Article

Spotting false news and doubting true news: a systematic review and meta-analysis of news judgements

Jan Pfänder ◉[1] **& Sacha Altay** ◉[2] ✉

How good are people at judging the veracity of news? We conducted a systematic literature review and pre-registered meta-analysis of 303 effect sizes from 67 experimental articles evaluating accuracy ratings of true and fact-checked false news ($N_{\text{Participants}}$ = 194,438 from 40 countries across 6 continents). We found that people rated true news as more accurate than false news (Cohen's d = 1.12 [1.01, 1.22]) and were better at rating false news as false than at rating true news as true (Cohen's d = 0.32 [0.24, 0.39]). In other words, participants were able to discern true from false news and erred on the side of skepticism rather than credulity. We found no evidence that the political concordance of the news had an effect on discernment, but participants were more skeptical of politically discordant news (Cohen's d = 0.78 [0.62, 0.94]). These findings lend support to crowdsourced fact-checking initiatives and suggest that, to improve discernment, there is more room to increase the acceptance of true news than to reduce the acceptance of fact-checked false news.

Many have expressed concerns that we live in a 'post-truth' era and that people cannot tell the truth from falsehoods anymore. In parallel, populist leaders around the world have tried to erode trust in the news by delegitimizing journalists and the news media[1]. Since the 2016 US presidential election, our systematic literature review shows that over 4,000 scientific articles have been published on the topic of false news. Across the world, numerous experiments evaluating the effect of interventions against misinformation or susceptibility to misinformation have relied on a similar design feature: having participants rate the accuracy of true and fact-checked false headlines, typically in a Facebook-like format, with an image, title, lede and source, or as an isolated title/claim. Taken together, these studies allow us to shed some light on the most common fears voiced about false news, namely, that people may fall for false news, distrust true news, or may be unable to discern true from false news. In particular, we investigated whether people rate true news as more accurate than fact-checked false news (discernment) and whether they were better at rating false news as inaccurate than at rating true news as accurate

(skepticism bias). We also tested various moderators of discernment and skepticism bias such as political congruence, the topic of the news, or the presence of a source.

Establishing whether people can spot false news is important to design interventions against misinformation. If people lack the skills to detect false news, interventions should focus on improving these skills. However, if people have the ability to spot false news yet buy into it anyway, the problem lies elsewhere and may be one of motivation or (in)attention that educational interventions will struggle to address.

Past work has reliably shown that people do not fare better than chance at detecting lies because most verbal and non-verbal cues people use to detect lies are unreliable[2]. Why would this be any different for detecting false news? People make snap judgements to evaluate the quality of the news they come across[3] and rely on seemingly imperfect proxies such as the source of information, police and fonts, the presence of hyperlinks, the quality of visuals, ads, or the tone of the text[4,5]. In experimental settings, participants report relying on intuitions and tacit knowledge to judge the accuracy of

[1]Institut Jean Nicod, Département d'études cognitives, ENS, EHESS, PSL University, CNRS, Paris, France. [2]Department of Political Science, University of Zurich, Zürich, Switzerland. ✉e-mail: sacha.altay@gmail.com

news headlines[6]. Yet, a scoping review of the literature on belief in false news (including a total of 26 articles) has shown that, in experiments, participants 'can detect deceitful messages reasonably well'[7]. Similarly, a survey on 150 misinformation experts has shown that 53% of experts agreed that 'people can tell the truth from falsehoods', while only 25% of experts disagreed with the statement[6]. Unlike the unreliable proxies people rely on to detect lies in interpersonal contexts, there are reasons to believe that some of the cues people use to detect false news may, on average, be reliable. For instance, the news outlets people trust the least do publish lower-quality news and more false news, as people's trust ratings of news outlets correlate strongly with fact-checkers' ratings in the United States and Europe[8,9]. Moreover, false news has some distinctive properties, such as being more politically slanted[10], being more novel, surprising, or disgusting, being more sensationalist, funnier, less boring and less negative[11,12], or being more interesting-if-true[13]. These features aim at increasing engagement, but they do so at the expense of accuracy and in many cases, people may pick up on it. This led us to pre-register the hypothesis that people would rate true news as more accurate than false news. Yet, legitimate concerns have been raised about the lack of data outside of the United States, especially in some Global South countries where the misinformation problem is arguably worse. Our meta-analysis covers 40 countries across 6 continents and directly addresses concerns about the over-representation of US data.

People rate true news as more accurate than false news

While many fear that people are exposed to too much misinformation, too easily fall for it and are overly influenced by it, a growing body of researchers is worried that people are exposed to too little reliable information, commonly reject it and are excessively resistant to it[14,15]. Establishing whether true news skepticism (excessively rejecting true news) is of similar magnitude as false news gullibility (excessively accepting false news) is important for future studies on misinformation: if people are excessively gullible, interventions should primarily aim at fostering skepticism, whereas if people are excessively skeptical, interventions should focus on increasing trust in reliable information. For these reasons, in addition to investigating discernment (H1), we also looked at skepticism bias (H2) by comparing the magnitude of true news skepticism to false news gullibility. Research in psychology has shown that people exhibit a 'truth bias'[16,17], a tendency to accept incoming statements rather than to reject them. Similarly, work on interpersonal communication has shown that, by default, people tend to accept communicated information[18]. However, there are reasons to think that the truth-default theory may not apply to news judgements. It has been hypothesized that people display a truth bias in interpersonal contexts because information in these contexts is, in fact, often true[16]. When it comes to news judgements, it is not clear whether people by default expect news stories to be true. Trust in the news and journalists is low worldwide[19], and a substantial part of the population holds cynical views of the news[20]. Similarly, populist leaders across the world have attacked the credibility of the news media and instrumentalized the concept of fake news to discredit quality journalism[21,22]. Disinformation strategies such as 'flooding the zone' with false information[23,24] have been shown to increase skepticism in news judgements[6]. Moreover, in many studies included in our meta-analysis, the news stories were presented in a social media format (most often Facebook), which could fuel skepticism in news judgements. People trust news[3]—and information more generally[25]—less on social media than on news websites. In line with these observations, some empirical evidence suggests that for news judgements, people display the opposite of a truth bias[26], namely, a skepticism bias, and view true news as false[6,27,28]. We thus predicted that when judging the accuracy of news, participants will err on the side of skepticism more than on the side of gullibility.
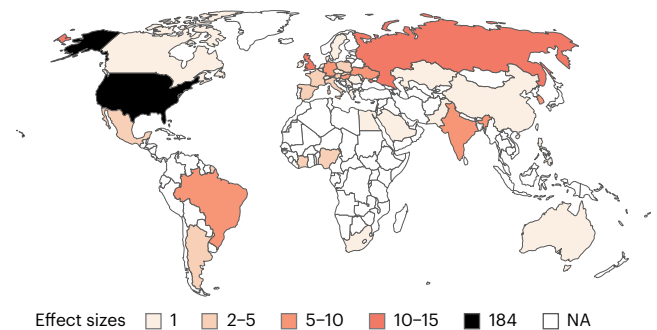

Fig. 1 | Map of the number of effect sizes per country. Countries are coloured based on the number of effect sizes. The darker a country, the more effect sizes. Countries in white are not covered in the sample.

People are better at rating false news as false than true news as true

Finally, we investigated potential moderators of H1 and H2, such as the country where the experiment was conducted, the format of the news headlines, the topic, whether the source of the news was displayed, and the political concordance of the news. Past work suggests that displaying the source of the news has a small effect at best on accuracy ratings[29], whereas little work has investigated differences in news judgements across countries, topics and formats. The effect of political concordance on news judgements is debated. Participants may be motivated to believe politically congruent (true and false) news, motivated to disbelieve politically incongruent news, or not be politically motivated at all but still display such biases[30]. We formulated research questions instead of hypotheses for our moderator analyses because of a lack of strong theoretical expectations.

Results

Descriptives

We conducted a systematic literature review and pre-registered the meta-analysis based on 67 publications, providing data on 195 samples (194,438 participants) and 303 effects (that is k, the meta-analytic observations). Our meta-analysis includes publications from 40 countries across 6 continents. However, 34% of all participants were recruited in the United States alone, and 54% in Europe. Only 6% of participants were recruited in Asia, and even less in Africa (2%; see Fig. 1 for the number of effect sizes per country). The average sample size was 997.12 (minimum = 19, maximum = 32,134, median = 482).

In total, participants rated the accuracy of 2,167 unique news items. On average, a participant rated 19.76 news items per study (min. = 2, max. = 240, median = 18). For 71 samples, news items were sampled from a pool of news (the pool size ranged from 12 to 255, with an average pool size of 57.46 items). The vast majority of studies (294 out of 303 effects) used a within-participant design for manipulating news veracity, with each participant rating both true and false news items. Almost all effect sizes are from online studies (286 out of 294).

Analytic procedures

All analyses were pre-registered unless explicitly stated otherwise (for deviations, see Methods). The choice of models was informed by simulations we conducted before having the data. To test H1, we calculated a discernment score by subtracting the mean accuracy ratings of false news from the mean accuracy ratings of true news, such that higher scores indicate better discernment. This differential measure of discernment is common in the literature on misinformation[31]. To test H2, we first calculated a judgement error for true and false news. Error is defined as the distance between optimal accuracy ratings and actual accuracy ratings (Fig. 2). We then calculated the skepticism bias as the difference between the two errors, subtracting the false

news error score from the true news error score. Note that we cannot use more-established Signal Detection Theory (SDT) measures because we rely on mean ratings and not individual ratings. However, in Supplementary Section H, we show that for the studies we have raw data on, our main findings hold when relying on d' (sensitivity) and c (response bias) from SDT.

To be able to compare effect sizes across different scales, we calculated Cohen's d, a common standardized mean difference. To account for statistical dependence between true and false news ratings arising from the within-participant design used by most studies (294 out of 303 effect sizes), we calculated the standard error following the Cochrane recommendations for crossover trials[32]. For the remaining 9 effect sizes from studies that used a between-participant design, we calculated the standard error assuming independence between true and false news ratings (see Methods). In Supplementary Section A, we show that our results hold across alternative standardized effect measures, including the one we originally pre-registered—a standardized mean change using change score standardization (SMCC). We chose to deviate from the pre-registration and use Cohen's d instead, because it is easier to interpret and corresponds to the standards recommended by the Cochrane manual[32]. In Supplementary Section A, we also provide effect estimates in units of the original scales separately for each scale.

We used multilevel meta-models with clustered standard errors at the sample level to account for cases in which the same sample contributed various effect sizes (that is, the meta-analytic units of observation). All confidence intervals reported in this paper are 95% confidence intervals. All statistical tests are two-tailed.

