



Thinking more or thinking differently? Using drift-diffusion modeling to illuminate why accuracy prompts decrease misinformation sharing

Hause Lin^{a,b,*}, Gordon Pennycook^{a,c}, David G. Rand^{b,d}

^a Hill/Levene Schools of Business, University of Regina, Canada

^b Sloan School, Massachusetts Institute of Technology, USA

^c Department of Psychology, University of Regina, Canada

^d Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, USA

ARTICLE INFO

Keywords:

Misinformation
Decision making
Dual process
Drift-diffusion model
Social media

ABSTRACT

Recent experiments have found that prompting people to think about accuracy reduces misinformation sharing intentions. The process by which this effect operates, however, remains unclear. Do accuracy prompts cause people to “stop and think,” increasing deliberation? Or do they change what people think about, drawing attention to accuracy? Since these two accounts predict the same behavioral outcomes (i.e., increased sharing discernment following a prompt), we used computational modeling of sharing decisions with response time data, as well as out-of-sample ratings of headline perceived accuracy, to test the accounts’ divergent predictions across six studies ($N = 5633$). The results suggest that accuracy prompts do not increase the *amount* of deliberation people engage in. Instead, they increase the weight participants put on accuracy *while* deliberating. By showing that prompting people makes them think better even without thinking more, our results challenge common dual-process interpretations of the accuracy-prompt effect. Our findings also highlight the importance of understanding how social media distracts people from considering accuracy, and provide evidence for scalable interventions that redirect people’s attention.

Misinformation, particularly on social media, has become a focus of widespread societal concern and academic research (Lazer et al., 2018; Pennycook & Rand, 2021). Substantial effort has been invested in developing approaches to combat online misinformation. One such approach involves prompting users to consider accuracy, which has been shown to improve the quality of news they subsequently share (Epstein et al., 2021; Pennycook et al., 2021; Pennycook, McPhetres, Zhang, Lu, & Rand, 2020; Roozenbeek, Freeman, & van der Linden, 2021). Although this effect generalizes across different accuracy prompts, news headlines, and countries (Arechar et al., 2022; Pennycook & Rand, 2022), how the effect operates remains unclear. Do accuracy prompts cause people to think *more*, or do they change what people think *about* (or both)? Here we investigate this issue using computational modeling.

One perspective is rooted in recent work indicating that people often fail to attend to accuracy when making sharing decisions on social media, even though they express a preference for sharing only accurate content and can often discern true from false news when asked to evaluate content accuracy (Pennycook, Epstein, et al., 2021). These

seemingly contradictory findings reflect the consequences of having to navigate an information-rich world (Lorenz-Spreen, Mønsted, Hövel, & Lehmann, 2019; Simon, 1971) with limited attention capacities (Buschman & Kastner, 2015; Navon & Gopher, 1979): Online social media environments are designed to engage users and capture their attention, which then distracts them from considering content quality (Hills, 2019; Lorenz-Spreen, Lewandowsky, Sunstein, & Hertwig, 2020). Thus, according to the *attention* account, inattention to accuracy causes people to share misinformation unwittingly. As a result, simple accuracy prompts that redirect attention to accuracy (e.g., asking people to evaluate the accuracy of a random news headline) will increase the quality of news shared (Pennycook & Rand, 2022).

Alternatively, accuracy prompts could cause people to share higher-quality news by encouraging people to deliberate more (even without increasing the relative share of attention directed to accuracy). This *deliberation* account follows from the dual-process perspective wherein intuitive “System 1” processes are distinguished from analytic “System 2” processes (De Neys, 2021; Evans, 2008; Pennycook, Fugelsang, & Koehler, 2015). Specifically, dual-process research suggests that merely

* Corresponding author at: University of Regina, Hill and Levene Schools of Business, 3737 Wascana Parkway, Regina, Saskatchewan S4S 0A2, Canada.
E-mail address: hauselin@gmail.com (H. Lin).

increasing deliberation (i.e., “triggering System 2”; Fazio, 2020) could increase judgment accuracy by making people question their faulty intuitions (Kahneman, 2011; Pennycook et al., 2015; Stanovich & West, 2000). For example, the Cognitive Reflection Test (CRT; Frederick, 2005) presents questions (e.g., “If you’re running a race and you pass the person in second place, what place are you in?”; Thomson & Oppenheimer, 2016) that cue an incorrect intuitive response (first place), but people who spend more time deliberating are more likely to answer correctly (second place) (Travers, Rolison, & Feeney, 2016). Notably, people who perform better on the CRT can not only better discern true from false news (Bago, Rand, & Pennycook, 2020; Pennycook & Rand, 2019; Ross, Rand, & Pennycook, 2021), but may also share higher-quality news (Mosleh, Pennycook, Arechar, & Rand, 2021; Pennycook et al., 2020; but see Osmundsen, Bor, Vahlstrup, Bechmann, & Petersen, 2021; Sirlin, Epstein, Arechar, & Rand, 2021). Furthermore, the accuracy-prompt effect on news discernment is stronger among those who score higher on the CRT (Pennycook & Rand, 2022). Thus, dual-process theories would suggest that accuracy prompts should encourage people to deliberate more, which increases the quality of news shared.

Distinguishing between the attention and deliberation accounts can provide theoretical insights into why people share misinformation and practical insights into designing misinformation interventions. Crucially, our approach allows us to test the predictions of popular dual-process theories, which often assume that thinking more should be

associated with thinking better. Namely, the key prediction is that prompting people to consider accuracy is effective specifically because it increases deliberation (which has been shown to facilitate evaluation of the accuracy of news; Pennycook, 2022; Pennycook & Rand, 2019).

