

Question framing affects accurate-inaccurate discrimination in responses to sharing questions, but not in responses to accuracy questions

Raoul Bell & Axel Buchner

Article - Version of Record



Suggested Citation:

Bell, R., & Buchner, A. (2024). Question framing affects accurate-inaccurate discrimination in responses to sharing questions, but not in responses to accuracy questions. *Scientific Reports*, 14, Article 29005. <https://doi.org/10.1038/s41598-024-80296-3>

Wissen, wo das Wissen ist.



UNIVERSITÄTS- UND
LANDESBIBLIOTHEK
DÜSSELDORF

This version is available at:

URN: <https://nbn-resolving.org/urn:nbn:de:hbz:061-20241210-124405-5>

Terms of Use:

This work is licensed under the Creative Commons Attribution 4.0 International License.

For more information see: <https://creativecommons.org/licenses/by/4.0>



OPEN Question framing affects accurate-inaccurate discrimination in responses to sharing questions, but not in responses to accuracy questions

Raoul Bell & Axel Buchner

Previous research suggests that even when people are capable of judging to the best of their knowledge whether claims are accurate or inaccurate, they do not sufficiently discriminate between accurate and inaccurate information when asked to consider whether they would share stories on social media. However, question framing ("To the best of your knowledge...", "Would you consider...?") differed between the questions about accuracy and the questions about sharing. Here we examine the effects of question framing on responses to accuracy questions and responses to sharing questions. The framing of accuracy questions had no effect on accurate-inaccurate discrimination. In contrast, accurate-inaccurate discrimination in response to sharing questions increased when participants were asked to respond, to the best of their knowledge, whether they would share claims compared to when they were asked whether they would consider sharing stories. At a theoretical level, the findings support the inattention-based account, according to which contextual cues shifting the focus toward accuracy can enhance accurate-inaccurate discrimination in sharing responses. At a methodological level, these findings suggest that researchers should carefully attend to the verbal framing of questions about sharing information on social media, as the framing may significantly influence participants' focus on accuracy.

Keywords Fake news, Truth judgment, Verbal framing effect, Accuracy prompt, Sharing intention

The widespread use of the internet makes it possible to easily access, provide and disseminate information. On the positive side, the efficient flow of information can help people to solve problems, make informed decisions, enhance education and promote democracy by facilitating political participation^{1–6}. On the negative side, however, this same ease of sharing information also allows for the spread of inaccurate content to undermine public discourse and individual decision making. One notable example is fake news—inaccurate content presented as authentic news, often intended to deceive internet users to further agendas that may be more or less obvious^{7–10}. Democracy relies on citizens being accurately informed and maintaining trust in the political process. Political fake news may lead to manipulated public opinion, swayed elections and polarized societies, thereby weakening democracy^{11–13}. Similarly, health-related fake news can mislead people about the severity and nature of health conditions as well as the effectiveness and safety of medical treatments and it can generally erode the trust in science, leading to false beliefs and risky health behaviors such as avoiding vaccinations, using unproven and potentially dangerous treatments or ignoring public health guidelines^{14,15}. Given these damaging societal and individual effects, it is crucial to understand why people share inaccurate information on social media. This topic is urgent, as fake news is likely to increase both in quantity and plausibility due to the ongoing advances in generative artificial intelligence (AI)¹⁶.

Pennycook and colleagues^{17,18} tested three competing accounts of why people share inaccurate information on social media. First, according to the *preference-based account*, people do not care whether the headlines they share are accurate or inaccurate because they prioritize other objectives such as entertaining their audience or promoting their political views, and thus people do not discriminate between accurate and inaccurate information because they are only interested in whether the information furthers their goals. Second, according

Department of Experimental Psychology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany. ✉email: raoul.bell@hhu.de

to the *confusion-based account*, people lack the ability to discriminate between accurate and inaccurate information. Third, according to the *inattention-based account* people generally consider it important to share only accurate information on social media and are able to discriminate between accurate and inaccurate information reasonably well when directly asked to judge the accuracy of the information, but they fail to pay sufficient attention to accuracy when sharing information on social media because they are preoccupied with other concerns such as whether their post will receive engagement or positive reactions.

Pennycook et al.¹⁸ directly asked their participants to indicate, on a scale ranging from “not at all important” to “extremely important”, how important it was for them to share only accurate information. The modal answer was “extremely important” which, along with accompanying experimental data, provides evidence against the preference-based account. In two highly influential studies with the same innovative basic design, participants in an accuracy condition were asked to decide, to the best of their knowledge, whether the claims in the headlines they were presented with were accurate. Participants in a sharing condition were asked whether they would consider sharing the stories online. Participants were able to discriminate between accurate and inaccurate statements fairly well when they were directly asked to evaluate the accuracy of claims to the best of their knowledge, ruling out the confusion-based account. In contrast, participants were less likely to discriminate between accurate and inaccurate statements when considering whether they would share stories. Pennycook et al.¹⁸ interpreted these findings as suggesting that people are well able to discriminate between accurate and inaccurate information but may nevertheless fail to sufficiently consider accuracy when sharing information on social media, in line with the inattention-based account [see also¹⁷].

The inattention-based account promises immediate and practical solutions to the challenge of fighting fake news in that contextual cues that cause people to focus more closely on accuracy should effectively cause social-media users to discriminate more between accurate and inaccurate information in their sharing responses. This short-term perspective is appealing for practitioners because it offers cost-effective solutions and it is also attractive for researchers because the effectiveness of short-term interventions can readily be assessed. It is therefore not surprising that the inattention-based account has proven to be extremely influential in guiding research and discussion on how to combat inaccurate information on social media^{19–22}. The account is also supported by findings showing that the discrimination between accurate and inaccurate statements in sharing decisions is affected by situational factors such as accuracy prompts^{17,18,23} and individual factors such as the tendency for cognitive reflection^{17,24}.

However, an alternative explanation exists for why participants in the studies of Pennycook et al.^{17,18} discriminated reasonably well between accurate and inaccurate information when responding to the questions about accuracy but were less likely to do so when responding to the questions about sharing information on social media. This explanation is much more prosaic than the inattention-based account and refers to the framing of the questions. Specifically, regarding accuracy, participants were asked “*To the best of your knowledge, is the claim in the above headline accurate?*”, whereas, regarding sharing, participants were asked “*Would you consider sharing this story online (for example, through Facebook or Twitter)?*” (emphasis added).