Main results

Discernment. Supporting H1, participants rated true news as more accurate than false news on average. Pooled across all studies, the average discernment estimate is large ($d = 1.12$ [1.01, 1.22], $z = 20.79$, $P < 0.001$). As shown in Fig. 3, 298 of 303 estimates are positive. Of the positive estimates, three have a confidence interval that includes zero, as does one of the negative estimates. Most of the variance in the effect sizes observed above is explained by between-sample heterogeneity ($I2_{between} = 92.04\%$). Within-sample heterogeneity is comparatively small ($I2_{within} = 7.93\%$), indicating that when the same participants were observed on several occasions (that is, the same sample contributed several effect sizes), on average, discernment performance was similar across those observations. The share of the variance attributed to sampling error is very small (0.03%), which is indicative of the large sample sizes and thus precise estimates.

Skepticism bias. We found support for H2, with participants better able to rate false news as inaccurate than to rate true news as accurate (that is, false news discrimination was on average higher than true news discrimination). However, the average skepticism bias estimate is small ($d = 0.32$ [0.24, 0.39], $z = 8.11$, $P < 0.001$). As shown in Fig. 3, 203 of 303 estimates are positive. Of the positive estimates, six have a confidence interval that includes zero, as do seven of the negative estimates. By contrast with discernment, most of the variance in skepticism bias is explained by within-sample heterogeneity ($I2_{within} = 60.96\%$; $I2_{between} = 38.99\%$; sampling error = 0.05%). Whenever we observe within-sample variation in our data, it is because several effects were available for the same sample. This is mostly the case for studies with multiple survey waves, or when effects were split by different news topics, suggesting that these factors may account for some of that variation. In the moderator analyses below, most variables vary between samples, thereby glossing over much of that within-sample variation, except for political concordance.

Moderators

Following the pre-registered analysis plan, we ran a separate meta-regression for each moderator by adding the respective moderator
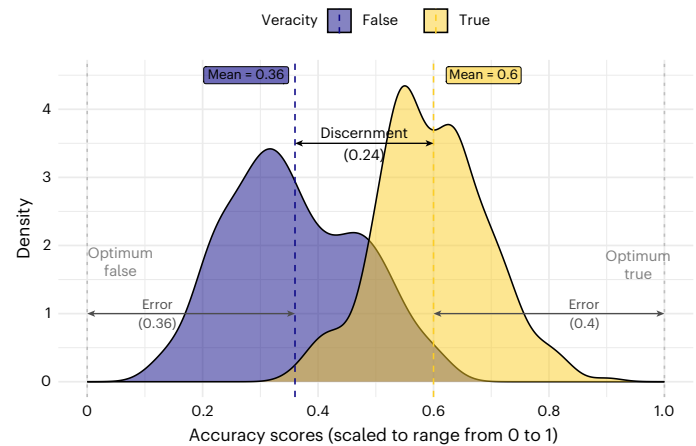

Fig. 2 | Illustration of outcome measures. Distributions of accuracy ratings for true and fact-checked false news, scaled to range from 0 to 1. The figure illustrates discernment (the distance between the mean for true news and the mean for false news) and the errors (distance to the right end for true news and to the left end for false news) from which the skepticism bias is computed. A larger error for true news compared with false news yields a positive skepticism bias. In this descriptive figure, unlike in the meta-analysis, ratings and outcomes sizes are not weighted by sample size.

variable as a fixed effect to the multilevel meta-models. We report regression tables and visualizations in Supplementary Section B. Here we report the regression coefficients as 'Deltas', since they designate differences between categories. For example, in the moderator analysis of political concordance on skepticism bias, 'concordant' marks the baseline category. The predicted value for this category can be read from the intercept (−0.2). The 'Delta' is the predicted difference between concordance and discordance (0.78). To obtain the predicted value for discordant news, one needs to add the 'Delta' to the intercept (−0.2 + 0.78 = 0.58).

Cross-cultural variability. For samples based in the United States (184/303 effect sizes), discernment was higher on average than for samples based in other countries (ΔDiscernment = 0.23 [0.02, 0.44], $z = 2.14$, $P = 0.033$; baseline discernment other countries pooled = 0.99 [0.84, 1.14], $z = 12.82$, $P < 0.001$). However, we did not find a statistically significant difference regarding skepticism bias (ΔSkepticism bias = 0.04 [−0.12, 0.19], $z = 0.47$, $P = 0.638$). A visualization of discernment (F1) and skepticism bias (F2) across countries can be found in Supplementary Section F.
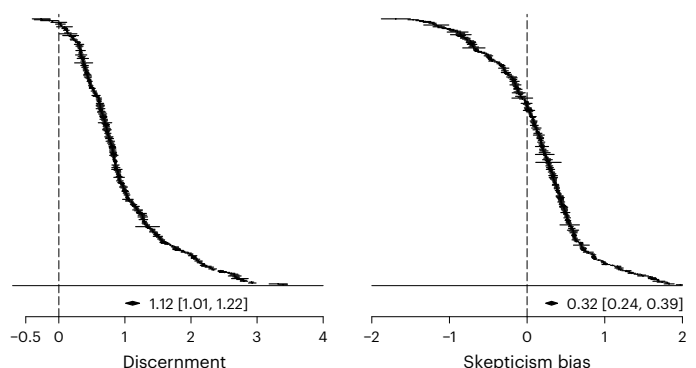

Fig. 3 | Forest plots for discernment and skepticism bias. All $n = 303$ effect sizes for both outcomes. Effects are weighted by their sample size. Effect sizes are calculated as Cohen's d. Horizontal bars represent 95% confidence intervals. The average estimate is the result of a multilevel meta model with clustered standard errors at the sample level.

Scales. The studies in our meta-analysis used a variety of accuracy scales, including both binary (for example, "Do you think the above headline is accurate? - Yes, No") and continuous ones (for example, "To the best of your knowledge, how accurate is the claim in the above headline? 1 = Not at all accurate, 4 = Very accurate").

Regarding discernment, two scale types differed from the most common 4-point scale (baseline discernment 4-point-scale = 1.28 [1.07, 1.49], $z = 11.96$, $P < 0.001$): both 6-point scales (ΔDiscernment = −0.41 [−0.7, −0.12], $z = −2.80$, $P = 0.006$) and binary scales (ΔDiscernment = −0.37 [−0.66, −0.08], $z = −2.50$, $P = 0.013$) yielded lower discernment. Regarding skepticism bias, studies using a 4-point scale (baseline skepticism bias 4-point scale = 0.51 [0.3, 0.72], $z = 0.75$, $P < 0.001$) reported a larger skepticism bias compared with studies using a binary or a 7-point scale (ΔSkepticism bias = −0.29 [−0.51, −0.06], $z = −2.47$, $P = 0.014$ for binary scales; −0.50 [−0.76, −0.23], $z = −3.67$, $P < 0.001$ for 7-point scales). Interpreting these observed differences is not straightforward; we further discuss differences between binary and Likert-scale studies in Supplementary Section D.

Format. Studies using headlines with pictures as stimuli (ΔSkepticism bias = 0.22 [0.04, 0.39], $z = 2.45$, $P = 0.015$; 65 effects), or headlines with pictures and a lede (ΔSkepticism bias = 0.33 [0.14, 0.52], $z = 3.40$, $P < 0.001$; 56 effects), displayed a stronger skepticism bias compared with studies relying on headlines with no picture/lede (baseline skepticism bias headlines only = 0.23 [0.13, 0.33], $z = 4.45$, $P < 0.001$; 163 effects). We do not find differences related to format for discernment, neither for headlines with pictures (ΔDiscernment = −0.01 [−0.28, 0.27], $z = −0.04$, $P = 0.969$), nor for headlines with pictures and a lede (ΔDiscernment = 0.11 [−0.12, 0.33], $z = 0.93$, $P = 0.353$).

Topic. We did not find statistically significant differences in discernment and skepticism bias across news topics, when distinguishing between the categories 'political' (ΔSkepticism bias = 0.03 [−0.13, 0.19], $z = 0.43$, $P = 0.671$; ΔDiscernment = −0.26 [−0.51, 0], $z = −1.98$, $P = 0.049$; 196 effects; 43 articles), 'covid-19' (baseline; 54 effects; 13 articles) and 'other' (ΔSkepticism bias = −0.02 [−0.2, 0.16], $z = −0.22$, $P = 0.825$; ΔDiscernment = −0.01 [−0.35, 0.34], $z = −0.03$, $P = 0.976$; 53 effects; 20 articles), a category regrouping all news not explicitly labelled as 'covid-19' or 'political', including news about health, cancer, science, economics, history or military matters.

Sources. In line with past findings, we did not observe a statistically significant difference in discernment between studies that displayed the source of the news items (ΔDiscernment = −0.22 [−0.47, 0.03], $z = −1.75$, $P = 0.082$; 112 effects) and studies that did not (147 effects; for 44 effects, this information was not explicitly provided). We did not find a difference regarding skepticism bias either (ΔSkepticism bias = 0.11 [−0.06, 0.29], $z = 1.30$, $P = 0.194$).

Political concordance. The moderators investigated above were (mostly) not experimentally manipulated within studies, but instead varied between studies, which impedes causal inference. Political concordance is an exception in this regard. It was manipulated within 31 different samples, across 14 different papers. In those experiments, typically, a pre-test establishes the political slant of news headlines (for example, pro-Republican vs pro-Democrat). Participants then rate the accuracy for news items of both political slants and provide information about their own political stance. The ratings of items are then grouped into concordant or discordant (for example, pro-Republican news rated by Republicans will be coded as concordant while pro-Republican news rated by Democrats will be coded as discordant).