Here, we formalize these accounts using the drift-diffusion model (DDM) framework. DDMs assume a decision-making process in which an initial inclination favoring one discrete choice over another gradually changes over time, “drifting” in one direction until a decision boundary is reached (Ratcliff & McKoon, 2008). Fig. 1 illustrates various simulated trajectories, which plot the relative decision value in favor of sharing (upper bound) or not sharing (lower bound) a true (Fig. 1A) or false (Fig. 1B) news headline, as a function of time during a single choice. As the participant deliberates, the decision value “drifts” toward one boundary (corresponding to a discrete choice), along a line influenced by, among other things, the headline’s perceived accuracy. A decision is made once the decision value reaches one of the boundaries.

Crucially, DDMs enable us to make separable inferences about the (i) amount of deliberation that occurs and (ii) extent to which that deliberation is influenced by accuracy. By decomposing the joint distributions of choices (sharing intentions) and response times into parameters that capture these different processes (Evans & Wagenmakers, 2019; Ratcliff & McKoon, 2008), DDMs allow us to more directly adjudicate between the attention and deliberation accounts of the accuracy-prompt effect. This modeling approach does not assume that longer response times (or more accurate responses) indicate more deliberation, which is

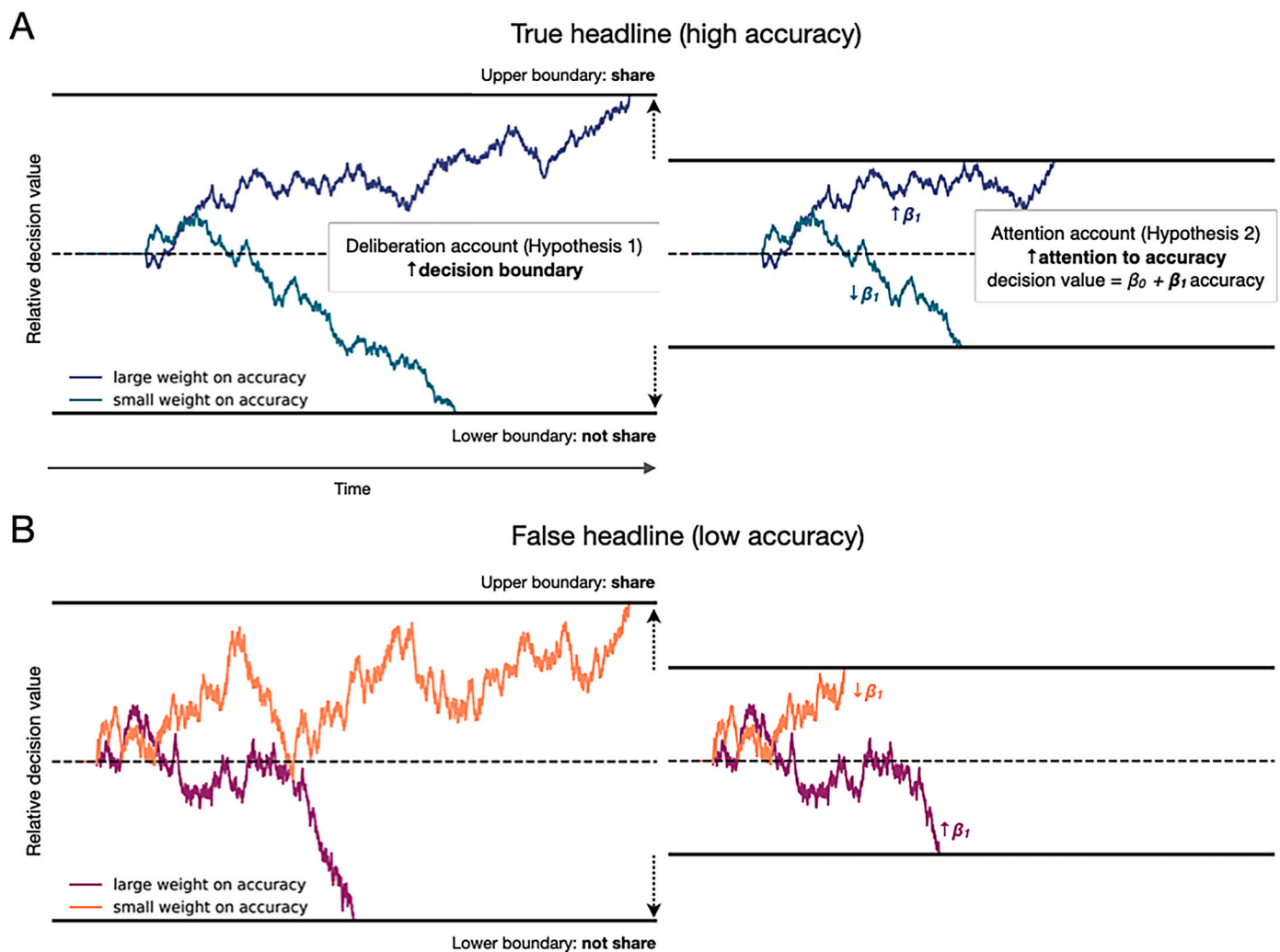


Fig. 1. Social media news-sharing diffusion model. Simulated trajectories for a true (A) and false (B) headline. Trajectories in the right panels are truncated versions (due to reduced boundaries) of the trajectories in the left panels.

potentially problematic because response times are determined by multiple processes, and thus can be difficult to interpret. For example, it is possible (and common) for experimental manipulations to have no effect on average response times and yet still affect the DDM parameter capturing the amount of deliberation (Lin, Saunders, Friese, Evans, & Inzlicht, 2020). Specifically, faster decisions can arise not only from engaging in less deliberation (smaller boundary; Fig. 1 right panels) but also from having an equivalent (or even higher) boundary coupled with a steeper drift rate (i.e., a decider with a steeper drift rate will reach an equivalent decision boundary more quickly than a decider with a more shallow drift rate).