Questions such as those just mentioned function as frames²⁵ that direct attention to particular aspects of the information relevant to the responses^{26,27}. According to classical framing theory, frames provide interpretive structures that guide individuals’ understanding of a situation and set expectations for their behavior^{25,28}. In line with this understanding, prompting individuals to “consider” sharing information may activate a framework suggesting a hypothetical, exploratory engagement, where respondents may be less likely to rigorously assess the accuracy of the information. In contrast, the frame “to the best of your knowledge” directly prompts individuals to retrieve relevant knowledge from memory, encouraging them to critically evaluate the accuracy of the information at hand. According to models of decision making, the consideration stage is an initial step in which potential alternative responses are retrieved from memory *before* being deliberated and critically evaluated^{29–31}. To illustrate, when seeing a strangely shaped shiny object flying past the window, many people may briefly consider, among other explanations, the unlikely hypothesis that it could be an alien spaceship. However, if pressured to respond to the best of their knowledge about whether the shiny object was actually an alien spaceship, many of those who were willing to consider the alien-spaceship hypothesis may come to the conclusion that the answer is “no”. Additionally, the two framings also differ in the use of the terms “claim” and “story”, as highlighted in the questions above. A “claim” may be associated with a need for accuracy whereas accuracy may not be a primary requirement for a “story”. Thus, the framing of the question alone may explain why participants put more focus on accuracy when judging *claims to the best of their knowledge* than when *considering stories*.

Given the extensive evidence for the influence of verbal framing on behavior^{32–37}, it is important to examine whether the framing of the questions is the cause of the dissociation between the responses to the accuracy questions and the responses to the sharing questions. Testing how question framing affects the responses to accuracy and sharing questions is important, not the least because questions with the framing scrutinized here have been used by the original authors in several other studies^{23,24,38–42} as well as by other research groups^{22,43–47}. According to the inattention-based account^{17,18}, people should be well able to discriminate between accurate and inaccurate statements in response to accuracy questions but should discriminate less between accurate and inaccurate statements in response to sharing questions. If, however, the differences in the framing of the questions were the driving factor for the dissociation, then the observed differences in accurate-inaccurate discrimination between accuracy questions and sharing questions should disappear if the framing is the same for accuracy questions and sharing questions. Specifically, people’s accurate-inaccurate discrimination should be generally better when asked to respond to claims to the best of their knowledge and worse when asked to consider stories, irrespective of whether the questions are about accuracy or sharing. This *question-framing hypothesis* is based on the assumption that asking participants to respond to claims to the best of their knowledge implies a plea to respond based on factual knowledge, which should lead to a focus on accuracy and, hence, to better

accurate-inaccurate discrimination. In contrast, asking participants to consider stories plausibly encourages less scrutiny of a statement, resulting in a comparatively lower emphasis on accuracy and, hence, a lower level of accurate-inaccurate discrimination. The question-framing hypothesis thus implies that the previously observed dissociation between sharing responses and accuracy responses^{17,18} is solely due to the difference in the verbal framing of the accuracy questions compared to the sharing questions.

Experiment 1 Methods

Participants

The participants were recruited by advertising the experiment on the homepage of the Psychology Department, on social media and via email, mostly targeting students of Heinrich Heine University Düsseldorf. We aimed at recruiting as many participants as possible within the three weeks during which the experiment was online. The data of 41 participants who had started to respond to accuracy or sharing questions had to be excluded because these participants did not complete the experiment or withdrew their consent to the use of their data. The data of four participants had to be excluded because these participants were under 18 years old and thus could not legally consent to the use of their data. The final data set comprised the data of $N=451$ participants (371 female, 76 male, 4 diverse) with a mean age of 24 ($SD=8$) years, most of whom (415) were native German speakers. Given that participants were mostly recruited among students, the sample was well-educated—98% of the participants had a qualification at the university-entrance level or higher. On average participants reported to spend 3 h ($SD=2$) daily on social media. When asked to indicate, on a scale from “not at all” to “extremely”, how important it was to share only accurate content on social media, the participants’ modal answer was that they found it “extremely” important. This is parallel to what was reported by Pennycook et al.¹⁸. When asked to indicate, on a scale from “never” to “all the time”, how often they shared news on social media, the modal response was that they “rarely” did so. Participants were randomly assigned to four groups that differed with respect to the type and framing of the questions about statements participants were presented with. Specifically, 111 participants were asked to decide, to the best of their knowledge, whether the claims in the headlines were accurate, 118 participants were asked whether they would consider that the stories could be accurate, 105 participants were asked to decide, to the best of their knowledge, whether they would share the claims in the headlines online and 117 participants were asked to consider whether they would share the stories online. A sensitivity analysis with G*Power⁵⁶ showed that, with a sample size of $N=451$ participants, a main effect of question framing (“To the best of your knowledge...”, “Would you consider...”) on the accurate-inaccurate discrimination as small as $\eta_p^2 = 0.03$ could be detected at an α level of 0.05 with a statistical power of $1 - \beta = 0.95$.

Ethics statement

The Ethics Committee of the Faculty of Mathematics and Natural Sciences at Heinrich Heine University Düsseldorf has granted ethical approval for the experiments reported in this article. The experiments were conducted in compliance with the Declaration of Helsinki. Prior to participation, informed content was obtained from all participants. At the end of the experiment, participants were informed that, in addition to accurate statements, inaccurate statements had been presented. They were then shown each inaccurate statement along with its corresponding accurate version. Finally, participants were given the opportunity to withdraw their consent to the use of their data.

Materials

A list of 50 accurate and 50 inaccurate statements were taken from a fact-checking service provided by the public broadcasting service in Germany with the intent to combat inaccurate information⁴⁸. In other words, we relied on the fact checkers of Germany’s public broadcasting service to evaluate the truthfulness of the statements. The statements covered a range of topics including medical, environmental and scientific subjects. Examples of accurate statements are “Noise increases the risk of heart attack” and “Offshore wind farms can change marine ecosystems”. Examples of inaccurate statements are “The sudden infant death syndrome is caused by vaccinations” and “Harvard study proves: Wind turbines are partly to blame for global warming”. Each statement was accompanied by a suitable picture to mimic the typical presentation of such statements on social media platforms [cf.⁴⁹]. The pictures accompanying each statement were sourced from internet resources and image databases specifically designed for psychological research^{50–52}. An example statement, as presented in the present experiments, is shown in Fig. 1.