Political concordance had no statistically significant effect on discernment (ΔDiscernment = 0.08 [−0.01, 0.17], $z = 1.72$, $P = 0.097$). It did, however, make a difference regarding skepticism bias (Fig. 4): when rating concordant items, there was no evidence that participants showed
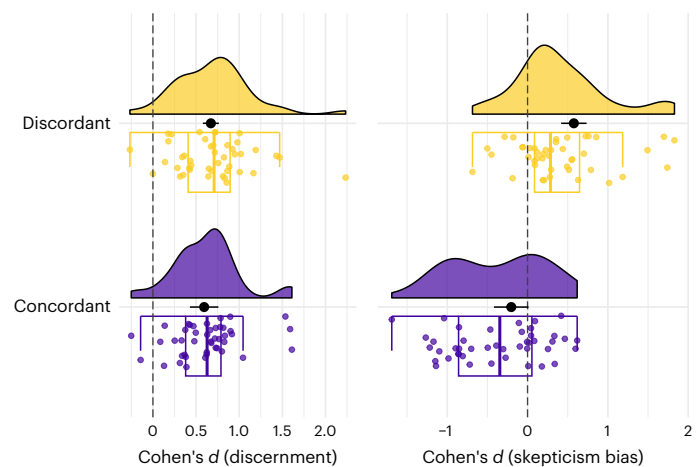

Fig. 4 | Effect of political concordance on discernment and skepticism bias. Distribution of the $n = 44$ effect sizes for politically concordant and discordant items. The black dots represent the predicted average of the meta-regression, the black horizontal bars the 95% confidence intervals. Note that the figure does not represent the different weights (that is, the varying sample sizes) of the data points, but these weights are taken into account in the meta-regression. For the boxplots, the box represents the interquartile range (IQR), that is, the distance between the first and third quartiles, the centre line indicates the median, and the outer lines (whiskers) extend to 1.5 times the IQR or the most extreme values within this range. Data beyond the end of the whiskers are considered 'outlying' points and are plotted individually.

a skepticism bias (baseline skepticism bias concordant items = −0.20 [−0.42, 0.01], $z = −1.93$, $P = 0.064$), while for discordant news items, participants displayed a positive skepticism bias (ΔSkepticism bias = 0.78 [0.62, 0.94], $z = 10.04$, $P < 0.001$). In other words, participants were not gullible when facing concordant news headlines (as would have been suggested by a negative skepticism bias), but were skeptical when facing discordant ones.

Individual-level data
In the results above, accuracy ratings were averaged across participants. It is unclear how these average results generalize to individuals. Do they hold for most participants? Or are they driven by a relatively small group of participants with either excellent discernment skills or extreme skepticism? For 22 articles ($N_{\text{Participants}} = 42,074$, $N_{\text{Observations}} = 813,517$), we have the raw data for all ratings that individual participants made on each news headline they saw. On these data, we ran a descriptive, non-pre-registered analysis: we calculated a discernment and skepticism bias score for each participant on the basis of all the news items they rated. To compare across different scales, we transposed all accuracy scores on a scale from 0 to 1, resulting in a range of possible values from −1 to 1 for both discernment and skepticism bias.

As shown in Fig. 5, 79.92% of individual participants had a positive discernment score, and 59.06% of participants had a positive skepticism bias score. Therefore, our main results based on mean ratings across participants seem to be representative of individual participants (see Supplementary Section C for further discussion).

Discussion
This meta-analysis sheds light on some of the most common fears voiced about false news. In particular, we investigated whether people are able to discern true from false news and whether they are better at judging the veracity of true news or false news (skepticism bias). Across 303 effect sizes ($N_{\text{Participants}} = 194,438$) from 40 countries across 6 continents, we found that people rated true news as much more accurate than fact-checked false news ($d_{\text{discernment}} = 1.12$ [1.01, 1.22]) and are
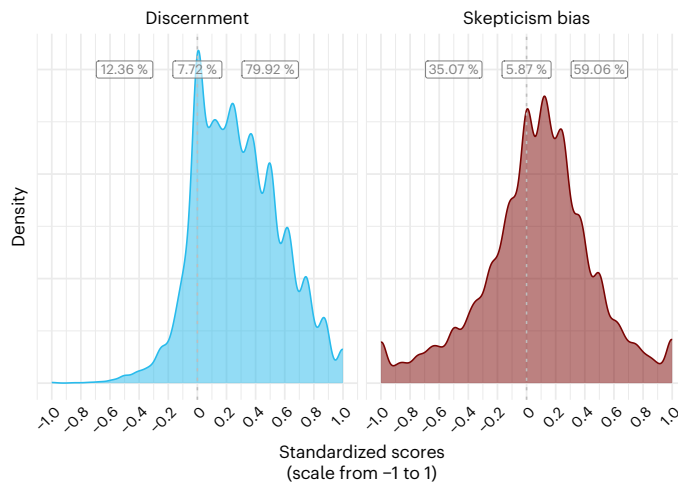
Fig. 5 | Outcomes at the participant level. Distribution of average discernment and skepticism bias scores of individual participants in the subset of studies that we have raw data on. We standardized original accuracy ratings to range from 0 to 1. The lowest possible score is −1 where, for discernment, an individual classified all news wrongly, and for skepticism bias, an individual classified all true news correctly (as true) and all false news incorrectly (as true). The highest possible score is 1 where, for discernment, an individual classified all news correctly, and for skepticism bias, an individual classified all true news incorrectly (as false) and all false news correctly (as false). The percentage labels (from left to right) represent the share of participants with a negative score, a score of exactly 0, and a positive score, for both measures.

slightly better at rating fact-checked false news as inaccurate than at rating true news as accurate ($d_{\text{skepticism bias}}$ = 0.32 [0.24, 0.39]).

The finding that people can discern true from false news when prompted to do so has important implications for interventions against misinformation. First, it suggests that most people do not lack the skills to spot false news (at least the kind of fact-checked false news used in the studies included in our meta-analysis). If people do not lack the skills to spot false news, why do they sometimes fall for false news? In some contexts, people may lack the motivation to use their discernment skills or may apply them selectively[33,34]. Thus, instead of teaching people how to spot false news, it may be more fruitful to target their motivations and incentives, either by manipulating features of the environment in which people encounter news[35,36], or by prompting people to use their critical thinking skills and pay more attention to accuracy[33]. For instance, it has been shown that design features of current social media environments sometimes impede discernment and not incentivize the sharing of accurate information[37].

Second, the fact that people can, on average, discern true from false news lends support to crowdsourced fact-checking initiatives. While fact-checkers cannot keep up with the pace of false news production, the crowd can, and it has been shown that even small groups of participants perform as well as professional fact-checkers[38,39]. The cross-cultural scope of our findings suggests that these initiatives may be fruitful in many countries worldwide. In every country included in the meta-analysis, participants on average rated true news as more accurate than false news (see Supplementary Section F). In line with past work[38], we have shown that this was not only true on average, but also for a large majority (79.92%) of participants for which we had individual-level data. Our results are also informative for the work of fact-checkers. Since people appear to be quite good at discerning true from false news, fact-checkers may want to focus on headlines that are less clearly false or true. However, we cannot rule out that people's current discernment skills stem in part from the current and past work of fact-checking organizations. Fact-checking remains important and complementary to, not in tension with, crowdsourcing efforts.

The fact that people disbelieve true news slightly more than they believe fact-checked false news speaks to the nature of the misinformation problem and how to fight it: the problem may be less that people are gullible and fall for falsehoods too easily, but instead that people are excessively skeptical and do not believe reliable information enough[15,40]. Even assuming that the rejection of true news and the acceptance of false news are of similar magnitude (and that both can be improved), given that true news are much more prevalent in people's news diet than false news[41], true news skepticism may be more detrimental to the accuracy of people's beliefs than false news acceptance[14]. This skepticism is concerning in the context of the low and declining trust and interest in news across the world[42], as well as the attacks of populist leaders on the news media[22] and growing news avoidance[43]. Interventions aimed at reducing misperceptions should therefore consider increasing the acceptance of true news in addition to reducing the acceptance of false news[14,44]. At the very least, when testing interventions, researchers should evaluate their effect on both true and false news, not just false news[45]. At best, interventions should use methods that allow estimation of discrimination while accounting for response bias, such as SDT, and make sure that apparent increases in discernment are not due to a more conservative response bias[28,46]. This is important given that recent evidence suggests that some interventions against misinformation, such as media literacy tips[47], fact-checking[48], or educational games aimed at inoculating people against misinformation[28], may reduce belief in false news at the expense of fostering skepticism towards true news.

We also investigated various moderators of discernment and skepticism bias. We found that discernment was greater in studies conducted in the United States compared with the rest of the world. This could be due to the inclusion of many countries from the Global South, where belief in misinformation and conspiracy theories has been documented to be higher[49]. In line with past work[29], the presence of a source had no statistically significant effects on discernment or skepticism bias. Neither did the topic of the news. Participants showed greater skepticism towards headlines presented in a social media format (with an image and lede) or along with an image, compared with plain headlines with just text. This suggests that the skepticism towards true news documented in this meta-analysis may be partially due to the social media format of the news headlines. Past work has shown that people report trusting news on social media less[3,19], and experimental manipulations have shown that the Facebook news format reduces belief in news[50,51], although the causal effects documented in these experiments are much smaller than the ones observed in surveys[52]. Low trust in news on social media may be a good thing, given that on average, news on social media may be less accurate than news on news websites, but it is also worrying since most of news consumption worldwide is shifting online and on social media in particular[43].

The political concordance of the news had no statistically significant effect on discernment, but participants were excessively skeptical of politically discordant news. That is, participants were equally skilled at discerning true from false news for concordant and discordant items, but rated all news as more false when it was politically discordant. This finding is in line with recent evidence on partisan biases in news judgements[53] and supports the idea that people are not excessively gullible of news they agree with, but are instead excessively skeptical of news they disagree with[15,54]. It suggests that interventions aimed at reducing partisan motivated reasoning, or at improving political reasoning in general, should focus more on increasing openness to opposing viewpoints than on increasing skepticism towards concordant viewpoints. Future studies should investigate whether the effect of congruence is specific to politics or if it holds across other topics, and compare it to a baseline of neutral items.

Our meta-analysis has two main conceptual limitations. First, participants evaluated the news stories in artificial settings that do not mimic the real world. For instance, the mere fact of asking participants

to rate the accuracy of the news stories may have increased discernment by increasing attention to accuracy[33]. When browsing on social media, people may be less discerning (and perhaps less skeptical) than in experimental settings because they would pay less attention to accuracy[37]. However, given people's low exposure to misinformation online[55], they may protect themselves from misinformation not by detecting it on the spot, but by relying on the reputation of the sources and avoiding unreliable sources[56]. Second, our results reflect choices made by researchers about news selection. The vast majority of studies in our meta-analysis relied on fact-checked false news, determined by fact-checking websites, such as Snopes or PolitiFact. By contrast, three papers[38,57,58] automated their news selection by scraping headlines from media outlets in real time. In these studies, both participants and fact-checkers (or the researchers themselves, in the case of ref. 57) rated the veracity of the headlines shortly after they got published. The three studies (53 effect sizes; 10,170 participants; all in the United States) find (1) lower discernment than our meta-analytic average and (2) a negative skepticism (that is, a credulity) bias (see Supplementary Section G for a detailed discussion). This highlights the importance of news selection in misinformation research: researchers need to think carefully about what population of news they sample from and be clear about the generalizability of their findings[40,59].