Furthermore, the attention and deliberation accounts make the same predictions regarding choices—that accuracy prompts will increase sharing discernment (i.e., proportion true shared minus proportion false headlines shared). Thus, it is not enough to simply examine choices and response times, whereas computational modeling can help to differentiate between these two accounts.

If accuracy prompts make people deliberate more—as per the *deliberation* account rooted in dual-process theories—the decision boundaries should be larger in the accuracy prompt (vs. control) condition (Fig. 1, left versus right panels), allowing for more deliberation to occur before acting (Hypothesis 1). Alternatively, if accuracy prompts increase how much accuracy influences the trajectory of deliberation (i.e., increase accuracy's influence on the slope along which the decision value drifts)—as per the *attention* account—then only weight placed on accuracy (β_1) (but not the amount of deliberation) should be increased in the accuracy prompt (vs. control) condition (Hypothesis 2). In this case, true (high perceived accuracy) headlines will tend to be shared (Fig. 1A, blue trajectories) whereas false (low perceived accuracy) headlines will not be shared (Fig. 1B, magenta trajectories).

1. Method

1.1. Participants

The data come from a subset of the studies we conducted between 2017 and 2020 (see Pennycook & Rand, 2022), but the analytic approach here is completely novel (preregistration: osf.io/6vmfa). We focused on six studies (Table 1) in which American participants were recruited online via Amazon Mechanical Turk or Lucid. Participants were randomly assigned to a (i) control condition in which they completed a news sharing task, or (ii) treatment condition in which they were prompted to think about accuracy prior to the news sharing task by being asked to evaluate the accuracy of a single neutral headline.

The total sample size across the six studies is 6961 ($n_{\text{control}} = 2867$, $n_{\text{treatment}} = 4094$). Prior to preregistration, we randomly sampled roughly 20% of participants ($n = 1328$) from each condition within each study to perform exploratory analyses, leaving 80% ($n = 5633$) of the data “unseen.” The analyses and results below focus on the “unseen” data from these 5633 participants ($n_{\text{control}} = 2321$, $n_{\text{treatment}} = 3312$). Data and code can be found here: <https://doi.org/10.17605/osf.io/3dunc>

1.2. Preregistered exclusion criteria

We first excluded trials with response times faster than 0.15 s or slower than 30 s (see Ratcliff & Tuerlinckx, 2002). We then used a robust outlier-detection approach to exclude trials with outlier response times (± 3 median absolute deviation; see Leys, Ley, Klein, Bernard, & Licata, 2013; Leys, Delacre, Mora, Lakens, & Ley, 2019). Finally, participants with fewer than 15 trials left were excluded. This process excluded 295 (5.24%) participants ($n_{\text{control}} = 128$, $n_{\text{treatment}} = 167$), leaving 5338 participants ($n_{\text{control}} = 2193$, $n_{\text{treatment}} = 3145$).

1.3. Task design and procedure

Only participants who indicated they used social media were allowed to participate. Participants were randomly assigned to either a control or a treatment. Then, they completed a news-sharing task. They were then shown a set of true and false news headlines taken from social media. Participants indicated how likely they would be to share the headline if they were to see it on social media.

All the false headlines were found using popular fact-checking sites (e.g., snopes.com) whereas the true headlines came from reputable mainstream news sources. Headlines were generally selected for inclusion based on prior pre-testing and were related to politics or COVID-19 (for headline selection procedures, see Pennycook, Binnendyk, Newton, & Rand, 2021). For studies that used political headlines, the headlines were balanced on partisan lean based on pre-test results.

1.4. Computational framework: drift-diffusion model

Consistent with work on value-based choices (Hutcherson, Bushong, & Rangel, 2015; Tusche & Hutcherson, 2018), we assumed that news-sharing decisions can be described using a multi-attribute extension of the DDM (Ratcliff & McKoon, 2008; Smith & Ratcliff, 2004). In our framework (Fig. 1), the two boundaries are sharing intentions (share [upper bound], do not share [lower bound]). For the studies where sharing intentions were measured with a six-point Likert scale (Table 1), we binarized sharing intentions around the scale midpoint (corresponding to being unlikely versus likely to share). Standard DDMs interpret the change over time in decision value for each trial or decision as reflecting the accumulation of evidence (“evidence” here refers to things that lead one to favor sharing or not sharing). To model the more general deliberation process, we interpret the decision value's change over time as capturing a deliberative process wherein the decider determines the relative decision value (“drift rate” parameter) of sharing versus not sharing a headline, and in doing so, adjusts away from an initial (intuitive) starting value prior to deliberation (“starting-point bias” parameter). A decision (share or not share) is made once the relative decision value signal reaches either the upper or lower bound (“boundary” parameter). As with standard DDMs, we also estimated the “non-decision time” parameter (e.g., time to encode stimulus and produce motor responses).