Statements for the main experiment were selected based on the outcomes of a separate *norming study* involving $N=240$ participants, each of whom received a mix of 50 accurate and 50 inaccurate statements in an order that was randomized separately for each individual. Participants were asked to assess the accuracy of each statement. The statements were presented in the same way as in the condition of the experiment proper in which accuracy was to be assessed using the *to-the-best-of-your-knowledge* framing. Participants were asked: “To the best of your knowledge, is the claim in the headline accurate?” with “yes” and “no” as response options. For the experiment proper, we selected the 20 accurate headlines that had received the highest proportion of “yes” responses ($M=0.73$, $SD=0.11$) and the 20 inaccurate headlines that had received the lowest proportion of “yes” responses ($M=0.23$, $SD=0.09$). These selection criteria ensured that accurate-inaccurate discrimination, given the *to-the-best-of-your-knowledge* framing, was closely aligned with that observed in studies upon which the present research was based^{17,18}.

Procedure

At the start of the experiment proper, participants were instructed to complete the task independently and without interruptions. If this was not possible, they were advised to exit the experiment by closing the browser



Climate change is shifting the seasons

Fig. 1. An example statement taken from a fact-checking service provided by the public broadcasting service in Germany⁴⁸. The picture was taken from the Open Affective Standardized Image Set [OASIS⁵¹]. The statement is the English translation of one of the German statements used in Experiments 1 and 2, which was used in Experiment 3.

window. After having consented to participate, participants were informed, depending on the framing of the questions in the condition they were assigned to, that they would see a series of headlines or stories taken from social media such as Facebook, X (formerly Twitter), WhatsApp, Telegram, Signal or Instagram. When translating the questions to German for the German-speaking participants, care was taken to ensure that the wordings of the questions was as close as possible to those used in previous studies^{17,18}. The statements were presented individually, one after another and in different random order for each participant, in the format illustrated in Fig. 1. Translated into English, the accuracy questions were “To the best of your knowledge, is the claim in the headline accurate?” and “Would you consider that the story could be accurate?”. The sharing questions were “To the best of your knowledge, would you share the claim in the headline online?” and “Would you consider sharing the story online?”. Participants answered the questions by selecting a “yes” or “no” checkbox. Which checkbox was presented as the left checkbox was counterbalanced across participants. After having selected a checkbox, a “continue” button was presented. When this button was clicked, the next question was presented.

When participants had responded to all questions, they were transferred to a page on which they were informed that some of the statements they had just seen had been inaccurate. They were told that the accurate statement would be provided in the next section of the experiment. Following current scientific recommendations on how to effectively correct misinformation⁵³, each inaccurate statement was first repeated before it was refuted. For instance, participants were informed: “*The following is inaccurate:* Sudden infant death syndrome is caused by vaccinations”. Each inaccurate statement was immediately followed by a detailed correction contrasting the inaccurate statement with accurate information. For instance, the correction to the statement just mentioned was: “*The following is correct:* Cases of sudden infant death syndrome have been falling steadily for years, although the number of vaccinations has remained at least the same. Studies show that the risk of sudden infant death is not increased after vaccinations. Berlin scientists found that the opposite is true: As vaccination rates increase, cases of sudden infant death syndrome fall.” The corrections were taken from the same fact-checking service as the accurate and inaccurate statements⁴⁸.

At the conclusion of the experiment, participants were asked to indicate how often they shared news on social media using a 5-point scale (1 = “never”, 2 = “rarely”, 3 = “sometimes”, 4 = “frequently”, 5 = “constantly”). They were also asked to rate the importance of sharing only accurate content on social media, using another 5-point scale (1 = “not at all”, 2 = “a little”, 3 = “medium”, 4 = “very”, 5 = “extremely”). Furthermore, participants were asked how much time (in hours per day) they spent on social media (for example, Facebook, X, WhatsApp, Telegram, Signal or Instagram). After having reported their age, gender, native language and education level, participants were informed about the purpose of the study. They were then given the option to withdraw their consent to the use of their data—which they had given at the start of the experiment—by selecting the “No, I withdraw my consent to the use of my data” option. Finally, participants were thanked for their participation.

Results

The significance level was set to $\alpha = 0.05$ for all analyses reported in this article. As the main hypotheses pertain to participants’ accurate-inaccurate discrimination, the discrimination index P_r of the two-high threshold model was calculated for each individual based on the proportions of “yes” responses to accurate and inaccurate statements. This discrimination measure was selected because it has been favorably evaluated in a series of validation experiments for the present 2×2 data structure compared to competing measurement models⁵⁴. The mean discrimination index P_r is displayed in Fig. 2. For completeness, the mean bias index B_p , indicating the tendency toward responding “yes” when in an uncertain state, is reported in Table 1.

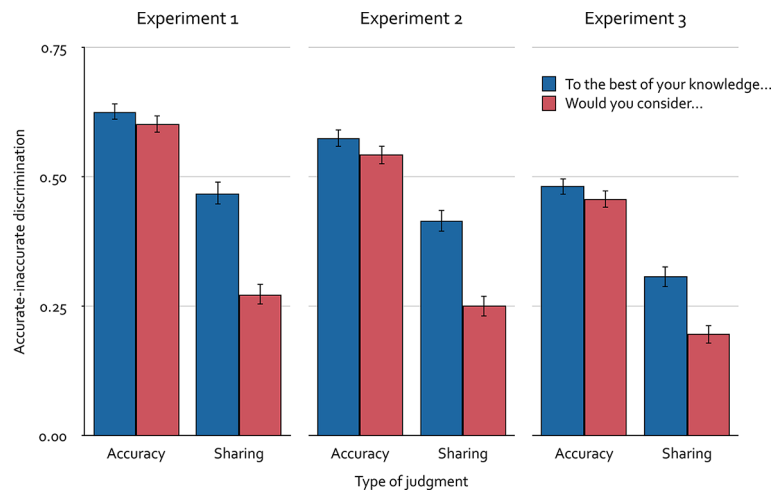


Fig. 2. Mean accurate-inaccurate discrimination index P_r^{54} as a function of type of judgment (accuracy, sharing) and question framing (“To the best of your knowledge...”, “Would you consider...?”) in Experiments 1, 2 and 3. Error bars represent the standard errors of the means.

	Accuracy				Sharing			
	“To the best of your knowledge...”		“Would you consider...”		“To the best of your knowledge...”		“Would you consider...”	
	Mean	(Standard error)	Mean	(Standard error)	Mean	(Standard error)	Mean	(Standard error)
Experiment 1	0.41	(0.02)	0.59	(0.02)	0.19	(0.02)	0.15	(0.02)
Experiment 2	0.37	(0.02)	0.49	(0.02)	0.32	(0.02)	0.26	(0.02)
Experiment 3	0.37	(0.02)	0.37	(0.02)	0.23	(0.02)	0.22	(0.02)

Table 1. Mean bias index B_i measuring the tendency toward responding “yes” when in an uncertain state⁵⁴, as a function of type of judgment (accuracy, sharing) and question framing (“To the best of your knowledge...”, “Would you consider...”). Values in parentheses represent the standard errors of the means.