Furthermore, our meta-analysis has methodological limitations, which we address in a series of robustness checks in the Supplementary Sections. We show that our results hold across alternative effect size estimators (Supplementary Section A), and that the results are similar when running a participant-level analysis on a subset of studies for which we have raw data (Supplementary Section C) or when relying on d' (sensitivity) and c (response bias) from SDT for that subset (Supplementary Section H). A comparison of binary and Likert-scale ratings suggests that the skepticism bias stems partly from mis-classifications and partly from degrees of confidence (Supplementary Section D).

In conclusion, we found that in experimental settings, people are able to discern mainstream true news from fact-checked false news, but when they err, they tend to do so on the side of skepticism more than on the side of gullibility (although the effect is small and probably contingent on false news selection). These findings lend support to crowdsourced fact-checking initiatives and suggest that, to improve discernment, there may be more room to increase the acceptance of true news than to reduce the acceptance of false news.

Methods

Data

We undertook a systematic review and meta-analysis of the experimental literature on accuracy judgements of news, following the PRISMA guidelines[60] (Extended Data Fig. 1). All records resulting from our literature searches can be found on the OSF project page (https://osf.io/96zbp/). We documented rejection decisions for all retrieved papers (see OSF project page).

Eligibility criteria. For a publication to be included in our meta-analysis, we set six eligibility criteria: (1) We considered as relevant all document types with original data (not only published ones, but also reports, preprints and working papers). When different publications were using the same data, a common scenario, we included only one publication (which we picked arbitrarily). (2) We only included articles that measured perceived accuracy (including 'accuracy', 'credibility', 'trustworthiness', 'reliability' or 'manipulativeness') and (3) did so for both true and false news. (4) We only included studies relying on real-world news items. Accordingly, we excluded studies in which researchers made up the false news items, or manipulated the properties of the true news items. (5) We could only include articles that provided us with the relevant summary statistics (means and standard deviations for both false and true news), or publicly available data that allowed us to calculate those. In cases where we were not able to retrieve the relevant

summary statistics either way, we contacted the authors. (6) Finally, to ensure comparability, we only included studies that provided a neutral control condition. For example, ref. 61, among other things, tested the effect of an interest prime vs an accuracy prime. A neutral control condition—one that is comparable to those of other studies—would have had no prime at all. We therefore excluded this paper. Rejection decisions for all retrieved papers are documented and can be accessed on the OSF project page (https://osf.io/96zbp/). We provide a list of all included articles in Supplementary Section J.

Deviations from eligibility criteria. We followed our eligibility criteria with 4 exceptions. We rejected one paper on the basis of a criterion that we had not previously set: scale asymmetry. Reference 62 asked participants: "According to your knowledge, how do you rate the following headline?", providing a very asymmetrical set of answer options ('1–not credible; 2–somehow credible; 3–quite credible; 4–credible; 5–very credible'). The paper provides 6 effect sizes, all of which strongly favour our second hypothesis (one effect being as large as $d = 2.54$). We decided to exclude this paper from our analysis because of its very asymmetric scale (that is, there is no clear scale midpoint and the labels do not symmetrically map onto a false/true dichotomy, which contrasts with the other scales included here). Further, we stretched our criterion for real-world news on three instances: refs. 63,64 used artificial intelligence trained on real-world news to generate false news, and ref. 65 had journalists create the false news items. We reasoned that asking journalists to write news should be similar enough to real-world news, and that large language models (LLMs) already produce news headlines that are indistinguishable from real news, so it should not make a big difference.

Literature search. Our literature review is based on two systematic searches. We conducted our first search on 2 March 2023 using Scopus (search string: '"false news" OR "fake news" OR "false stor*" AND "accuracy" OR "discernment" OR "credibilit*" OR "belief" OR "susceptib*"') and Google Scholar (search string: '"Fake news"|"False news"|"False stor*"|"Accuracy"|"Discernment"|"Credibility"|"Belief"|"Suceptib*", no citations, no patents'). On Scopus, given the initially high volume of papers (12,425), we excluded papers not written in English, that were not articles or conference papers, and that were from disciplines that are probably irrelevant for the present search (for example, Dentistry, Veterinary, Chemical Engineering, Chemistry, Nursing, Pharmacology, Microbiology, Materials Science, Medicine) or unlikely to use an experimental design (for example, Computer Science, Engineering, Mathematics; see Supplementary Section I for detailed search string). After these filters were applied, we ended up with 4,002 results. The Google Scholar search was intended to identify important preprints or working papers that the Scopus search would have missed. We only considered the first 980 results of that search—a limit imposed by the 'Publish or Perish' software we used to store Google Scholar search results in a data frame.

After submitting a manuscript version, reviewers remarked that not including the terms 'misinformation' or 'disinformation' in our search string might have omitted relevant results. On 22 March 2024, we therefore conducted a second, pre-registered (https://doi.org/10.17605/OSF.IO/YN6R2, registered on 12 March 2024) search using an extended query string (search string for both Scopus and Google Scholar: ' "false news" OR "fake news" OR "false stor*" OR "misinformation" OR "disinformation") AND ("accuracy" OR "discernment" OR "credibilit*" OR "belief" OR "suceptib*" OR "reliab*" OR "vulnerabi*" '; see Supplementary Section I for detailed search string). After removing duplicates (642 between the first and the second Scopus search and 269 between the first and the second Google Scholar search), the second search yielded an additional 1,157 results for Scopus and 711 results for Google Scholar. In total, the Scopus searches yielded 5,159, and the Google Scholar searches 1,691 unique results.

We identified and removed 338 duplicates between the Google Scholar and the Scopus searches, and ended up with 6,512 documents for screening. We had two screening phases: first, titles, then abstracts. For the results from the second literature search, both authors screened the results independently. In case of conflicting decisions, an article passed onto the next stage (that is, received abstract screening or full text assessment). For the results from the second literature search, screening was done on the basis of titles and abstracts only, so that the screeners would not be influenced by information on the authors or the publishing journal. The vast majority of documents (6,248) had irrelevant titles and were removed during that phase. Most irrelevant titles were not about false news or misinformation (for example, "Formation of a tourist destination image: Co-occurrence analysis of destination promotion videos"), and some were about false news or misinformation but were not about belief or accuracy (for example, "Freedom of Expression and Misinformation Laws During the COVID-19 Pandemic and the European Court of Human Rights"). We stored the remaining 264 records in the reference management system Zotero for retrieval. Of those, we rejected a total of 217 papers that did not meet our inclusion criteria. We rejected 87 papers on the basis of their abstract and 130 after assessment of the full text. We documented all rejection decisions (available on the OSF project page, https://osf.io/96zbp/). We included the remaining 47 papers from the systematic literature search. To complement the systematic search results, we conducted forward and backward citation search through Google Scholar. We also reviewed additional studies that we had on our computers and papers we found via Twitter (mostly working papers). Taken together, we identified an additional 47 papers via those methods. Of these, we excluded 27 papers after full text assessment because they did not meet our inclusion criteria. For these papers, we also documented our exclusion decisions, which can be found together with the ones for the systematic search on the OSF project page (https://osf.io/96zbp/). We included the remaining 20 papers. In total, we included 67 papers in our meta-analysis[6,8,12,13,26,29,34,37,44,53,57–59,63–116]: 47 peer reviewed and 20 grey literature (reports and working papers). We retrieved the relevant summary statistics directly from the paper for 21 papers, calculated them ourselves on the basis of publicly available raw data for 31 papers, and got them from the authors after request for 15 papers.

Statistical methods

Unless stated otherwise, all the analyses were pre-registered 28 April 2023 (https://doi.org/10.17605/OSF.IO/SVC7U). Our choice of statistical models was informed by simulations, which can also be found on the OSF project page. We conducted all analyses in R v.4.2.2 (31 October 2022)[117] using Rstudio (v.2024.9.0.375)[118] and the tidyverse package (v.2.0.0)[119]. We relied on the functions escalc() for effect size calculations, rma.mv() for models, and robust() for clustered standard errors, all from the metafor package (v.4.6.0)[120].

Deviations from pre-registration. We pre-registered standardized mean changes using change score standardization (SMCC) as an estimator for our effect sizes[121]. However, in line with the Cochrane guidelines[32], we chose to rely on the more common Cohen's d for the main analysis. We report results from the pre-registered SMCC (along with other alternative estimators) in Supplementary Section A. All estimators yielded similar results. We did not pre-register considering scale symmetry, and proportion of true news and false news selection (taken from fact-checking sites vs verified by researchers) as moderator variables. We report the results regarding these variables in Supplementary Section B.

Outcomes. We have two complementary measures of assessing the quality of people's news judgement. The first measure is discernment. It measures the overall quality of news judgement across true and false news. We calculated discernment by subtracting the mean accuracy

ratings of false news from the mean accuracy ratings of true news, such that more positive scores indicate better discernment. However, discernment is a limited diagnostic of the quality of people's news judgement. Imagine a study A in which participants rate 50% of true news and 20% of false news as accurate and a study B rating 80% of true news and 50% of false news as accurate. In both cases, the discernment is the same: participants rated true news as more accurate by 30 percentage points than false news. However, the performance by news type is very different. In study A, people do well for false news—they only mistakenly classify 20% as accurate—but are at chance for true news. In study B, it is the opposite. We therefore used a second measure: skepticism bias. For any given level of discernment, it indicates whether people's judgements were better for true news or for false news and to what extent. First, we calculated an error for false and true news separately, which we defined as the distance of participants' actual ratings to the best possible ratings. For example, for study A, the mean error for true news is 50% (100%–50%), because in the best possible scenario, participants would have classified 100% of true news as true. The error for false news in Study A is 20% (20%–0%), because the best possible performance for participants would have been to classify 0% of false news as accurate. We calculated skepticism bias by subtracting the mean error for false news from the mean error for true news. For example, for Study A, the skepticism bias is 30% (50%–20%). A positive skepticism bias indicates that people doubt true news more than they believe false news.

Skepticism bias can only be (meaningfully) interpreted on scales using symmetrical labels, that is, the intensity of the labels to qualify true and false news are equivalent (for example, 'True' vs 'False' or 'Definitely fake' [1] to 'Definitely real' [7]). Of the effects included in the meta-analysis, 69% used scales with perfectly symmetrical labels, while 26% used imperfectly symmetrical scale labels, that is, the intensity of the labels to qualify true and false news are similar but not equivalent (for example, [1] not at all accurate, [2] not very accurate, [3] somewhat accurate, [4] very accurate; here for instance 'not at all accurate' is stronger than 'very accurate'). We could only compute symmetry for scales that explicitly labelled scale points, resulting in missing values for 5% of effects. In Supplementary Section B, we show that scale symmetry has no statistically significant effect on skepticism bias.

