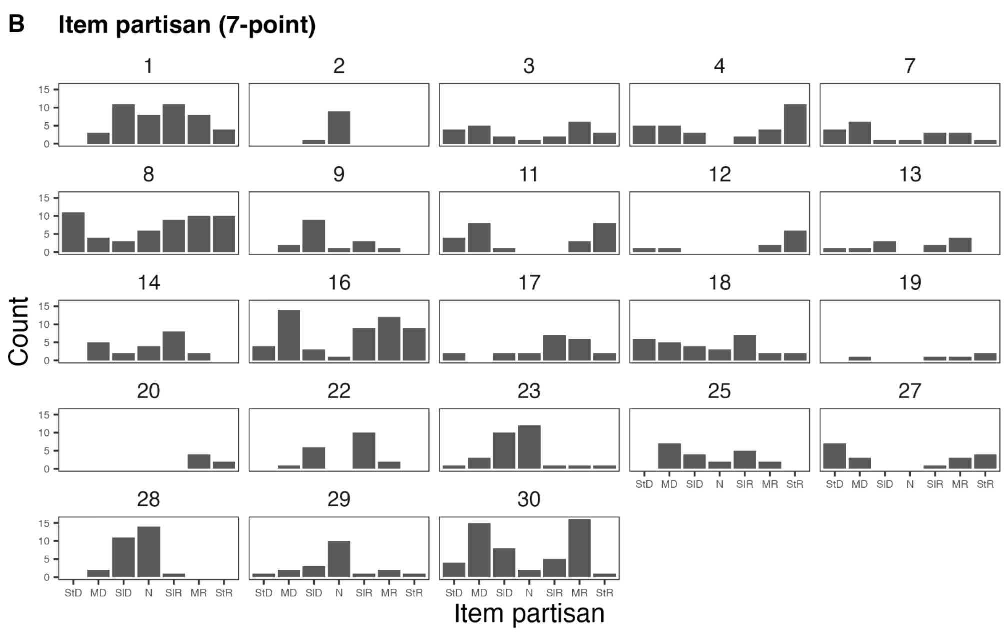
Correlations between the demographic and psychological factors. The diagonal shows histograms of each of the factors. The lower triangle visualises their associations. The upper triangle depicts the Pearson correlation coefficients. The factors were all very weakly correlated, with the strongest association being with age and political identity (0.11). Gender (M/F) = coded Male to Female. Party (D/R) = political identity, coded Democrat to Republican. Congruency = ideological congruency.



Supplementary Figure 13.

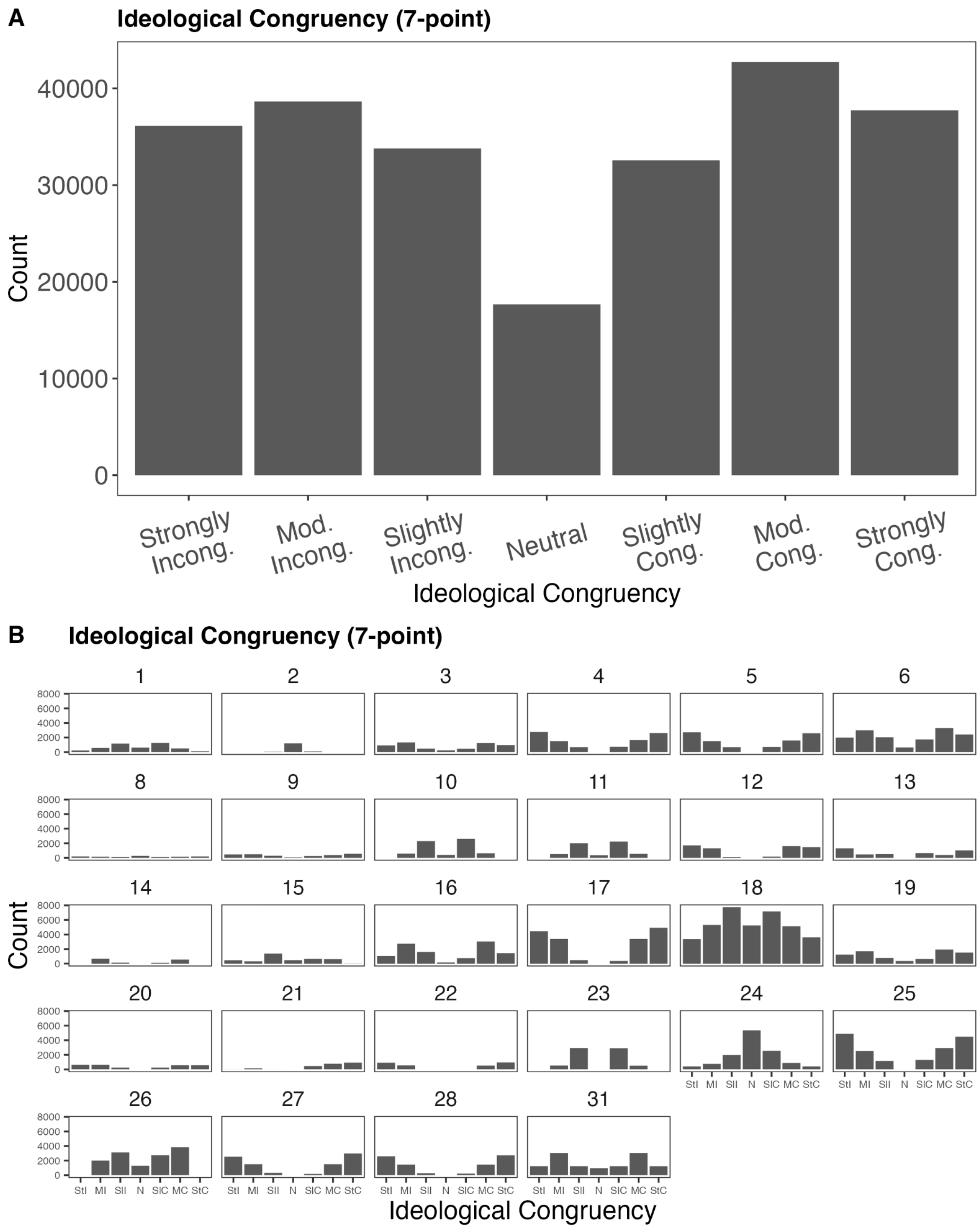
Histogram for partisan leanings of news headlines across all studies (A) and separated by study. The study numbers correspond to study numbers in Table 2. Dem: Democrat. Rep: Republican. StD: Strongly Democrat. MD: Moderately Democrat. SlD: Slightly Democrat. N: Neutral. SlD: Slightly Republican. MD: Moderately Republican. StD: Strongly Republican.