A 2×2 between-subjects analysis of variance was performed to examine how type of judgment (accuracy, sharing) and question framing (“To the best of your knowledge...”, “Would you consider...”) influenced accurate-inaccurate discrimination, as measured by the discrimination index P_r . Accurate-inaccurate discrimination was significantly better when responding to accuracy questions than when responding to sharing questions, $F(1, 447) = 169.32, p < 0.001, \eta_p^2 = 0.27$. Additionally, the framing “To the best of your knowledge...” on average led to better accurate-inaccurate discrimination than the framing “Would you consider...”, $F(1, 447) = 34.38, p < 0.001, \eta_p^2 = 0.07$. Importantly, there was a significant interaction between type of judgment and question framing, $F(1, 447) = 20.96, p < 0.001, \eta_p^2 = 0.04$. The way the accuracy questions were framed did not influence accurate-inaccurate discrimination, $t(227) = 1.06, p = 0.29, d = 0.14$. In contrast, accurate-inaccurate discrimination was better when the framing of the sharing questions was “To the best of your knowledge...” than when it was “Would you consider...”, $t(220) = 6.51, p < 0.001, d = 0.88$. The results thus suggest that question framing has no influence on accurate-inaccurate discrimination when responding to accuracy questions but affects accurate-inaccurate discrimination when responding to sharing questions.

Discussion

The framing of questions did not affect accurate-inaccurate discrimination uniformly but instead had distinct effects depending of whether participants responded to questions about accuracy or to questions about sharing. In response to accuracy questions, accurate-inaccurate discrimination was equally good regardless of how the questions were framed. In response to the sharing questions, in contrast, the requirement to respond to claims to the best of their knowledge rather than to consider stories caused participants to discriminate better between accurate and inaccurate information.

Experiment 2

A limitation of Experiment 1 is that, even though the experiment was conducted online, the participants were mostly students. Students may differ in how they respond to the accuracy and sharing questions from participants recruited via online panels which were primarily focused on in previous research^{17,18}. Specifically, it seems conceivable that students are particularly responsive to verbal framing effects, for example, because they are often exposed to nuanced language in academic settings, which could heighten their sensitivity to how information is presented. For this reason and because direct replications are generally important⁵⁵, Experiment

2 was designed to serve as a direct replication of Experiment 1, with the main difference that the participants in Experiment 2 were recruited through an online panel provider.

Methods

Participants

To recruit a sample of active online users with residency in Germany, participants were recruited via the online panel provider *Cint* (<https://www.cint.com/>). We aimed at recruiting about 600 participants and ended data collection at the end of the day this criterion was reached. The data of 25 participants who had started to respond to accuracy or sharing questions had to be excluded because these participants did not complete the experiment or withdrew their consent to the use of their data. The data of two participants had to be excluded because these participants were under 18 years old and thus could not legally consent to the use of their data. The final data set comprised the data of $N = 595$ participants (323 female, 268 male, 4 diverse) with a mean age of 47 ($SD = 16$) years, most of whom (567) were native German speakers. The sample was characterized by a diverse range of education. As an indication, 54% (compared to 98% in Experiment 1) had a qualification at the university-entrance level or higher. When asked to indicate, on a scale from “not at all” to “extremely”, how important it was to share only accurate content on social media, the participants’ modal answer was that they found it “extremely” important. This is again parallel to what was reported by Pennycook et al.¹⁸. When asked to indicate, on a scale from “never” to “all the time”, how often they shared news on social media, the modal response was that they “rarely” did so. Participants were randomly assigned to four groups that differed with respect to the type and framing of the questions about statements participants were presented with. Specifically, 142 participants were asked to decide, to the best of their knowledge, whether the claims in the headlines were accurate, 148 participants were asked whether they would consider that the stories could be accurate, 154 participants were asked to decide, to the best of their knowledge, whether they would share the claims in the headlines online and 151 participants were asked to consider whether they would share the stories online. A sensitivity analysis with G*Power⁵⁶ showed that, with a sample size of $N = 595$ participants, a main effect of question framing (“To the best of your knowledge...”, “Would you consider...”) on accurate-inaccurate discrimination as small as $\eta_p^2 = 0.02$ could be detected at an α level of 0.05 with a statistical power of $1 - \beta = 0.95$.

Materials and procedure

Materials and procedure were identical to those of Experiment 1 except that the question about daily social media usage was omitted.

Results

The mean discrimination index P_p , representing accurate-inaccurate discrimination, is displayed in Fig. 2. For completeness, the mean bias index B_p , indicating the tendency toward responding “yes” when in an uncertain state⁵⁴, is reported in Table 1.

A 2×2 between-subjects analysis of variance was performed to examine how type of judgment (accuracy, sharing) and question framing (“To the best of your knowledge...”, “Would you consider...”) influenced accurate-inaccurate discrimination, as measured by the discrimination index P_p . Accurate-inaccurate discrimination was significantly better when responding to accuracy questions than when responding to sharing questions, $F(1, 591) = 139.81, p < 0.001, \eta_p^2 = 0.19$. Additionally, the framing “To the best of your knowledge...” on average led to better accurate-inaccurate discrimination than the framing “Would you consider...”, $F(1, 591) = 26.43, p < 0.001, \eta_p^2 = 0.04$. Importantly, there was a significant interaction between type of judgment and question framing, $F(1, 591) = 11.97, p < 0.001, \eta_p^2 = 0.02$. The way the accuracy questions were framed did not influence accurate-inaccurate discrimination, $t(288) = 1.36, p = 0.177, d = 0.16$. In contrast, accurate-inaccurate discrimination was better when the framing of the sharing questions was “To the best of your knowledge...” compared to when it was “Would you consider...”, $t(303) = 5.54, p < 0.001, d = 0.63$. The results thus confirm that question framing has no influence on accurate-inaccurate discrimination when responding to accuracy questions but affects accurate-inaccurate discrimination when responding to sharing questions.