Effect sizes. The studies in our meta-analysis used a variety of response scales, including both binary (for example, "Do you think the above headline is accurate? - Yes, No") and continuous ones (for example "To the best of your knowledge, how accurate is the claim in the above headline? 1 = Not at all accurate, 4 = Very accurate"). To be able to compare across the different scales, we calculated standardized effects, that is, effects expressed in units of standard deviations. Precisely, we calculated Cohen's d as

$$\text{Cohen's } d = \frac{\bar{x}_{\text{true}} - \bar{x}_{\text{false}}}{\text{SD}_{\text{pooled}}} \tag{1}$$

with

$$\text{SD}_{\text{pooled}} = \sqrt{\frac{SD_{\text{true}}^2 + SD_{\text{false}}^2}{2}} \tag{2}$$

The vast majority of experiments (294 out of 303 effects) in our meta-analysis manipulated news veracity within participants, that is, having participants rate both false and true news. Following the Cochrane manual, we accounted for the dependency between ratings that this design generates when calculating the standard error for Cohen's d. Precisely, we calculated the standard error for within-participant designs as

$$\text{SE}_{\text{Cohen's } d(\text{within})} = \sqrt{\frac{2(1 - r_{\text{true,false}})}{n} + \frac{\text{Cohen's } d^2}{2n}} \tag{3}$$

where *r* is the correlation between true and false news. Ideally, for each effect size (that is, the meta-analytic units of observation) in our data, we need the estimate of *r*. However, this correlation is generally not reported in the original papers. We could only obtain it for a subset of samples for which we computed the summary statistics ourselves based on the raw data. On the basis of this subset of correlations, we calculated an average correlation, which we then imputed for all effect size calculations. This approach is in line with the Cochrane recommendations for crossover trials[32]. In our case, this average correlation was 0.26.

For the 9 (out of 303) effects from studies that used a between-participant design, we calculated the standard error as

$$SE_{Cohen's\,d(between)} = \sqrt{\frac{n_{true} + n_{false}}{n_{true}n_{false}} + \frac{Cohen's\,d^2}{2(n_{true} + n_{false})}} \quad (4)$$

For all effect size calculations, we defined the sample size *n* as the number of instances of news ratings. That is, we multiplied the number of participants by the number of news items rated per participant.

Models. In our models for the meta-analysis, each effect size was weighted by the inverse of its standard error, thereby giving more weight to studies with larger sample sizes. We used random effects models, which assume that there is not only one true effect size but a distribution of true effect sizes[122]. These models assume that variation in effect sizes is not due to sampling error alone, and thereby allow modelling other sources of variance. We estimated the overall effect of our outcome variables using a three-level meta-analytic model with random effects at the sample and the publication level. This approach allowed us to account for the hierarchical structure of our data, in which samples (level three) contributed multiple effects (level two), level one being the participant level of the original studies (see ref. 122). A common case where a sample provided several effect sizes occurred when participants rated both politically concordant and discordant news. In this case, if possible, we entered summary statistics separately for the concordant and discordant items, yielding two effect sizes (that is, two different rows in our data frame). Another case where multiple effects per sample occurred was when follow-up studies were conducted on the same participants (but different news items). While our multilevel models account for this hierarchical structure of the data, they do not account for dependencies in sampling error. When one same sample contributes several effect sizes, one should expect their respective sampling errors to be correlated[122]. To account for dependency in sampling errors, we computed cluster-robust standard errors, confidence intervals and statistical tests for all meta-analytic estimates.

To assess the effect of moderator variables, we used meta-regressions. We calculated a separate regression for each moderator, by adding the moderator variable as a fixed effect to the multilevel meta-models presented above. We pre-registered a list of six moderator variables to test. These included the 'country' of participants (levels: United States vs all other countries), 'political concordance' (levels: politically concordant vs politically discordant), 'news family' (levels: political, including both concordant and discordant vs covid related vs other, including categories as diverse as history, environment, health, science and military-related news items), the 'format' in which the news were presented (levels: headline only vs headline and picture vs headline, picture and lede), whether news items were accompanied by a 'source' or not, and the 'response scale' used (levels: 4-point vs binary vs 6-point vs 7-point vs other, for all other numeric scales that were not frequent). We ran an additional regression for two non-pre-registered variables, namely, the 'symmetry of scales' (levels: perfectly symmetrical vs imperfectly symmetrical) and 'false news selection' (levels: taken from fact-checking sites vs verified by researchers). We further descriptively checked whether the 'proportion of true news' among all news would yield differences.

Publication bias. We ran standard procedures for detecting publication bias. However, we did not a priori expect publication bias to be present because our variables of interest were not those of interest to the researchers of the original studies: researchers generally set out to test factors that alter discernment, and not the state of discernment in the control group. No study measured skepticism bias in the way we define it here.

Regarding discernment, we found evidence that smaller studies tended to report larger effect sizes, according to Egger's regression test (Extended Data Fig. 2; see also Supplementary Section E). We did not find evidence for asymmetry regarding skepticism bias. However, it is unclear how meaningful these results are. As illustrated by the funnel plot, there is generally high between-effect size heterogeneity: even when focusing only on the most precise effect sizes (top of the funnel), the estimates vary substantially. It thus seems reasonable to assume that most of the dispersion of effect sizes does not arise from the studies' sampling error, but from the studies estimating different true effects. Further, even the small studies are relatively high powered, suggesting that they would have yielded significant, publishable results even with smaller effect sizes. Lastly, Egger's regression test can lead to an inflation of false positive results when applied to standardized mean differences[122,123].

We did not find any evidence to suspect *P*-hacking for either discernment or skepticism bias from visually inspecting *P*-curves for both outcomes (Extended Data Fig. 3).

Reporting summary
Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability
The extracted data used to produce our results are available on OSF at https://doi.org/10.17605/OSF.IO/96ZBP (ref. 124).

Code availability
The code used to create all results (including tables and figures) of this manuscript is also available on OSF at https://doi.org/10.17605/OSF.IO/96ZBP (ref. 124).

References
1. Egelhofer, J. L., Boyer, M., Lecheler, S. & Aaldering, L. Populist attitudes and politicians' disinformation accusations: effects on perceptions of media and politicians. *J. Commun.* **72**, 619–632 (2022).
2. Brennen, T. & Magnussen, S. Lie detection: what works? *Curr. Dir. Psychol. Sci.* **32**, 395–401 (2023).
3. Mont'Alverne, C. et al. The trust gap: how and why news on digital platforms is viewed more sceptically versus news in general. *Reuters Institute* https://doi.org/10.60625/risj-skfk-h856 (2022).
4. Metzger, M. J. Making sense of credibility on the Web: models for evaluating online information and recommendations for future research. *J. Am. Soc. Inf. Sci. Technol.* **58**, 2078–2091 (2007).
5. Ross Arguedas, A. et al. Snap judgements: how audiences who lack trust in news navigate information on digital platforms. *Reuters Institute* https://doi.org/10.60625/risj-64ja-0s18 (2022).
6. Altay, S., Lyons, B. & Modirrousta-Galian, A. Exposure to higher rates of false news erodes media trust and fuels skepticism in news judgment. *Mass Commun. Soc.* https://doi.org/10.1080/15205436.2024.2382776 (2024).
7. Bryanov, K. & Vziatysheva, V. Determinants of individuals' belief in fake news: a scoping review determinants of belief in fake news. *PLoS ONE* **16**, e0253717 (2021).
8. Pennycook, G. & Rand, D. G. Lazy, not biased: susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition* **188**, 39–50 (2019).

9. Schulz, A., Fletcher, R. & Popescu, M. Are news outlets viewed in the same way by experts and the public? A comparison across 23 European countries. *Reuters Institute* https://doi.org/10.60625/risj-nbg4-8b23 (2020).

10. Mourão, R. R. & Robertson, C. T. Fake news as discursive integration: an analysis of sites that publish false, misleading, hyperpartisan and sensational information. *Journal. Stud.* **20**, 2077–2095 (2019).

11. Vosoughi, S., Roy, D. & Aral, S. The spread of true and false news online. *Science* **359**, 1146–1151 (2018).

12. Chen, X., Pennycook, G. & Rand, D. What makes news sharable on social media? *J. Quant. Descr. Digit. Media* https://doi.org/10.51685/jqd.2023.007 (2023).

13. Altay, S., de Araujo, E. & Mercier, H. "If this account is true, it is most enormously wonderful": interestingness-if-true and the sharing of true and false news. *Digit. Journal.* **10**, 373–394 (2022).

14. Acerbi, A., Altay, S. & Mercier, H. Research note: fighting misinformation or fighting for information? *Harv. Kennedy Sch. Misinformation Rev.* https://doi.org/10.37016/mr-2020-87 (2022).

15. Mercier, H. *Not Born Yesterday: The Science of Who We Trust and What We Believe* (Princeton Univ. Press, 2020).

16. Brashier, N. M. & Marsh, E. J. Judging truth. *Annu. Rev. Psychol.* **71**, 499–515 (2020).

17. Street, C. N. H. & Masip, J. The source of the truth bias: heuristic processing? *Scand. J. Psychol.* **56**, 254–263 (2015).

18. Levine, T. R. Truth-Default Theory (TDT): a theory of human deception and deception detection. *J. Lang. Soc. Psychol.* **33**, 378–392 (2014).

19. Newman, N., Fletcher, R., Robertson, C. T., Eddy, K. & Nielsen, R. K. Digital News Report 2022. *Reuters Institute* https://doi.org/10.60625/risj-x1gn-m549 (2022).

20. Mihailidis, P. & Foster, B. The cost of disbelief: fracturing news ecosystems in an age of rampant media cynicism. *Am. Behav. Sci.* **65**, 616–631 (2021).

21. Egelhofer, J. L. & Lecheler, S. Fake news as a two-dimensional phenomenon: a framework and research agenda. *Ann. Int. Commun. Assoc.* **43**, 97–116 (2019).

22. Van Duyn, E. & Collier, J. Priming and fake news: the effects of elite discourse on evaluations of news media. *Mass Commun. Soc.* **22**, 29–48 (2019).

23. Paul, C. & Matthews, M. *The Russian "Firehose of Falsehood" Propaganda Model* (Rand, 2016).

24. Ulusoy, E. et al. Flooding the zone: how exposure to implausible statements shapes subsequent belief judgments. *Int. J. Public Opin. Res.* **33**, 856–872 (2021).