Crucially, a variety of factors may influence the drift rate parameter, which reflects the relative decision value of sharing versus not sharing. Here, we focus specifically on the effect of the headline's perceived accuracy to test the *attention* account's prediction that accuracy prompts will increase the weight put on perceived accuracy when deliberating. Specifically, on each trial i , the relative decision value (v_t , i.e., drift rate) at each instantaneous time t depends on the following:

$$v_t = \beta_0 + \beta_1 \text{accuracy}_i + \varepsilon_t$$

where accuracy_i is the mean perceived accuracy rating (obtained from pre-tests with many participants) of a headline on trial i , β_0 is the mean decision value (or drift rate), β_1 is the weight placed on accuracy when deliberating, and ε_t is Gaussian noise on trial i . Note that headline perceived accuracy was measured with Likert scales and within-participant centered before model fitting to ensure β_0 has a meaningful interpretation (i.e., capturing the drift rate for headlines with average perceived accuracy; note this parameter was not the focus of our investigation or analyses).

1.5. Preregistered models and analyses

We used a differential evolution algorithm (Shinn, Lam, & Murray, 2020) to estimate (separately for each participant) the five parameters described above: decision boundary, weight on accuracy (β_1), mean drift rate or decision value (β_0), starting-point bias, and non-decision time. As

Table 1
Summary of studies

Study	Platform	Headlines	Trials	Response	Sample size		Sharing Discernment		RT	
					Control	Treatment	Control	Treatment	Control	Treatment
B	MTurk	24	16–24	6-point Likert	459	461	0.04 (0.19)	0.10 (0.20)	8.38 (3.78)	8.43 (3.24)
C	MTurk	24	15–24	6-point Likert	488	495	0.03 (0.23)	0.12 (0.27)	7.89 (3.11)	8.48 (3.68)
G	Lucid	20	15–20	6-point Likert	472	240	0.02 (0.22)	0.08 (0.25)	8.49 (4.12)	8.80 (4.16)
H	MTurk	20	15–20	6-point Likert	290	1453	0.15 (0.26)	0.16 (0.28)	6.92 (2.96)	6.96 (2.97)
L	Lucid	20	15–20	Binary	296	296	0.07 (0.24)	0.11 (0.26)	6.98 (4.13)	7.10 (4.12)
P	Lucid	20	15–20	6-point Likert	188	200	0.10 (0.24)	0.13 (0.26)	5.72 (3.88)	5.98 (3.70)

Note. See Pennycook and Rand (2022) for details on each study including its original publication. The study identifiers correspond to those in Pennycook and Rand (2022). Sharing discernment: Proportion true headlines shared minus proportion false headlines shared. Standard deviations are in parentheses.

noted in our theoretical accounts and preregistration, our focus is on the boundary and weight-on-accuracy (β_1) parameters; the remaining parameters were estimated because they are standard DDM parameters that are required to ensure good model fits.

To estimate the overall effects of condition across studies, we fit Bayesian mixed-effects models (maximal models with varying intercepts and slopes) using the R package *brms* (Bürkner, 2017): $brm(outcome \sim condition * platform + (1 + condition | study))$, where outcome is boundary, β_1 , sharing discernment, or response time; condition is control or treatment (coded -0.5 and 0.5), study is study ID, and platform is the recruitment platform (Lucid [0] or Amazon MTurk [1]; this dummy-coded variable was then z-scored). The population-level or fixed-effects priors were $N(0, SD_{outcome[control]} / 3)$. For each model, we ran four Markov chain Monte Carlo chains with 5000 samples and discarded the first 2500 samples (as burn-in). For each effect in the Bayesian model, we report the mean of the posterior samples and the 95% highest posterior-density interval (i.e., narrowest interval containing the specified mass). We also report Bayes factors (BF) by estimating Savage-Dickey density ratios (evidence against the point null $b = 0$), which provide relatively conservative estimates of the amount of evidence for the alternative (BF > 1) or null (BF < 1) hypothesis (Lee & Wagenmakers, 2014).

We also performed a preregistered robustness check where we used headline accuracy ratings that were separately obtained from participants classified as Democrats or Republicans to fit the models; that is, for Democrat (Republican) participants, the accuracy rating for any given headline was the mean rating obtained from Democrat (Republican) participants during pre-testing. This rating corresponds to the “believability” of headlines: If participants are paying greater attention to accuracy, believability should have a greater influence on their choices to share.

2. Results

Beginning with choice data descriptives (Table 1), we summarize past findings that accuracy prompts increased the quality of participants’ sharing intentions (Pennycook & Rand, 2022). Sharing discernment (proportion true headlines shared minus proportion false headlines shared) was higher in the treatment than control condition ($b = 0.05$ [0.004, 0.08], $d = 0.20$ [0.02, 0.34], BF = 3.90). There was also no meaningful difference between conditions in overall response times ($b = 0.23$ [−0.12, 0.60], $d = 0.06$ [−0.03, 0.17], BF = 0.47). This null effect provides a first indication that the prompts may not be causing people to deliberate more, although as discussed above, greater deliberation is not necessarily always associated with longer response times.

We now turn to our main analyses, which use DDMs to directly test the predictions of the deliberation and attention accounts. The DDM parameter values by condition and recruitment platform are shown in Fig. 2A. We found a null effect of condition on the boundary parameter ($b = -0.06$ [−0.24, 0.14], $d = -0.02$ [−0.11, 0.06], BF = 0.13; Fig. 2B left). Thus, contrary to Hypothesis 1 (deliberation account), accuracy prompts do not appear to make people deliberate more about what to

share. There were also null effects of platform ($b = 0.05$ [−0.43, 0.56], $d = 0.03$ [−0.20, 0.24], BF = 0.24) and condition-platform interaction ($b = -0.05$ [−0.22, 0.12], $d = -0.02$ [−0.10, 0.06], BF = 0.11).