Discussion

Even though, in Experiment 2, we used a panel provider to recruit participants with a more diverse educational background than that of the participants of Experiment 1 who were mostly students, the results were strikingly similar to those of Experiment 1. As in Experiment 1, the hypothesis that question framing affects accurate-inaccurate discrimination in response to accuracy questions has to be rejected. By contrast, accurate-inaccurate discrimination in response to sharing questions is increased when participants are asked to respond to claims to the best of their knowledge rather than to consider stories.

Experiment 3

A limitation of Experiments 1 and 2 is that both experiments were conducted in German with German participants. Therefore, the results may have depended on nuances of translation. This issue is more serious than it might appear at first glance. This is so because a word-by-word translation of the German phrase for “to the best of your knowledge” would be “according to the best of one’s knowledge and conscience”. The German translation thus refers not only to knowledge but also to conscience which may be conceived of as a considerable intensification of the request expressed by the phrase. To test whether the framing effects observed in Experiments 1 and 2 were due to the peculiarity of how “to the best of your knowledge” is translated into German, we performed a conceptual replication of Experiments 1 and 2 with the main difference that the experiment was conducted in English using a phrasing of the questions as in the studies of Pennycook et al.^{17,18} and relying on an online panel provider to recruit participants residing in Great Britain.

Methods

Participants

To collect data from a sample of active online users with residency in Great Britain, participants were recruited via the online panel provider *Bilendi* (<https://www.bilendi.de/>). We aimed at recruiting about 600 participants and ended data collection at the end of the day this criterion was reached. The data of 53 participants who had started to respond to accuracy or sharing questions had to be excluded because these participants did not complete the experiment or withdrew their consent to the use of their data. The data of one participant had to be excluded because this participant was under 18 years old and thus could not legally consent to the use of their data. The final data set comprised the data of $N=610$ participants (291 female, 318 male, 1 diverse) with a mean age of 57 ($SD=17$) years, most of whom (580) were native English speakers. The sample was characterized by a diverse range of education. As an indication, 52% had at least a qualification corresponding to a university entrance qualification in Germany which is roughly comparable to the same figure obtained for the German online-panel sample in Experiment 2. When asked to indicate, on a scale ranging from “not at all” to “extremely”, how important it was to share only accurate content on social media, the participants’ modal answer was that they found it “extremely” important. This is again parallel to what was reported by Pennycook et al.¹⁸ When asked to indicate, on a scale from “never” to “all the time”, how often they shared news on social media, the modal response was that they “rarely” did so. Participants were randomly assigned to four groups that differed with respect to the type and framing of the questions about statements participants were presented with. Specifically, 155 participants were asked to decide, to the best of their knowledge, whether the claims in the headlines were accurate, 155 participants were asked whether they would consider the stories to be accurate, 150 participants were asked to decide, to the best of their knowledge, whether they would share the claims in the headlines online and 150 participants were asked to consider whether they would share the stories online. A sensitivity analysis with G*Power⁵⁶ showed that, with a sample size of $N=610$ participants, a main effect of question framing (“To the best of your knowledge...”, “Would you consider...”) on accurate-inaccurate discrimination as small as $\eta_p^2=0.02$ could be detected at an α level of 0.05 with a statistical power of $1-\beta=0.95$.

Materials and procedure

Materials and procedure were identical to those of Experiment 2, except that the experiment was conducted in English. Depending on the framing condition, the accuracy questions were “To the best of your knowledge, is the claim in the headline accurate?” or “Would you consider the story to be accurate?”; the sharing questions were “To the best of your knowledge, would you share the claim in the headline online?” or “Would you consider sharing the story online?”. This ensured that the questions “To the best of your knowledge, is the claim in the headline accurate?” and “Would you consider sharing the story online?” corresponded to those used in previous studies^{17,18}. Minor modifications were made to the statements to adapt the stimulus material for participants living in Great Britain. For instance, the speed limit mentioned in one of the statements was changed from 30 kilometers per hour to 20 miles per hour.

Results

The mean discrimination index P_r , representing accurate-inaccurate discrimination, is displayed in Fig. 2. For completeness, the mean bias index B_p , indicating the tendency toward responding “yes” when in an uncertain state⁵⁴, is reported in Table 1.

A 2×2 between-subjects analysis of variance was performed to examine how type of judgment (accuracy, sharing) and question framing (“To the best of your knowledge...”, “Would you consider...”) influenced accurate-inaccurate discrimination, as measured by the discrimination index P_r . Accurate-inaccurate discrimination was significantly better when responding to accuracy questions than when responding to sharing questions, $F(1, 606)=155.58, p<0.001, \eta_p^2=0.20$. Additionally, the framing “To the best of your knowledge...” on average led to better accurate-inaccurate discrimination than the framing “Would you consider...”, $F(1, 606)=15.39, p<0.001, \eta_p^2=0.02$. Importantly, there was a significant interaction between type of judgment and question framing, $F(1, 606)=6.32, p=0.012, \eta_p^2=0.01$. The way the accuracy questions were framed did not influence accurate-inaccurate discrimination, $t(308)=1.11, p=0.269, d=0.13$. In contrast, accurate-inaccurate discrimination was better when the framing of the sharing questions was “To the best of your knowledge...” than when it was “Would you consider...”, $t(298)=4.15, p<0.001, d=0.48$. The results thus confirm again that question framing has no influence on accurate-inaccurate discrimination when responding to accuracy questions but affects accurate-inaccurate discrimination when responding to sharing questions.

Discussion

When the instructions and questions are presented in their original English phrasing, parallel to those used by Pennycook et al.^{17,18}, the conclusions remain consistent with those of the previous two experiments. The hypothesis that question framing affects accurate-inaccurate discrimination in response to accuracy questions has to be rejected. In contrast, when responding to sharing questions, accurate-inaccurate discrimination is better when participants are required to respond to claims to the best of their knowledge rather than when they are required to consider stories. Therefore, the findings of Experiments 1 and 2 cannot be attributed to peculiarities in the translation of the phrase “to the best of your knowledge” into German.

General discussion

Understanding why people fail to discriminate between accurate and inaccurate information on social media is crucial for developing effective strategies to combat fake news^{7,8}. According to the influential inattention-based account⁹, people fail to sufficiently discriminate between accurate and inaccurate information in their sharing responses even though they find it important to share only accurate information and can, in principle,

discriminate fairly well between accurate and inaccurate information when directly asked to judge the accuracy of the information. The reduction in accurate-inaccurate discrimination in response to sharing questions relative to accuracy questions is attributed to a lack of attention: People do not spontaneously pay enough attention to accuracy when sharing information on social media. A foundational finding supporting this account is that people show a lower degree of accurate-inaccurate discrimination when asked whether they would consider sharing stories than when asked to judge the accuracy of claims to the best of their knowledge. However, in previous studies^{17,18} the framing for sharing questions differed systematically from the framing for accuracy questions. Therefore, we examined, in three experiments, how the framing of the questions affects accuracy and sharing responses.