25. Fletcher, R. & Nielsen, R.-K. People dont trust news media-and this is key to the global misinformation debate. In *Understanding and Addressing the Disinformation Ecosystem* 13–17 (Annenberg School for Communication, 2017).

26. Luo, M., Hancock, J. T. & Markowitz, D. M. Credibility perceptions and detection accuracy of fake news headlines on social media: effects of truth-bias and endorsement cues. *Commun. Res.* **49**, 171–195 (2022).

27. Batailler, C., Brannon, S. M., Teas, P. E. & Gawronski, B. A signal detection approach to understanding the identification of fake news. *Perspect. Psychol. Sci.* **17**, 78–98 (2022).

28. Modirrousta-Galian, A. & Higham, P. A. Gamified inoculation interventions do not improve discrimination between true and fake news: reanalyzing existing research with receiver operating characteristic analysis. *J. Exp. Psychol. Gen.* **152**, 2411 (2023).

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

30. Tappin, B. M., Pennycook, G. & Rand, D. G. Bayesian or biased? Analytic thinking and political belief updating. *Cognition* **204**, 104375 (2020).

31. Guay, B., Berinsky, A. J., Pennycook, G. & Rand, D. How to think about whether misinformation interventions work. *Nat. Hum. Behav.* **7**, 1231–1233 (2023).

32. Higgins, J. P. et al. *Cochrane Handbook for Systematic Reviews of Interventions* (John Wiley & Sons, 2019).

33. Pennycook, G. et al. Shifting attention to accuracy can reduce misinformation online. *Nature* **592**, 590–595 (2021).

34. Rathje, S., Roozenbeek, J., Van Bavel, J. J. & van der Linden, S. Accuracy and social motivations shape judgements of (mis) information. *Nat. Hum. Behav.* **7**, 892–903 (2023).

35. Capraro, V. & Celadin, T. "I think this news is accurate": endorsing accuracy decreases the sharing of fake news and increases the sharing of real news. *Pers. Soc. Psychol. Bull.* **49**, 1635–1645 (2023).

36. Globig, L. K., Holtz, N. & Sharot, T. Changing the incentive structure of social media platforms to halt the spread of misinformation. *eLife* **12**, e85767 (2023).

37. Epstein, Z., Sirlin, N., Arechar, A., Pennycook, G. & Rand, D. The social media context interferes with truth discernment. *Sci. Adv.* **9**, eabo6169 (2023).

38. Allen, J., Arechar, A. A., Pennycook, G. & Rand, D. G. Scaling up fact-checking using the wisdom of crowds. *Sci. Adv.* **7**, eabf4393 (2021).

39. Martel, C., Allen, J. N. L., Pennycook, G. & Rand, D. G. Crowds can effectively identify misinformation at scale. *OSF* https://doi.org/10.31234/osf.io/2tjk7 (2022).

40. Altay, S., Berriche, M. & Acerbi, A. Misinformation on misinformation: conceptual and methodological challenges. *Soc. Media Soc.* https://doi.org/10.1177/20563051221150412 (2023).

41. Allen, J., Howland, B., Mobius, M., Rothschild, D. & Watts, D. J. Evaluating the fake news problem at the scale of the information ecosystem. *Sci. Adv.* **6**, eaay3539 (2020).

42. Altay, S., Fletcher, R. & Nielsen, R. K. News participation is declining: evidence from 46 countries between 2015 and 2022. *New Media Soc.* https://doi.org/10.1177/14614448241247822 (2024).

43. Newman, N., Fletcher, R., Eddy, K., Robertson, C. T. & Nielsen, R. K. Digital news report 2023. *Reuters Institute* https://doi.org/10.60625/risj-p6es-hb13 (2023).

44. Altay, S., De Angelis, A. & Hoes, E. Media literacy tips promoting reliable news improve discernment and enhance trust in traditional media. *Commun. Psychol.* **2**, 74 (2024).

45. Guay, B., Berinsky, A., Pennycook, G. & Rand, D. How to think about whether misinformation interventions work. *Nat. Hum. Behav.* **7**, 1231–1233 (2023).

46. Higham, P. A., Modirrousta-Galian, A. & Seabrooke, T. Mean rating difference scores are poor measures of discernment: the role of response criteria. *Curr. Opin. Psychol.* **56**, 101785 (2024).

47. Hoes, E., Aitken, B., Zhang, J., Gackowski, T. & Wojcieszak, M. *Prominent misinformation interventions reduce misperceptions but increase skepticism. Nat. Hum. Behav.* **8**, 1545–1553 (2024).

48. Bachmann, I. & Valenzuela, S. Studying the downstream effects of fact-checking on social media: experiments on correction formats, belief accuracy, and media trust. *Soc. Media Soc.* https://doi.org/10.1177/20563051231179694 (2023).

49. Alper, S. When conspiracy theories make sense: the role of social inclusiveness. *Gale OneFile: Health and Medicine* https://link.gale.com/apps/doc/A664289697/HRCA?u=anon~2d0546d6&sid=sitemap&xid=e90dddc5 (2021).

50. Besalú, R. & Pont-Sorribes, C. Credibility of digital political news in Spain: comparison between traditional media and social media. *Soc. Sci.* **10**, 170 (2021).

51. Karlsen, R. & Aalberg, T. Social media and trust in news: an experimental study of the effect of Facebook on news story credibility. *Digit. Journal.* **11**, 144–160 (2023).

52. Agadjanian, A. et al. A platform penalty for news? How social media context can alter information credibility online. *J. Inf. Technol. Politics* **20**, 338–348 (2023).

53. Gawronski, B., Ng, N. L. & Luke, D. M. Truth sensitivity and partisan bias in responses to misinformation. *J. Exp. Psychol. Gen.* **152**, 2205–2236 (2023).

54. Trouche, E., Johansson, P., Hall, L. & Mercier, H. Vigilant conservatism in evaluating communicated information. *PLoS ONE* **13**, e0188825 (2018).

55. Altay, S., Kleis Nielsen, R. & Fletcher, R. Quantifying the "infodemic": People turned to trustworthy news outlets during the 2020 coronavirus pandemic. *J. Quant. Descr. Digit. Media* https://doi.org/10.51685/jqd.2022.020 (2022).

56. Altay, S., Hacquin, A.-S. & Mercier, H. Why do so few people share fake news? It hurts their reputation. *New Media Soc.* **24**, 1303–1324 (2022).

57. Garrett, R. K. & Bond, R. M. Conservatives' susceptibility to political misperceptions. *Sci. Adv.* **7**, eabf1234 (2021).

58. Aslett, K. et al. Online searches to evaluate misinformation can increase its perceived veracity. *Nature* **625**, 548–556 (2024).

59. Pennycook, G., Binnendyk, J., Newton, C. & Rand, D. G. A practical guide to doing behavioral research on fake news and misinformation. *Collabra Psychol.* **7**, 25293 (2021).

60. Page, M. J. et al. THE PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372**, n71 (2021).

61. Calvillo, D. P. & Smelter, T. J. An initial accuracy focus reduces the effect of prior exposure on perceived accuracy of news headlines. *Cogn. Res. Princ. Implic.* **5**, 55 (2020).

62. Baptista, J. P., Correia, E., Gradim, A. & Piñeiro-Naval, V. The influence of political ideology on fake news belief: the Portuguese case. *Publications* **9**, 23 (2021).

63. Maertens, R. et al. The Misinformation Susceptibility Test (MIST): a psychometrically validated measure of news veracity discernment. *Behav. Res. Methods* **56**, 1863–1899 (2024).

64. Roozenbeek, J. et al. Susceptibility to misinformation about COVID-19 around the world. *R. Soc. Open Sci.* **7**, 201199 (2020).

65. Bryanov, K. et al. What drives perceptions of foreign news coverage credibility? A cross-national experiment including Kazakhstan, Russia, and Ukraine. *Polit. Commun.* **40**, 115–146 (2023).

66. Ali, A. & Qazi, I. A. Digital literacy and vulnerability to misinformation: evidence from Facebook users in Pakistan. *J. Quant. Descr. Digit. Media* https://doi.org/10.51685/jqd.2022.025 (2022).

67. Allen, J., Arechar, A. A., Pennycook, G. & Rand, D. G. Scaling up fact-checking using the wisdom of crowds. *Sci. Adv.* **7**, eabf4393 (2021). (2021).

68. Altay, S. & Gilardi, F. People are skeptical of headlines labeled as ai-generated, even if true or human-made, because they assume full AI automation. *OSF* https://doi.org/10.31234/osf.io/83k9r (2023).

69. Altay, S., Nielsen, R. K. & Fletcher, R. The impact of news media and digital platform use on awareness of and belief in COVID-19 misinformation. Preprint at *PsyArXiv* https://doi.org/10.31234/osf.io/7tm3s (2022).

70. Arechar, A. A. et al. Understanding and combatting misinformation across 16 countries on six continents. *Nat. Hum. Behav.* **7**, 1502–1513 (2023).

71. Badrinathan, S. Educative Interventions to combat misinformation: evidence from a field experiment in India. *Am. Polit. Sci. Rev.* **115**, 1325–1341 (2021).

72. Bago, B., Rand, D. G. & Pennycook, G. Fake news, fast and slow: deliberation reduces belief in false (but not true) news headlines. *J. Exp. Psychol. Gen.* **149**, 1608–1613 (2020).

73. Bago, B., Rosenzweig, L. R., Berinsky, A. J. & Rand, D. G. Emotion may predict susceptibility to fake news but emotion regulation does not seem to help. *Cogn. Emot.* **36**, 1166–1180 (2022).

74. Basol, M. et al. Towards psychological herd immunity: cross-cultural evidence for two prebunking interventions against COVID-19 misinformation. *Big Data Soc.* **8**, 205395172110138 (2021).

75. Brashier, N. M., Pennycook, G., Berinsky, A. J. & Rand, D. G. Timing matters when correcting fake news. *Proc. Natl. Acad. Sci. USA* **118**, e2020043118 (2021).

76. Bronstein, M. V., Pennycook, G., Bear, A., Rand, D. G. & Cannon, T. D. Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking. *J. Appl. Res. Mem. Cogn.* **8**, 108–117 (2019).

77. Clayton, K. et al. Real solutions for fake news? Measuring the effectiveness of general warnings and fact-check tags in reducing belief in false stories on social media. *Polit. Behav.* **42**, 1073–1095 (2020).

78. Clemm von Hohenberg, B. Truth and bias, left and right: testing ideological asymmetries with a realistic news supply. *Public Opin. Q.* **87**, 267–292 (2023).