Conversely, there was a positive effect of condition on the weight put on accuracy (Fig. 2B right), providing support for Hypothesis 2 (attention account): Accuracy prompts increased the amount of attention people paid to headline accuracy when deciding what to share ($b = 0.19$ [0.02, 0.37], $d = 0.19$ [0.01, 0.36], BF = 3.35¹). Once again, we found null effects of recruitment platform ($b = -0.02$ [−0.25, 0.22], $d = -0.02$ [−0.26, 0.21], BF = 0.32) and condition-platform interaction ($b = 0.01$ [−0.14, 0.15], $d = 0.01$ [−0.15, 0.15], BF = 0.20).

Thus, our results indicate that accuracy prompts increased weight-on-accuracy, although the effect is small and there was substantial variation across the six studies (effects were positive in all but one study, though only one study had 95% posterior-density intervals that did not include 0; Fig. S2). When we used a different modeling approach, the effects were also small and varied across studies (see Supplementary Information; Figs. S6 and S7). In addition, the effects were similar for the study where sharing intentions were measured with a binary response scale (boundary: $b = -0.13$ [−0.52, 0.25]; weight-on-accuracy: $b = 0.38$ [0.06, 0.72]), suggesting that the effects of treatment on the DDM parameters were robust to response format (i.e., binary response or binarized Likert data). Finally, the preregistered robustness check (using headline accuracy ratings separately obtained from participants classified as Democrats or Republicans) also found a similar pattern of results, suggesting that the effects are robust to how the perceived headline accuracy ratings were measured. Relative to the control condition, the treatment condition had larger weight-on-accuracy parameter estimates ($b = 0.17$ [0.01, 0.34], $d = 0.18$ [0.01, 0.35], BF = 2.94) but not boundary estimates ($b = -0.07$ [−0.29, 0.14], $d = -0.03$ [−0.13, 0.06], BF = 0.15).

3. Discussion

Here we used computational modeling to test two cognitive mechanisms by which accuracy prompts might increase the quality of news people share. We find small effects that are consistent with the attention account: Accuracy prompts increase sharing discernment by shifting peoples’ attention to accuracy while deciding what to share, rather than increasing the amount of deliberation per se (i.e., no treatment effects on boundary parameter or response times). These results challenge common dual-process interpretations of the accuracy-prompt effect, which assume that prompting people to consider accuracy is effective specifically because it increases deliberation. Instead, they show that people

¹ If we directly tested our directional preregistered hypotheses with one-sided BFs (evidence for $b > 0$ against $b < 0$; Morey & Wagenmakers, 2014), the BFs would provide stronger evidence: weight-on-accuracy parameter BF = 38.80; boundary parameter BF = 0.34; sharing discernment BF = 56.80. The Bayes factors in the main text are evidence against the point null $b = 0$ (Savage-Dickey density ratios; Wagenmakers, Lodewyckx, Kuriyal, & Grasman, 2010), which are more conservative estimates.