According to the question-framing hypothesis, question framing should affect accuracy and sharing responses in a parallel manner. Specifically, the framing “*To the best of your knowledge...*” should cause participants to focus more on accuracy than the framing “*Would you consider ...*” which should be less likely to encourage people to scrutinize the veracity of statements, resulting in a decreased emphasis on accuracy. Importantly, according to the question-framing hypothesis this framing effect should occur for both accuracy questions and sharing questions. Contradicting this prediction implied by the question-framing hypothesis, the results obtained in the present series of experiments demonstrate that question framing did not affect accurate-inaccurate discrimination uniformly but instead had distinct effects depending of whether participants responded to questions about accuracy or to questions about sharing. Question framing did not affect accurate-inaccurate discrimination in response to accuracy questions. In sharp contrast, question framing consistently affected accurate-inaccurate discrimination in response to sharing questions. Specifically, accurate-inaccurate discrimination in response to sharing questions was better when the framing stimulated participants to respond to claims to the best of their knowledge than when they were asked to consider stories. This pattern of results was consistently observed across all three experiments, regardless of whether participants were recruited among university students (Experiment 1) or through online panels in Germany (Experiment 2) or in Great Britain (Experiment 3).

Why does question framing affect accurate-inaccurate discrimination in response to sharing questions but not in response to accuracy questions? Responses to sharing questions may be more susceptible to framing than responses to accuracy questions because sharing on social media may involve multiple goals such as providing accurate content, seeking social approval, fostering social engagement and capturing attention. The competition among multiple goals, only one of which is to share accurate information, may explain why responses to sharing questions are more easily influenced by contextual cues than responses to accuracy questions, where the content of the question itself implies that accuracy is the singular goal. This interpretation is in fact fully consistent with Pennycook et al.’s^{17,18} inattention-based account. An important implication of the inattention-based account is that people fail to discriminate between accurate and inaccurate information in their sharing responses not because they lack the motivation or ability to do so but because they fail to sufficiently attend to accuracy, becoming distracted by competing goals such as that of fostering social engagement. Therefore, contextual cues that prompt people to focus more on accuracy should enhance accurate-inaccurate discrimination in sharing responses. As a practical consequence, any contextual cue that directs the focus of attention toward accuracy can be expected to improve the quality of information shared on social media. Here, asking participants to respond to whether they would share claims on social media to the best of their knowledge rather than to consider sharing stories is such a contextual cue that prompts people to focus more on accuracy. Consequently, accurate-inaccurate discrimination is better when the framing of the sharing questions appeals to people’s knowledge about facts than when it does not, consistent with the assumptions of the inattention-based account. When responding to accuracy questions, the focus is naturally on accuracy, as this is the subject of these questions, even without additional cues to emphasize accuracy. Therefore, asking people to judge the accuracy of claims to the best of their knowledge yields similar outcomes to asking them to consider the accuracy of stories, as both approaches encourage a focus on accuracy.

The present findings fit nicely with those of Capraro and Celadin²¹ who have already shown that the verbal framing of the response options of questions about sharing affects the proportions of accurate and inaccurate statements that are shared. Specifically, when the response option to the question “If you were to see the above article on Facebook, would you consider sharing it?” was phrased as “Yes; I think this news is accurate” instead of only “Yes”, then accurate-inaccurate discrimination in people’s sharing responses was increased. The results of the present study extend these previous findings by showing that subtle changes in the verbal framing of questions cause people to better discriminate between accurate and inaccurate information when deciding what to share on social media. These findings should encourage further investigation into how nuanced framing manipulations can enhance accurate-inaccurate discrimination in sharing decisions on social media. For instance, platforms could ask users “Are you confident that this information is accurate?” rather than simply offering a share button, leveraging framing to reduce the spread of misinformation.

However, as a limitation of the prospect to decrease reliance on fake news through framing interventions, it must be acknowledged that, given the ongoing advances in generative AI¹⁶, interventions aimed at shifting people’s focus on accuracy may eventually lose their effectiveness. These interventions fundamentally rely on people’s ability to distinguish between accurate and inaccurate information, but as AI-generated misinformation becomes increasingly difficult to discriminate from accurate information, features that prompt people to focus on accuracy will likely become less and less effective. For the time being, however, stimulating people to focus on accuracy may still serve as a useful tool in combating misinformation.

As another limitation of the present study, it must be noted that the question-framing manipulation used here comprised not only the phrases “*To the best of your knowledge*” and “*Would you consider...*” but also the terms “*claim*” and “*story*”, respectively. As mentioned in the introduction, a “*claim*” may be associated with a need for accuracy whereas accuracy may not be a primary requirement for a “*story*”. As a result, we only

examined the joint effects of both elements, which implies that the effect of the verbal framing of the questions about sharing on accurate-inaccurate discrimination could be due either to one of these elements or to both elements. Whereas the distinction between these two elements of the framing was not the focus of the present study, future research could explore these nuances further by dissecting the effect of the individual elements of the framing on accurate-inaccurate discrimination.

To summarize, the present study shows that question framing has distinct effects on how well people discriminate between accurate and inaccurate headlines in response to accuracy questions and sharing questions. While the framing of questions does not affect the participants' accurate-inaccurate discrimination in response to questions about accuracy, asking participants to respond to claims to the best of their knowledge instead of considering stories increases the accurate-inaccurate discrimination in responses to questions about the sharing on social media. At a theoretical level, the present results support the inattention-based account⁹. According to this account, people fail to discriminate between accurate and inaccurate information on social media because they do not pay sufficient attention to accuracy. Asking participants to respond to claims to the best of their knowledge, rather than to consider stories, can be interpreted as a contextual cue that shifts the focus toward accuracy, thereby enhancing accurate-inaccurate discrimination in sharing responses. At a methodological level, these findings indicate that while the phrasing of questions about accuracy may not affect the results, researchers should carefully attend to the questions about sharing on social media, as the verbal framing of such questions may significantly influence the participants' focus on accuracy. To align the decision-making process with how it occurs on social media, it may be advisable to straightforwardly ask participants, "Do you want to share this information?" or to design experiments that closely mirror the social-media interface [cf.^{21,57}].

Data availability

The data of the experiments are available at: <https://osf.io/xc4wf/>.