79. Erlich, A. & Garner, C. Is pro-Kremlin disinformation effective? Evidence from Ukraine. *Int. J. Press Polit.* **28**, 5–28 (2023).

80. Espina Mairal, S., Bustos, F., Solovey, G. & Navajas, J. Interactive crowdsourcing to fact-check politicians. *J. Exp. Psychol. Appl.* **30**, 3–15 (2024).

81. Lee, E. J. & Jang, J. How political identity and misinformation priming affect truth judgments and sharing intention of partisan news. *Digit. Journal.* **11**, 226–245 (2023).

82. Faragó, L., Krekó, P. & Orosz, G. Hungarian, lazy, and biased: the role of analytic thinking and partisanship in fake news discernment on a Hungarian representative sample. *Sci. Rep.* **13**, 178 (2023).

83. Fazio, L. et al. Combating misinformation: a megastudy of nine interventions designed to reduce the sharing of and belief in false and misleading headlines. *OSF* https://doi.org/10.31234/osf.io/uyjha (2024).

84. Gottlieb, J., Adida, C. & Moussa, R. Reducing misinformation in a polarized context: experimental evidence from Côte d'Ivoire. *OSF Preprints* https://doi.org/10.31219/osf.io/6x4wy (2022).

85. Guess, A. M. et al. A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proc. Natl. Acad. Sci. USA* **117**, 15536–15545 (2020).

86. Guess, A., McGregor, S., Pennycook, G. & Rand, D. Unbundling digital media literacy tips: results from two experiments. *OSF* https://doi.org/10.31234/osf.io/u34fp (2024).

87. Hameleers, M. et al. Mistakenly misinformed or intentionally deceived? Mis- and disinformation perceptions on the Russian War in Ukraine among citizens in 19 countries. *Eur. J. Polit. Res.* https://doi.org/10.1111/1475-6765.12646 (2023).

88. Hlatky, R. Unintended consequences? Russian disinformation and public opinion. *OSF* https://doi.org/10.31219/osf.io/85vmt (2024).

89. Koetke, J., Schumann, K., Porter, T. & Smilo-Morgan, I. Fallibility salience increases intellectual humility: implications for people's willingness to investigate political misinformation. *Pers. Soc. Psychol. Bull.* **49**, 806–820 (2023).

90. Kreps, S. E. & Kriner, D. L. Assessing misinformation recall and accuracy perceptions: evidence from the COVID-19 pandemic. *Harv. Kennedy Sch. Misinformation Rev.* https://doi.org/10.37016/mr-2020-123 (2023).

91. Lutzke, L., Drummond, C., Slovic, P. & Árvai, J. Priming critical thinking: simple interventions limit the influence of fake news about climate change on Facebook. *Glob. Environ. Change* **58**, 101964 (2019).

92. Lyons, B., King, A. J. & Kaphingst, K. A health media literacy intervention increases skepticism of both inaccurate and accurate cancer news among U.S. adults. *OSF Preprints* https://doi.org/10.31219/osf.io/hm9ty (2024).

93. Lyons, B., Modirrousta-Galian, A., Altay, S. & Salovich, N. A. Reduce blind spots to improve news discernment? Performance feedback reduces overconfidence but does not improve subsequent discernment. *OSF Preprints* https://doi.org/10.31219/osf.io/kgfrb (2024).

94. Lyons, B., Montgomery, J. & Reifler, J. Partisanship and older Americans' engagement with dubious political news. *OSF* https://doi.org/10.31219/osf.io/etb89 (2023).

95. Lühring, J. et al. Emotions in misinformation studies: distinguishing affective state from emotional response and misinformation recognition from acceptance. Preprint at *PsyArXiv* https://doi.org/10.31234/osf.io/udqms (2023).

96. Martel, C., Pennycook, G. & Rand, D. G. Reliance on emotion promotes belief in fake news. *Cogn Res Princ Implic* **5**, 47 (2020).

97. Modirrousta-Galian, A., Higham, P. A. & Seabrooke, T. Effects of inductive learning and gamification on news veracity discernment. *J. Exp. Psychol. Appl.* **29**, 599–619 (2023).

98. Modirrousta-Galian, A., Higham, P. A. & Seabrooke, T. Wordless wisdom: the dominant role of tacit knowledge in true and fake news discrimination. *J. Appl. Res. Mem. Cogn.* https://doi.org/10.1037/mac0000151 (2024).

99. Muda, R., Pennycook, G., Hamerski, D. & Białek, M. People are worse at detecting fake news in their foreign language. *J. Exp. Psychol. Appl.* **29**, 712–724 (2023).

100. OECD. An international effort using behavioural science to tackle the spread of misinformation. *OECD Public Governance Policy Papers 21; OECD Public Governance Policy Papers, Vol. 21* https://doi.org/10.1787/b7709d4f-en (2022).

101. Orosz, G., Paskuj, B., Faragó, L. & Krekó, P. A prosocial fake news intervention with durable effects. *Sci. Rep.* **13**, 3958 (2023).

102. Pehlivanoglu, D. et al. The role of analytical reasoning and source credibility on the evaluation of real and fake full-length news articles. *Cogn. Res. Princ. Implic.* **6**, 24 (2021).

103. Pennycook, G. & Rand, D. G. Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *J. Pers.* **88**, 185–200 (2020).

104. Pennycook, G., Cannon, T. D. & Rand, D. G. Prior exposure increases perceived accuracy of fake news. *J. Exp. Psychol. Gen.* **147**, 1865–1880 (2018). (2018).

105. Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G. & Rand, D. G. Fighting COVID-19 misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention. *Psychol. Sci.* **31**, 770–780 (2020).

106. Pereira, F. B., Bueno, N. S., Nunes, F. & Pavão, N. Inoculation reduces misinformation: experimental evidence from multidimensional interventions in Brazil. *J. Exp. Polit. Sci.* **11**, 239–250 (2024). (2023).

107. Peren Arin, K., Mazrekaj, D. & Thum, M. Ability of detecting and willingness to share fake news. *Sci. Rep.* **13**, 7298 (2023).

108. Roozenbeek, J. et al. Susceptibility to misinformation is consistent across question framings and response modes and better explained by myside bias and partisanship than analytical thinking. *Judgm. Decis. Mak.* **17**, 27 (2022).

109. Rosenzweig, L. R., Bago, B., Berinsky, A. J. & Rand, D. G. Happiness and surprise are associated with worse truth discernment of COVID-19 headlines among social media users in Nigeria. *Harv. Kennedy Sch. Misinformation Rev.* https://doi.org/10.37016/mr-2020-75 (2021).

110. Ross, B., Heisel, J., Jung, A.-K. & Stieglitz, S. Fake news on social media: the (in)effectiveness of warning messages. In *Proc. 39th International Conference on Information Systems* Article 16 (Association for Information Systems, 2018).

111. Ross, R. M., Rand, D. G. & Pennycook, G. Beyond "fake news": analytic thinking and the detection of false and hyperpartisan news headlines. *Judgm. Decis. Mak.* **16**, 22 (2021).

112. Shirikov, A. Fake news for all: how citizens discern disinformation in autocracies. *Polit. Commun.* **41**, 45–65 (2024).

113. Smelter, T. J. & Calvillo, D. P. Pictures and repeated exposure increase perceived accuracy of news headlines. *Appl. Cogn. Psychol.* **34**, 1061–1071 (2020).

114. Stagnaro, M., Pink, S., Rand, D. G. & Willer, R. Increasing accuracy motivations using moral reframing does not reduce Republicans' belief in false news. *Harv. Kennedy Sch. Misinformation Rev*. 10.37016/mr-2020–128 (2023).

115. Sultan, M. et al. Time pressure reduces misinformation discrimination ability but does not alter response bias. *Sci. Rep.* **12**, 22416 (2022).

116. Winter, S. et al. (Don't) stop believing: a signal detection approach to risk and protective factors for engagement with politicized (mis)information in social media. Preprint at *PsyArXiv* https://doi.org/10.31234/osf.io/84c36 (2024).

117. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2022).

118. Posit Team. *RStudio: Integrated Development Environment for R* (PBC, 2023).

119. Wickham, H. et al. Welcome to the tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).

120. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* **36**, 1–48 (2010).

121. Gibbons, R. D., Hedeker, D. R. & Davis, J. M. Estimation of effect size from a series of experiments involving paired comparisons. *J. Educ. Stat.* **18**, 271–279 (1993).

122. Harrer, M., Cuijpers, P., Furukawa, T. A. & Ebert, D. D. *Doing Meta-analysis With R: A Hands-on Guide* (CRC Press, 2021).

123. Pustejovsky, J. E. *Simulating Correlated Standardized Mean Differences for Meta-analysis* https://jepusto.com/posts/Simulating-correlated-SMDs/ (2019).

124. Pfänder, J. & Altay, S. News accuracy meta analysis. *OSF* https://doi.org/10.17605/OSF.IO/96ZBP (2025).

Acknowledgements

Author contributions

J.P. conceptualized the project, performed systematic literature search, developed methodology, acquired software, conducted formal analysis, data curation and visualization, wrote the original draft, and reviewed and edited the manuscript. S.A. conceptualized the project, performed systematic literature search, wrote the original draft, and reviewed and edited the manuscript.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41562-024-02086-1.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41562-024-02086-1.