Supplementary Figure 14.

Histogram for ideological congruence (7-point) across all studies (A) and separated by study. The study numbers correspond to study numbers in Table 2. Incon: Incongruent. Cong: Congruent. Mod: Moderately. StI: Strongly Incongruent. MI: Moderately Incongruent. SlI: Slightly Incongruent. N: Neutral. SlC: Slightly Congruent. MC: Moderately Congruent. StC: Strongly Congruent.



Supplementary Table 1.

Search terms for Web of Science, Scopus, and PsycINFO. For Web of Science and Scopus, to be indexed, studies needed to mention at least one term from each category of interest (i.e., news AND misinformation AND veracity AND human). For PsycINFO, to be indexed, studies needed to mention at least one term from each category (i.e., subject headings AND custom terms AND custom terms). Custom terms refer to the search class ".tw.", which includes the following: words that appear in the table of contents, title, abstract, and key concepts.

Web of Science and Scopus categories

News	Misinformation	Veracity	Human
news	misinfo*	veracity	Men OR women*
headline*	disinfo*	accura*	male* OR female*
	fake*	real	adult*
	false	trustworthy	subject*
		credib*	participant*
		manipulative	individual*
		correct*	Democrat* OR Republican*
		true	Liberal* OR Conservative*
		susceptib*	"United States" OR "United States of America" OR US OR USA OR U.S. OR U.S.A.
		discern*	

PsycINFO categories

Subject headings	Custom terms	Custom terms
faking	misinformation	veracity
news media	fake news	accura*

information	fact-checking	real
social media	correction	trustworthy
messages	news media	credib*
exposure		manipulative
COVID-19		correct*
coronavirus		true
truth		susceptib*
false beliefs		discern*
skepticism		
political issues		

Supplementary Table 2.

List of extracted variables of interest

Variable	Description
Study and participant	
paper_ref	Paper reference (DOI)
study_id	Study identification number (1 to n). Note that one paper reference can contain multiple study_ids.
study_year	Year study conducted
study_platform	Name of platform used to conduct the study (e.g., Prolific)
study_base_rate	Base rate of true and false news headlines (e.g., 50% true: 50% false)
study_treatment	The name of the study/treatment the data is extracted from (e.g., control, active control, veracity-only)
part_id	Participant identification number (1 to n)
base_rate_informed	Whether participants are informed of study base rate or not (i.e., yes, no, NA)
News headline	
item_id	News headline identification number (1 to n; unique if possible)
item_veracity	Binary of item_veracity (i.e., false, true)
item_type	Type of news headline (e.g., political, covid, health)
item_type_GPT	Type of news headline as generated by GPT4 (i.e., political, covid, health)
item_presentation	Presentation mode of headline (e.g., headline only, headline and image)
item_generation	How news headlines were generated (i.e., human vs. AI)