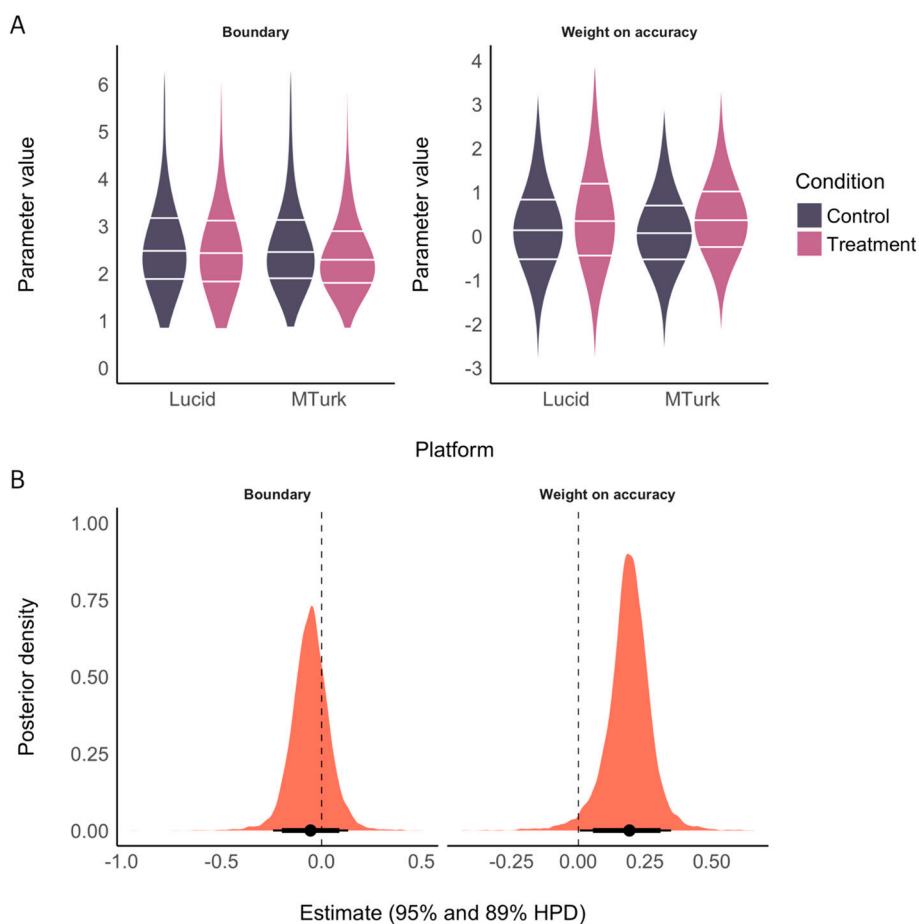


Fig. 2. Boundary and weight-on-accuracy parameters. (A) Parameter values as a function of recruitment platform and condition. Violin plots show the distribution of parameter values. White horizontal lines show the 25%, 50%, and 75% quantiles. (B) Bayesian posterior distributions for the effect of condition on boundary (left) and weight-on-accuracy (right) parameters. Positive estimates indicate larger parameter values in the treatment (vs. control) condition. Error bars are 95% and 89% highest posterior-density (HPD) intervals.

can think better even without thinking more.

These results resonate with recent misinformation research that suggests that “thinking better” is different from—and perhaps more important than—“thinking more.” For example, the disposition to specifically question one’s beliefs according to evidence (“actively open-minded thinking”) better predicted the ability to discern true from false COVID-19 news headlines than both CRT performance and one’s preference for effortful thinking or deliberation (Newton, Feeney, & Pennycook, 2021; see also Roozenbeek et al., 2022).

Our findings contribute to the body of evidence challenging the necessity of deliberation for increasing accuracy that is a common component of dual-process theories (although perhaps not a necessary one; see Evans, 2012). For example, although analytic thinking correlates positively with decision-making performance even after accounting for cognitive ability (Toplak, West, & Stanovich, 2011), the positive effects of analytic thinking can result from cognitive processes unrelated to increased deliberation. Indeed, when people are put under time pressure and are expected to provide intuitive or biased responses (due to insufficient time to trigger analytic processes), analytic thinking (e.g., CRT performance) still predicts decision-making performance (Bago & De Neys, 2019; Furlan, Agnoli, & Reyna, 2016; Raelison, Thompson, & De Neys, 2020; Thompson, Pennycook, Trippas, & Evans, 2018), and people might rely even more on strategic decision processes (Roberts, Teoh, & Hutcherson, 2022). Similarly, the present results highlight how people can think better and share higher-quality news without necessarily having deliberated more.

Attention-based explanations are therefore more consistent with our accuracy-prompt results than dual-process accounts. People have limited attention capacities, but have to navigate information-rich online social media environments (Hills, 2019; Lorenz-Spreen et al., 2020).

Given such processing constraints, it is likely that people flexibly direct attention toward the most salient or relevant information in a context-sensitive manner. Thus, prompting people redirects their attentional priorities to accuracy, but does not increase the overall amount of deliberation and available attention. This interpretation is consistent with studies showing that strategic attention deployment—rather than dual-process theories—explains changes in generosity and choice biases (Roberts et al., 2022; Teoh & Hutcherson, 2022; Teoh, Yao, Cunningham, & Hutcherson, 2020).

While our data reveal that accuracy prompts redirect attention to perceived headline accuracy, they do not offer insights into what other factors participants are attending to, or whether the prompts change other attentional priorities. Future work could test these predictions by examining whether accuracy prompts also direct attention away from headline features like political consistency and sensationalism.

It is important to bear in mind the difference between sharing and accuracy judgments, which have been shown to be largely distinct (Pennycook & Rand, 2021; Sirlin et al., 2021). Our findings challenge a dual-process interpretation of the accuracy prompt effect on news sharing. But a separate body of work has demonstrated correlational (Arechar et al., 2022; Pennycook et al., 2020; Pennycook & Rand, 2019; Pennycook & Rand, 2021; Ross et al., 2021) and causal (Bago et al., 2020; Martel, Pennycook, & Rand, 2020) that deliberation *does* increase people’s ability to discern true from false news when judging accuracy. These different findings further emphasize the gap between accuracy judgments and sharing intentions. Future work should also seek to reconcile correlational findings of more deliberative people sharing higher quality news (Mosleh et al., 2021; Pennycook et al., 2020; Ross et al., 2021; but see Osmundsen et al., 2021; Sirlin et al., 2021) with our results. Careful consideration of these conflicting findings may help to

resolve important inconsistencies in the dual-process literature.

In sum, our results indicate that accuracy prompts do *not* work by causing people to “stop and think” (United Nations, 2021), and therefore challenge common dual-process interpretations of the accuracy-prompt effect. Moreover, it is notoriously difficult to increase deliberation. Redirecting attention, however, might be relatively easier. Thus, the accuracy-prompt approach may be more effective and scalable than interventions that try to increase people’s propensity to engage in analytical thinking. For example, social media platforms could ask users to periodically consider content accuracy—particularly if platforms had reasons to believe the user is about to be exposed to misinformation.

CRedit authorship contribution statement

Hause Lin: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. **Gordon Pennycook:** Conceptualization, Methodology, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. **David G. Rand:** Conceptualization, Methodology, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Data availability

Data and code can be found at <https://doi.org/10.17605/osf.io/3ducn>

Acknowledgments

We thank Adam Bear, Rahul Bhui, Cameron Martel, Henrik Singmann, and Ben Tappin for their feedback and comments. We gratefully acknowledge funding from the Ethics and Governance of Artificial Intelligence Initiative of the Miami Foundation, the William and Flora Hewlett Foundation, the Reset initiative of Luminare (part of the Omidyar Network), the John Templeton Foundation, the TDF Foundation, the Canadian Heritage Digital Citizen Contribution Program, the Social Sciences and Humanities Research Council of Canada, and Google.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cognition.2022.105312>.