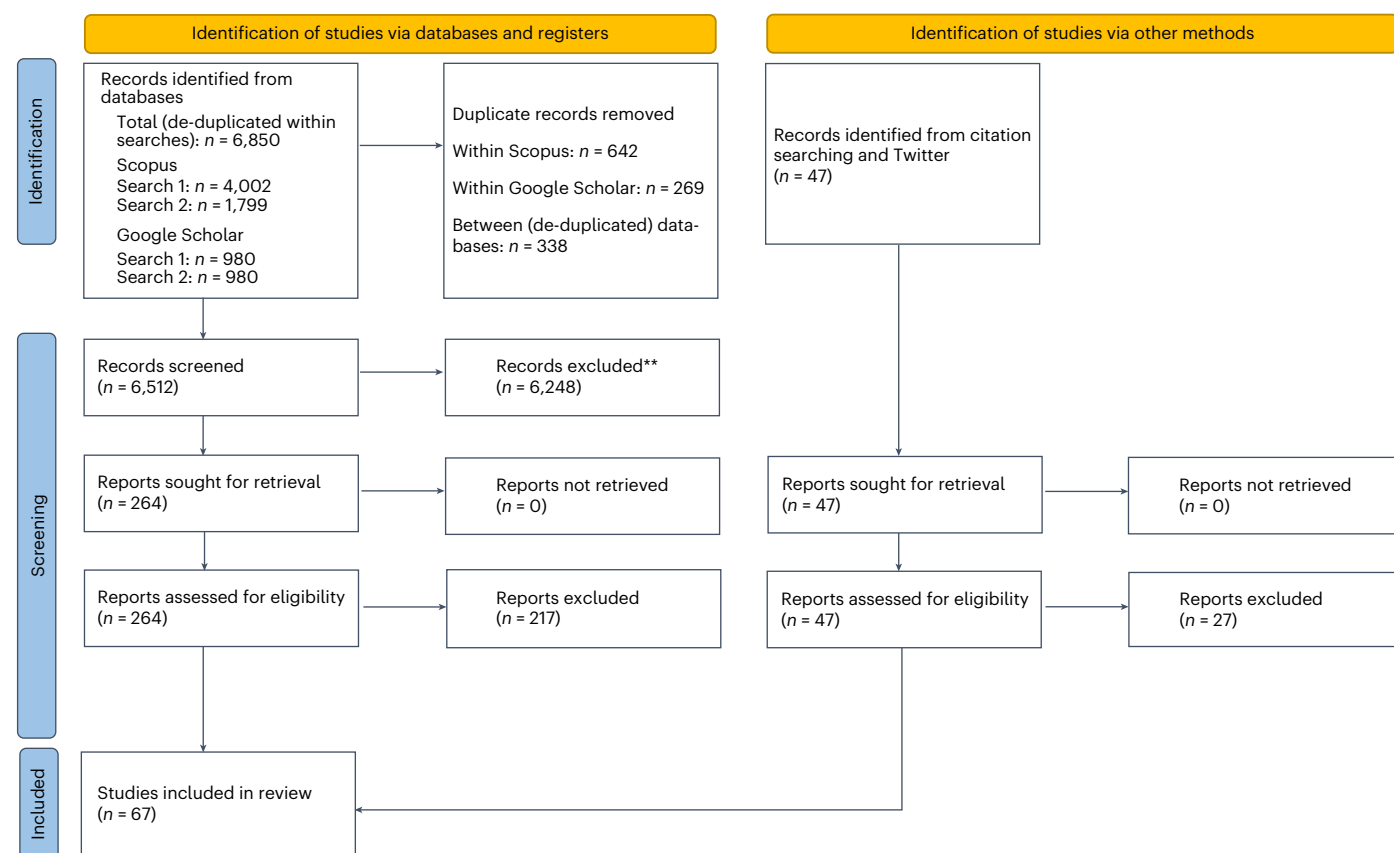**Correspondence and requests for materials** should be addressed to Sacha Altay.

Peer review information *Nature Human Behaviour* thanks Chang Lu, Jonathan van 't Riet and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Identification of studies via databases and registers

Identification of studies via other methods

Identification

Records identified from databases
Total (de-duplicated within searches): n = 6,850
Scopus
Search 1: n = 4,002
Search 2: n = 1,799
Google Scholar
Search 1: n = 980
Search 2: n = 980

Duplicate records removed
Within Scopus: n = 642
Within Google Scholar: n = 269
Between (de-duplicated) databases: n = 338

Records identified from citation searching and Twitter (n = 47)

Screening

Records screened
(n = 6,512)

Records excluded**
(n = 6,248)

Reports sought for retrieval
(n = 264)

Reports not retrieved
(n = 0)

Reports sought for retrieval
(n = 47)

Reports not retrieved
(n = 0)

Reports assessed for eligibility
(n = 264)

Reports excluded
(n = 217)

Reports assessed for eligibility
(n = 47)

Reports excluded
(n = 27)

Included

Studies included in review
(n = 67)

Extended Data Fig. 1 | PRISMA flow diagram. A flow diagram for the systematic literature review, based on the 2020 PRISMA template.
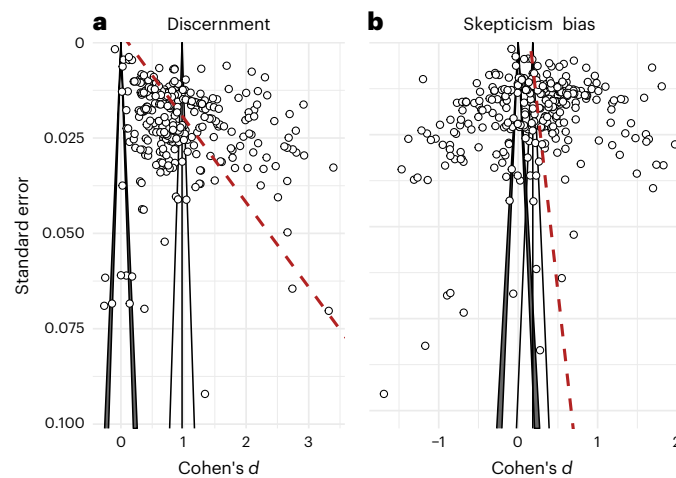
Extended Data Fig. 2 | Funnel plots for discernment and skepticism bias. Dots represent effect sizes for both discernment (**a**) and skepticism bias (**b**). In the absence of publication bias and heterogeneity, one would then expect to see the points forming a funnel shape, with the majority of the points falling inside of the pseudo-confidence region centred around the average effect (solid vertical line) estimate, with contours corresponding to the 95% confidence interval, i.e. ±1.96 s.e. (the standard error value from the *y* axis). The second funnel represents the expected shape under the Null-hypothesis, with the shaded area marking statistical significance between the 5% and the 1% level, and everything outside the contours corresponding to the 1% level. The dashed red regression line illustrates the estimate of the Egger's regression test. For discernment (but not for skepticism bias), the slope differs significantly from zero (Supplementary Section E).
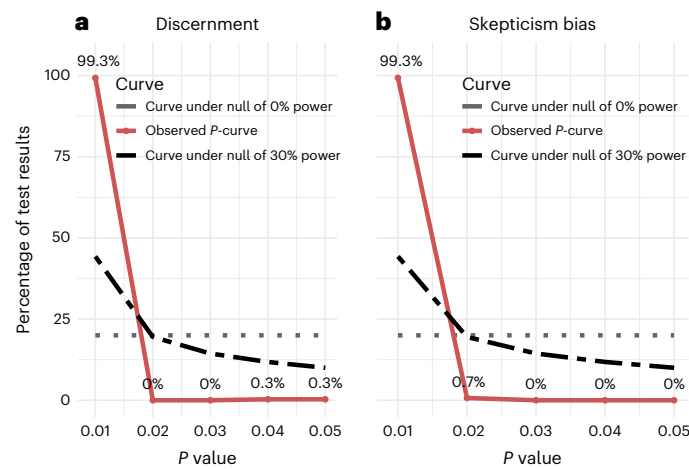
Extended Data Fig. 3 | *P*-curves for discernment and skepticism bias. a,b, The *P*-curves show the percentage of effect sizes for a given *P* value within the range 0.01–0.05. All values smaller than 0.01 are rounded off to that value. The dashed/dotted reference lines indicate the expected percentage of studies for a given *P* value, assuming that there is a true effect and certain statistical power to detect it (either 0%, grey dotted line, or 30%, black dashed line, power). The observed *P*-curve is negatively sloped and heavily right skewed (the tail points to the right) for both outcomes, which suggests no evidence of *P*-hacking.

nature portfolio

Corresponding author(s):   Sacha Altay

Last updated by author(s):   Oct 21, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	We used R version 4.2.2 (2022-10-31), Rstudio version 2024.9.0.375, and the tidyverse package version 2.0.068 for extracting data from papers.The extracted data used to produce our results are available on the OSF project page (https://osf.io/96zbp/).
Data analysis	We conducted all analyses in R version 4.2.2 (2022-10-31) using Rstudio version 2024.9.0.375 and the tidyverse package version 2.0.068. For effect size calculations, we rely on the escalc(), for models on therma.mv(), for clustered standard errors on the robust() function, all from the metafor package version 4.6.069. The code used to create all results (including tables and figures) of this manuscript is also available on the OSF project page (https://osf.io/96zbp/).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

The extracted data used to produce our results are available on the OSF project page (https://osf.io/96zbp/). For our literature search, we relied on the Scopus database (https://www.scopus.com/search/form.uri?display=basic#basic).

Research involving human participants, their data, or biological material

Reporting on sex and gender	No reporting on sex or gender.
Reporting on race, ethnicity, or other socially relevant groupings	No reporting on race, ethnicity, or other socially relevant groupings.
Population characteristics	No population covariates were measured, except for country.
Recruitment	Most studies included in our meta-analysis recruited participants online, some of which used representative samples.
Ethics oversight	None.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We conducted a meta-analysis, i.e. took a purely quantitative approach in our systematic review.
Research sample	We selected all our studies systematically, following the PRISMA guidelines. Most samples of included studies were recruited online, via platforms such as Amazon MTurk or Prolific. Some studies recruited representative samples of the respective country population.
Sampling strategy	We selected all our studies systematically, following the PRISMA guidelines. Our inclusion criteria are documented in the article.
Data collection	We retrieved the relevant summary statistics directly from the paper for 21 papers, calculated them ourselves based on publicly available raw data for 31 papers, and got them from the authors after request for 15 papers. For the second systematic search, both authors screened search results independently. In case of conflicting decisions, an article passed onto the next stage (i.e. receive abstract screening or full text assessment).
Timing	We conducted our first search on March 2, 2023. After submitting a manuscript version, reviewers remarked that not including the terms "misinformation" or "disinformation" in our search string might have omitted relevant results. On March 22nd, 2024, we therefor conducted a second, pre-registered (https://osf.io/yn6r2) search using an extended query string.
Data exclusions	We document all exclusion phases following the PRISMA protocol in our article. Specifically, we stored 264 records in the reference management system Zotero for retrieval. Of those, we rejected a total of 217 papers that did not meet our inclusion criteria. We rejected 87 papers based on their abstract and 130 after assessment of the full text. All exclusion decisions are documented in a publicly available spreadsheet that can be found on the OSF project page (https://osf.io/96zbp/).
Non-participation	We rely on published data and that information was not available in most cases.
Randomization	Almost all studies used a within-participant design. Between-participant studies randomized participants to see either true or false

Randomization | news.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒ ☐	Antibodies
☒ ☐	Eukaryotic cell lines
☒ ☐	Palaeontology and archaeology
☒ ☐	Animals and other organisms
☒ ☐	Clinical data
☒ ☐	Dual use research of concern
☒ ☐	Plants

Methods

n/a	Involved in the study
☒ ☐	ChIP-seq
☒ ☐	Flow cytometry
☒ ☐	MRI-based neuroimaging

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*