Received: 4 September 2024; Accepted: 18 November 2024

Published online: 22 November 2024

References

- Means, B., Toyama, Y., Murphy, R. & Baki, M. The effectiveness of online and blended learning: a meta-analysis of the empirical literature. *Teachers Coll. Record*. **115**, 1–47. <https://doi.org/10.1177/016146811311500307> (2013).
- Ward, A. F. Supernormal: How the internet is changing our memories and our minds. *Psychol. Inq.* **24**, 341–348. <https://doi.org/10.1080/1047840X.2013.850148> (2013).
- Jost, J. T. et al. How social media facilitates political protest: information, motivation, and social networks. *Political Psychol.* **39**, 85–118. <https://doi.org/10.1111/pops.12478> (2018).
- Yamashiro, J. K. & Roediger, H. L. Expanding cognition: a brief consideration of technological advances over the past 4000 years. *J. Appl. Res. Memory Cognition*. **8**, 15–19. <https://doi.org/10.1016/j.jarmac.2018.12.004> (2019).
- Lorenz-Spreen, P., Oswald, L., Lewandowsky, S. & Hertwig, R. A systematic review of worldwide causal and correlational evidence on digital media and democracy. *Nat. Hum. Behav.* **7**, 74–101. <https://doi.org/10.1038/s41562-022-01460-1> (2023).
- Storm, B. C. & Soares, J. S. Memory in the digital age. In *The Oxford Handbook of Human Memory* (eds. Kahana, M. J. & Wagner, A. D.). (Oxford University Press, in press).
- Lazer, D. M. J. et al. The science of fake news. *Science* **359**, 1094–1096. <https://doi.org/10.1126/science.aao2998> (2018).
- Kozyreva, A., Lewandowsky, S. & Hertwig, R. Citizens versus the internet: confronting digital challenges with cognitive tools. *Psychol. Sci. Public. Interest*. **21**, 103–156. <https://doi.org/10.1177/1529100620946707> (2020).
- Pennycook, G. & Rand, D. G. The psychology of fake news. *Trends Cogn. Sci.* **25**, 388–402. <https://doi.org/10.1016/j.tics.2021.02.007> (2021).
- Loftus, E. & Klemfuss, F. (ed Z, J.) Misinformation – past, present, and future. *Psychol. Crime. Law* **30** 312–318 <https://doi.org/10.1080/1068316X.2023.2219813> (2024).
- McKay, S. & Tenove, C. Disinformation as a threat to deliberative democracy. *Polit. Res. Q.* **74**, 703–717. <https://doi.org/10.1177/1065912920938143> (2021).
- Lewandowsky, S. et al. Misinformation and the epistemic integrity of democracy. *Curr. Opin. Psychol.* **54**, 101711. <https://doi.org/10.1016/j.copsyc.2023.101711> (2023).
- Schincariol, A. et al. Fake memories: a meta-analysis on the effect of fake news on the creation of false memories and false beliefs. *Memory Mind Media*. **3** <https://doi.org/10.1017/mem.2024.14> (2024). e17, Article e17.
- Swire-Thompson, B. & Lazer, D. Public health and online misinformation: challenges and recommendations. *Annu. Rev. Public Health*. **41**, 433–451. <https://doi.org/10.1146/annurev-publhealth-040119-094127> (2020).
- Mangiulli, I. et al. False memory and COVID-19: how people fall for fake news about COVID-19 in digital contexts. *Front. Psychol.* **13**, 972004. <https://doi.org/10.3389/fpsyg.2022.972004> (2022).
- Capraro, V. et al. The impact of generative artificial intelligence on socioeconomic inequalities and policy making. *PNAS Nexus*. **3** <https://doi.org/10.1093/pnasnexus/pgae191> (2024).
- 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. <https://doi.org/10.1177/0956797620939054> (2020).
- Pennycook, G. et al. Shifting attention to accuracy can reduce misinformation online. *Nature* **592**, 590–595. <https://doi.org/10.1038/s41586-021-03344-2> (2021).
- Pennycook, G. & Rand, D. G. Accuracy prompts are a replicable and generalizable approach for reducing the spread of misinformation. *Nat. Commun.* **13**, 2333. <https://doi.org/10.1038/s41467-022-30073-5> (2022).
- Arechar, A. A. et al. Understanding and combatting misinformation across 16 countries on six continents. *Nat. Hum. Behav.* **7**, 1502–1513. <https://doi.org/10.1038/s41562-023-01641-6> (2023).
- 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. <https://doi.org/10.1177/01461672221117691> (2023).
- Gawronski, B., Ng, N. L. & Luke, D. M. Truth sensitivity and partisan bias in responses to misinformation. *J. Exp. Psychol. Gen.* **152**, 2205–2236. <https://doi.org/10.1037/xge0001381> (2023).
- Epstein, Z. et al. Developing an accuracy-prompt toolkit to reduce COVID-19 misinformation online. *Harv. Kennedy School Misinformation Rev.* **2** <https://doi.org/10.37016/mr-2020-71> (2021).
- 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. <https://doi.org/10.1016/j.cognition.2018.06.011> (2019).
- Goffman, E. *Frame Analysis: An Essay on the Organization of Experience* (Harvard University Press, 1974).