References

- Arechar, A. A., Allen, J. N. L., Cole, R., Epstein, Z., Garimella, K., Gully, A., ... Zhang, J. (2022). *Understanding and reducing online misinformation across 16 countries on six continents*. <https://doi.org/10.31234/osf.io/a9frz>
- Bago, B., & De Neys, W. (2019). The smart system 1: Evidence for the intuitive nature of correct responding on the bat-and-ball problem. *Thinking & Reasoning*, 25(3), 257–299. <https://doi.org/10.1080/13546783.2018.1507949>
- 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*. <https://doi.org/10.1037/xge0000729>
- Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1), 1–28. <https://doi.org/10.18637/jss.v080.i01>
- Buschman, T. J., & Kastner, S. (2015). From behavior to neural dynamics: An integrated theory of attention. *Neuron*, 88(1), 127–144. <https://doi.org/10.1016/j.neuron.2015.09.017>
- De Neys, W. (2021). On dual- and single-process models of thinking. *Perspectives on Psychological Science*, 117(12). <https://doi.org/10.1177/1745691620964172>, 1745691620964172.
- Epstein, Z., Berinsky, A. J., Cole, R., Gully, A., Pennycook, G., & Rand, D. G. (2021). Developing an accuracy-prompt toolkit to reduce COVID-19 misinformation online. *Harvard Kennedy School Misinformation Review*, 2(3), 1–12. <https://doi.org/10.37016/mr-2020-71>
- Evans, J. S. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59, 255–278. <https://doi.org/10.1146/annurev.psych.59.103006.093629>

- Evans, J. S. B. T. (2012). *Dual process theories of deductive reasoning: Facts and fallacies* (pp. 115–133). New York: Oxford University Press.
- Evans, N. J., & Wagenmakers, E.-J. (2019). Evidence accumulation models: Current limitations and future directions. <https://psyarxiv.com/74df9/download?format=pdf>.
- Fazio, L. (2020). Pausing to consider why a headline is true or false can help reduce the sharing of false news. *Harvard Kennedy School Misinformation Review*, 1(2). <https://doi.org/10.37016/mr-2020-009>
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19(4), 25–42. <https://doi.org/10.1257/089533005775196732>
- Furlan, S., Agnoli, F., & Reyna, V. F. (2016). Intuition and analytic processes in probabilistic reasoning: The role of time pressure. *Learning and Individual Differences*, 45, 1–10. <https://doi.org/10.1016/j.lindif.2015.11.006>
- Hills, T. T. (2019). The dark side of information proliferation. *Perspectives on Psychological Science*, 14(3), 323–330. <https://doi.org/10.1177/1745691618803647>
- Hutcherson, C. A., Bushong, B., & Rangel, A. (2015). A neurocomputational model of altruistic choice and its implications. *Neuron*, 87(2), 451–462. <https://doi.org/10.1016/j.neuron.2015.06.031>
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., ... Zittrain, J. L. (2018). The science of fake news. *Science*, 359(6380), 1094–1096. <https://doi.org/10.1126/science.aao2998>
- Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.
- Leys, C., Delacre, M., Mora, Y. L., Lakens, D., & Ley, C. (2019). How to classify, detect, and manage univariate and multivariate outliers, with emphasis on pre-registration. *International Review of Social Psychology*, 32(1). <https://doi.org/10.5334/irsp.289>
- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4), 764–766. <https://doi.org/10.1016/j.jesp.2013.03.013>
- Lin, H., Saunders, B., Friese, M., Evans, N. J., & Inzlicht, M. (2020). Strong effort manipulations reduce response caution: A preregistered reinvention of the ego-depletion paradigm. *Psychological Science*, 1-17. <https://doi.org/10.1177/0956797620904990>
- Lorenz-Spreen, P., Lewandowsky, S., Sunstein, C. R., & Hertwig, R. (2020). How behavioural sciences can promote truth, autonomy and democratic discourse online. *Nature Human Behaviour*, 4(11), 1102–1109. <https://doi.org/10.1038/s41562-020-0889-7>
- Lorenz-Spreen, P., Mønsted, B. M., Hövel, P., & Lehmann, S. (2019). Accelerating dynamics of collective attention. *Nature Communications*, 10(1). <https://doi.org/10.1038/s41467-019-09311-w>
- Martel, C., Pennycook, G., & Rand, D. G. (2020). Reliance on emotion promotes belief in fake news. *Cognitive Research: Principles and Implications*, 5(1), 1–20. <https://doi.org/10.31234/osf.io/a2ydw>
- Morey, R. D., & Wagenmakers, E.-J. (2014). Simple relation between Bayesian order-restricted and point-null hypothesis tests. *Statistics & Probability Letters*, 92, 121–124. <https://doi.org/10.1016/j.spl.2014.05.010>
- Mosleh, M., Pennycook, G., Arechar, A. A., & Rand, D. G. (2021). Cognitive reflection correlates with behavior on Twitter. *Nature Communications*, 12(1), 1–10. <https://doi.org/10.1038/s41467-020-20043-0>
- Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review*, 86(3), 214–255. <https://doi.org/10.1037/0033-295X.86.3.214>
- Newton, C., Feeney, J., & Pennycook, G. (2021). *On the disposition to think analytically: Four distinct intuitive-analytic thinking styles*. <https://doi.org/10.31234/osf.io/r5wez>
- Osmundsen, M., Bor, A., Vahlstrup, P. B., Bechmann, A., & Petersen, M. B. (2021). Partisan polarization is the primary psychological motivation behind “fake news” sharing on Twitter. *American Political Science Review*, 1-17. <https://doi.org/10.1017/S0003055421000290>
- Pennycook, G. (2022). A framework for understanding reasoning errors: From fake news to climate change and beyond. *Advances in Experimental Social Psychology*. <https://doi.org/10.31234/osf.io/j3w7d>
- Pennycook, G., Binnendyk, J., Newton, C., & Rand, D. G. (2021). A practical guide to doing behavioral research on fake news and misinformation. *Collabra: Psychology*, 7(1). <https://doi.org/10.1525/collabra.25293>
- 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., Fugelsang, J. A., & Koehler, D. J. (2015). Everyday consequences of analytic thinking. *Current Directions in Psychological Science*, 24(6), 425–432. <https://doi.org/10.1177/0963721415604610>
- 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(2019), 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>
- Pennycook, G., & Rand, D. G. (2022). Accuracy prompts are a replicable and generalizable approach for reducing the spread of misinformation. *Nature Communications*, 13(1), 2333. <https://doi.org/10.1038/s41467-022-30073-5>
- Raoelison, M., Thompson, V. A., & De Neys, W. (2020). The smart intuitor: Cognitive capacity predicts intuitive rather than deliberate thinking. *Cognition*, 204, Article 104381. <https://doi.org/10.1016/j.cognition.2020.104381>

- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20(4), 873–922. <https://doi.org/10.1162/neco.2008.12-06-420>
- Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, 9(3), 438–481. <https://doi.org/10.3758/BF03196302>
- Roberts, I. D., Teoh, Y. Y., & Hutcherson, C. A. (2022). Time to pay attention? Information search explains amplified framing effects under time pressure. *Psychological Science*, 33(1), 90–104. <https://doi.org/10.1177/09567976211026983>
- Roozenbeek, J., Freeman, A. L. J., & van der Linden, S. (2021). How accurate are accuracy-nudge interventions? A preregistered direct replication of pennycook et al. (2020). *Psychological Science*, 32(7), 1169–1178. <https://doi.org/10.1177/09567976211024535>
- Roozenbeek, J., Maertens, R., Herzog, S. M., Geers, M., Kurvers, R. H. J. M., Sultan, M., & van der Linden, S. (2022). Susceptibility to misinformation is consistent across question framings and response modes and better explained by open-mindedness and partisanship than analytical thinking. *Judgment and Decision making*, 17(3), 547–573. <https://journal.sjdm.org/22/220228/jdm220228.pdf>
- Ross, R. M., Rand, D., & Pennycook, G. (2021). Beyond “fake news”: The role of analytic thinking in the detection of inaccuracy and partisan bias in news headlines. *Judgment and Decision Making*, 16(2), 484–504.
- Shinn, M., Lam, N. H., & Murray, J. D. (2020). A flexible framework for simulating and fitting generalized drift-diffusion models. *eLife*, 9. <https://doi.org/10.7554/eLife.56938>
- Simon, H. A. (1971). Designing organizations for an information-rich world. In *Computers, communications, and the public interest* (pp. 37–52). Johns Hopkins University Press.
- Sirlin, N., Epstein, Z., Arechar, A. A., & Rand, D. G. (2021). Digital literacy is associated with more discerning accuracy judgments but not sharing intentions. *Harvard Kennedy School Misinformation Review*. <https://doi.org/10.37016/mr-2020-83>
- Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, 27(3), 161–168. <https://doi.org/10.1016/j.tins.2004.01.006>
- Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Sciences*, 23(5), 645–665. <https://doi.org/10.1017/s0140525x00003435>
- Teoh, Y. Y., & Hutcherson, C. A. (2022). The games we play: Prosocial choices under time pressure reflect context-sensitive information priorities. *Psychological Science*. <https://doi.org/10.1177/09567976221094782>
- Teoh, Y. Y., Yao, Z., Cunningham, W. A., & Hutcherson, C. A. (2020). Attentional priorities drive effects of time pressure on altruistic choice. *Nature Communications*, 11(1), 3534. <https://doi.org/10.1038/s41467-020-17326-x>
- Thompson, V. A., Pennycook, G., Trippas, D., & Evans, J. S. B. T. (2018). Do smart people have better intuitions? *Journal of Experimental Psychology: General*, 147(7), 945–961. <https://doi.org/10.1037/xge0000457>
- Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision making*, 11(1).
- Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, 39(7), 1275–1289. <https://doi.org/10.3758/s13421-011-0104-1>
- Travers, E., Rolison, J. J., & Feeney, A. (2016). The time course of conflict on the cognitive reflection test. *Cognition*, 150, 109–118. <https://doi.org/10.1016/j.cognition.2016.01.015>
- Tusche, A., & Hutcherson, C. A. (2018). Cognitive regulation alters social and dietary choice by changing attribute representations in domain-general and domain-specific brain circuits. *eLife*, 7, Article e31185. <https://doi.org/10.7554/eLife.31185.001>
- United Nations. (2021). Verified #PledgeToPause. <https://shareverified.com/pledge-to-pause/>.
- Wagenmakers, E. J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the savage-dickey method. *Cognitive Psychology*, 60(3), 158–189. <https://doi.org/10.1016/j.cogpsych.2009.12.001>