item_partisan	Political leaning of news headlines (i.e., Republican, Democratic)
item_partisan_GPT	Political leaning of news headlines as generated by (i.e., Strongly Republican, Moderately Republican, Lean Republican, Neutral, Lean Democratic, Moderately Democratic, Strongly Democratic)
Demographic	
part_age	Participant age in years (minimum 18)
part_gender	Participant gender (i.e., male, female)
part_education_raw	Participant education raw (e.g., high school to doctoral level)
part_education	Participant education (i.e., Secondary, Undergraduate, and Graduate) turned into ridit scores.
part_political_identity_raw	Participant political identity (e.g., Strong Democrat, Moderate Democrat, Lean Democrat, Lean Republican, Moderate Republican, Strong Republican)
part_political_identity_binary	Binary of part_political_identity (i.e., Democrat, Republican)
Psychological	
CRT_type	Type of CRT administered (e.g., numeric, non-numeric, mixed)
CRT_number_questions	Number of CRT questions asked and answered (%)
CRT_correct	Number of correct CRT responses
CRT	Proportion of correct CRT responses (0-1)
ideological_congruency	Whether the political leaning of news headline is congruent (incongruent) with the participants' political leaning (i.e., Strongly incongruent, Moderately incongruent, Lean incongruent, Neutral, Lean congruent, Moderately congruent, Strongly congruent)
familiarity_response_mode	Response mode of familiarity question (e.g., binary, 6-point

	Likert scale)
familiarity_response_raw	Familiarity response (e.g., unfamiliar, familiar)
familiarity_response_binary	Binary of familiarity_response_raw (i.e., unfamiliar, familiar)
Veracity judgment	
veracity_framing	Question framing of veracity question (e.g., real-fake, credible)
veracity_response_mode	Response mode of veracity question (e.g., binary, 6-point Likert scale)
veracity_response_raw	Participant response (e.g. false/true [for binary response mode] or 1-6 [for 6-point Likert scale])
veracity_response_binary	Binary of veracity_response_raw (i.e., false, true)

Supplementary Table 3.

Aggregate descriptive statistics ($N_{participants}$ = 11,561)

Variable	Descriptor	Value
Age	Mean (SD)	41.29 (15.68)
	Range	18 - 88
Gender (%)	Female (Male)	53.91 (46.09)
Education (%)	Secondary	39.17
	Undergraduate	44.31
	Graduate	16.57
Political identity (%)	Republican (Democrat)	41.85 (58.15)

Note: There is missing data across the following variables: Education (10.71%; 4/31 studies) and Political identity (6.52%; 3/31 studies).

Supplementary Section 1: Re-coding Partisan Leaning of Headlines

To calculate the partisan leanings of each headline, the following two prompts were used and in the following order.

- Prompt 1: “You will be presented with headlines and will have to answer the following question regarding the headline: Assuming the headline is entirely accurate, how favourable is it to Democrats (liberals) or Republicans (conservatives)? In a forced choice manner, you can ONLY respond using these two options, either as Favourable for Democrats (democrat_leaning) or Favourable for Republicans (republican_leaning). Your response should strictly adhere to this binary choice (democrat_leaning or republican_leaning) without exceptions, even if the headline appears neutral or unrelated to US politics. In your answer, first provide the partisan leaning (democrat_leaning or republican_leaning). This should be separated by ‘<separate>’, which will follow your explanation. Here's the headline: [HEADLINE TEXT]”
- Prompt 2: “Now, answer the same question using the following seven options: either as Strongly Favourable for Democrats (Strongly Democrat), Moderately Favourable for Democrats (Moderately Democrat), Slightly Favourable for Democrats (Slightly Democrat), Neutral (Neutral), Slightly Favourable for Republicans (Slightly Republican), Moderately Favourable for Republicans (Moderately Republican), or Strongly Favourable for Republicans (Strongly Republican). Note that you can now respond with a ‘Neutral’, making it no longer a forced choice between democratic_leaning or republican_leaning. In your answer, first provide the partisan leaning (Strongly Democrat, Moderately Democrat, Slightly Democrat, Neutral, Slightly Republican, Moderately Republican, and Strongly Republican). This should be separated by ‘<separate>’, which will follow your explanation. Here's the headline: [HEADLINE TEXT].”

Note that the separator (“<separate>”) was used to distinguish between the categorisation and the explanation during data processing.

Supplementary Section 2: Full Model Specification

```
model <- brm(veracity_response_binary ~
              d_prime*age + # age
              d_prime*gender  + # gender
              d_prime*education  + # education
              d_prime*pol_party  + # political identity
              d_prime*CRT + # CRT
              d_prime*congruency + # congruency
              d_prime*CRT*congruency + # MR

              (1 + d_prime | part_id) +
              (1 + d_prime*age + d_prime*gender  + d_prime*education  + d_prime*pol_party  +
              d_prime*CRT + d_prime*congruency + d_prime*CRT*congruency || study_id) +
              (1 | item_id),

            data = dat_model_mean_impute,
            family = bernoulli(link = "probit"),
            init = 0,
            iter = 10000,
            chains = 4
)
```