26. Harper, F. M., Moy, D. & Konstan, J. A. Facts or friends? distinguishing informational and conversational questions in social Q&A sites. In *Proceedings of the 27th Conference on Human Factors in Computing Systems*. 759–768 (Association for Computing Machinery, (2009)). <https://doi.org/10.1145/1518701.1518819>
27. Wu, Q., Lee, C. S. & Goh, D. H. L. Understanding user-generated questions in social Q&A: a goal-framing approach. *J. Association Inform. Sci. Technol.* **74**, 990–1009. <https://doi.org/10.1002/asi.24770> (2023).
28. Bateson, G. A theory of play and fantasy. *Psychiatric Res. Rep.* **2**, 39–51 (1955).
29. Kardes, F. R., Kalyanaram, G., Chandrashekar, M. & Dornoff, R. J. Brand retrieval, consideration set composition, consumer choice, and the pioneering advantage. *J. Consum. Res.* **20**, 62–75. <https://doi.org/10.1086/209333> (1993).
30. Shapiro, S., MacInnis, D. J. & Heckler, S. E. The effects of incidental ad exposure on the formation of consideration sets. *J. Consum. Res.* **24**, 94–104. <https://doi.org/10.1086/209496> (1997).
31. Hauser, J. R. Consideration-set heuristics. *J. Bus. Res.* **67**, 1688–1699. <https://doi.org/10.1016/j.jbusres.2014.02.015> (2014).
32. Loftus, E. F. & Palmer, J. C. Reconstruction of automobile destruction: an example of the interaction between language and memory. *J. Verbal Learn. Verbal Behav.* **13**, 585–589. [https://doi.org/10.1016/S0022-5371\(74\)80011-3](https://doi.org/10.1016/S0022-5371(74)80011-3) (1974).
33. Loftus, E. F. & Zanni, G. Eyewitness testimony: the influence of the wording of a question. *Bull. Psychon. Soc.* **5**, 86–88. <https://doi.org/10.3758/BF03336715> (1975).
34. Fiedler, K. & Hertel, G. Content-related schemata versus verbal-framing effects in deductive reasoning. *Soc. Cogn.* **12**, 129–147. <https://doi.org/10.1521/soco.1994.12.2.129> (1994).
35. Capraro, V. & Vanzo, A. The power of moral words: loaded language generates framing effects in the extreme dictator game. *Judgm. Decis. Mak.* **14**, 309–317. <https://doi.org/10.1017/S1930297500004356> (2019).
36. Mieth, L., Buchner, A. & Bell, R. Moral labels increase cooperation and costly punishment in a prisoner's dilemma game with punishment option. *Sci. Rep.* **11**, 10221. <https://doi.org/10.1038/s41598-021-89675-6> (2021).
37. Capraro, V., Halpern, J. Y. & Perc, M. From outcome-based to language-based preferences. *J. Econ. Lit.* **62**, 115–154. <https://doi.org/10.1257/jel.2022.1613> (2024).
38. Pennycook, G., Cannon, T. D. & Rand, D. G. Prior exposure increases perceived accuracy of fake news. *J. Exp. Psychol. Gen.* **147**, 1865–1880. <https://doi.org/10.1037/xge0000465> (2018).
39. Pennycook, G., Bear, A., Collins, E. T. & Rand, D. G. The implied truth effect: attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. *Manage. Sci.* **66**, 4944–4957. <https://doi.org/10.1287/mnsc.2019.3478> (2020).
40. 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. <https://doi.org/10.1111/jopy.12476> (2020).
41. Sirlin, N., Epstein, Z., Arechar, A. A. & Rand, D. G. Digital literacy is associated with more discerning accuracy judgments but not sharing intentions. *Harvard Kennedy School Misinformation Review* **2**. (2021). <https://doi.org/10.37016/mr-2020-83>
42. Epstein, Z., Sirlin, N., Arechar, A., Pennycook, G. & Rand, D. The social media context interferes with truth discernment. *Sci. Adv.* **9**, eabo6169. <https://doi.org/10.1126/sciadv.abo6169> (2023).
43. Bowes, S. M. & Tasimi, A. Clarifying the relations between intellectual humility and pseudoscience beliefs, conspiratorial ideation, and susceptibility to fake news. *J. Res. Pers.* **98**, 104220. <https://doi.org/10.1016/j.jrp.2022.104220> (2022).
44. Lawson, M. A. & Kakkar, H. Of pandemics, politics, and personality: the role of conscientiousness and political ideology in the sharing of fake news. *J. Exp. Psychol. Gen.* **151**, 1154–1177. <https://doi.org/10.1037/xge0001120> (2022).
45. Nurse, M. S., Ross, R. M., Isler, O. & Van Rooy, D. Analytic thinking predicts accuracy ratings and willingness to share COVID-19 misinformation in Australia. *Mem. Cognit.* **50**, 425–434. <https://doi.org/10.3758/s13421-021-01219-5> (2022).
46. Arin, K. P., Mazrekaj, D. & Thum, M. Ability of detecting and willingness to share fake news. *Sci. Rep.* **13**, 7298. <https://doi.org/10.1038/s41598-023-34402-6> (2023).
47. Lawson, M. A., Anand, S. & Kakkar, H. Tribalism and tribulations: the social costs of not sharing fake news. *J. Exp. Psychol. Gen.* **152**, 611–631. <https://doi.org/10.1037/xge0001374> (2023).
48. MDR. MDR-WISSEN-Faktenchecks [MDR Knowledge fact checks]. (2023). <https://www.mdr.de/wissen/faktencheck/index.html>
49. Nadarevic, L., Reber, R., Helmecke, A. J. & Köse, D. Perceived truth of statements and simulated social media postings: an experimental investigation of source credibility, repeated exposure, and presentation format. *Cognitive Research: Principles and Implications* **5**. (2020). <https://doi.org/10.1186/s41235-020-00251-4>
50. Marchewka, A., Żurawski, Ł., Jednoróg, K. & Grabowska, A. The Nencki Affective Picture System (NAPS): introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behav. Res. Methods.* **46**, 596–610. <https://doi.org/10.3758/s13428-014-0379-1> (2014).
51. Kurdi, B., Lozano, S. & Banaji, M. R. Introducing the Open Affective standardized Image Set (OASIS). *Behav. Res. Methods.* **49**, 457–470. <https://doi.org/10.3758/s13428-016-0715-3> (2017).
52. Carretié, L., Tapia, M., López-Martín, S., Albert, J. & EmoMadrid An emotional pictures database for affect research. *Motivation Emot.* **43**, 929–939. <https://doi.org/10.1007/s11031-019-09780-y> (2019).
53. Cameron, K. A. et al. Patient knowledge and recall of health information following exposure to facts and myths message format variations. *Patient Educ. Couns.* **92**, 381–387. <https://doi.org/10.1016/j.pec.2013.06.017> (2013).
54. Snodgrass, J. G. & Corwin, J. Pragmatics of measuring recognition memory: applications to dementia and amnesia. *J. Exp. Psychol. Gen.* **117**, 34–50. <https://doi.org/10.1037/0096-3445.117.1.34> (1988).
55. Simons, D. J. The value of direct replication. *Perspect. Psychol. Sci.* **9**, 76–80. <https://doi.org/10.1177/1745691613514755> (2014).
56. Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**(2), 175–191. <https://doi.org/10.3758/BF03193146> (2007).
57. Bhardwaj, V., Martel, C. & Rand, D. G. Examining accuracy-prompt efficacy in combination with using colored borders to differentiate news and social content online. *Harvard Kennedy School (HKS) Misinformation Review* **4**. (2023). <https://doi.org/10.37016/mr-2020-113>

Author contributions

RB and AB developed the experiments. RB collected and analyzed the data with AB. RB wrote the first draft of the manuscript. AB revised the manuscript and gave approval for publication.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to R.B.

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.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024