Associative Inference Can Increase People's Susceptibility to Misinformation

Sian Lee, Haeseung Seo, Dongwon Lee, Aiping Xiong

The Pennsylvania State University, USA
{szl43, hxs378, dongwon, axx29}@psu.edu

Abstract

Associative inference is an adaptive, constructive process of memory that allows people to link related information to make novel connections. We conducted three online human-subjects experiments investigating participants' susceptibility to associatively inferred misinformation and its interaction with their cognitive ability and how news articles were presented. In each experiment, participants completed recognition and perceived accuracy rating tasks for the snippets of news articles in a tweet format across two phases. At Phase 1, participants viewed real news only. At Phase 2, participants viewed both real and fake news. Critically, we varied whether the fake news at Phase 2 was inferred from (i.e., associative inference), associated with (i.e., association only), or irrelevant to (i.e., control) the corresponding real news pairs at Phase 1. Both recognition and perceived accuracy results showed that participants in the associative inference condition were more susceptible to fake news than those in the other conditions. Furthermore, hashtags embedded within the tweets made the obtained effects evident only for the participants of higher cognitive ability. Our findings reveal that associative inference can be a basis for individuals' susceptibility to misinformation, especially for those of higher cognitive ability. We conclude by discussing the implications of our results for understanding and mitigating misinformation on social media platforms.

Introduction

The ubiquitousness of social media platforms and people's extended use of them for news consumption (Matsa and Shearer 2018) have partially helped bring a proliferation of misinformation (Allcott and Gentzkow 2017; Lazer et al. 2018). The extensive spread of fake news on social media platforms (Vosoughi, Roy, and Aral 2018) can have serious negative impacts on individuals and society, including election manipulation (Rosenberg 2016), spreading false treatments of COVID-19 (Perlow 2020), as well as reducing people's COVID-19 vaccination intention (Loomba et al. 2021).

Fake news especially becomes "successful" when it could cast doubts on truth or deceive news consumers. Detecting such fake news and preventing its harm, thus, have deep intellectual values as well as broad societal impacts. To help mitigate the negative effects of fake news on social media platforms, we seek to understand **cognitive processes** that increase people's susceptibility to *remember* and *believe in* misinformation.

Prior studies have examined various cognitive factors impacting people's *false memory* of fake news (Loftus 2005). Research has shown that repeated exposure can increase people's familiarity with misinformation and thus enhance their recollection afterward (Foster et al. 2012). Compared to individuals with higher cognitive ability, those with lower cognitive ability were more likely to remember fake news in agreement with their political stance (Murphy et al. 2019).

The impacts of repetition (Begg, Anas, and Farinacci 1992; Pillai and Fazio 2021) and cognitive ability (Bago, Rand, and Pennycook 2020) have also been evident in people's *misbelief* in fake news. People increase their belief in repeated fake headlines even if they were labeled as disputed (Pennycook, Cannon, and Rand 2018; Seo, Xiong, and Lee 2019). Individuals' cognitive ability levels were negatively correlated with the accuracy ratings of fake news but positively correlated with the accuracy ratings of real news (Pennycook and Rand 2019).

We examine the impact of **associative inference**, an adaptive, constructive memory process that allows people to link together acquired knowledge (Zeithamova and Preston 2010) or experience (Carpenter and Schacter 2017) that shares a common feature to make novel connections. When one person views *false inferences* based on associated real news articles (see Figure 1), feelings of familiarity comparable to or stronger than repetition could be created, increasing her or his susceptibility. Different from a piece of fake news that is completely fabricated, associatively inferred misinformation represents a dedicated attempt for deception by leveraging people's prior knowledge of existing real news.

Few studies have been conducted on people's susceptibility to associatively inferred fake news (Lee et al. 2020; Xiong et al. 2022). While Xiong et al. found that associative inference could increase participants' susceptibility to misinformation in the familiarity judgment (i.e., recognition), the existing studies do not provide conclusive results on participants' misbelief evaluation. We fill the gap in this work.

In addition, using hashtags, prefixed by the # symbol with a keyword or keywords, has become a common tagging method to help social media users associate tweet messages and share certain events or contexts (Bonilla and Rosa 2015; Bruns and Burgess 2011). Potentially, usage of two or more hashtags visually enhances the parts following # (see Figure 1), which may serve as explicit cues to accentuate the

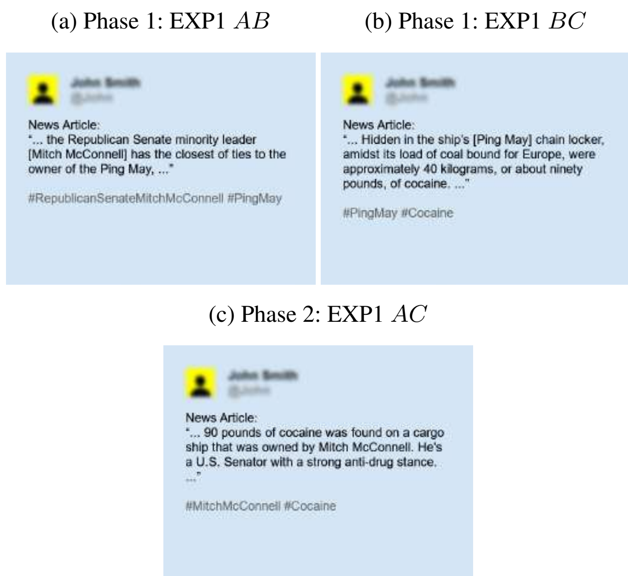


Figure 1: An associatively-inferred triad ($AB\&BC \rightarrow AC$) of Experiments 1 (EXP1). The top row shows an overlapping, real-news pair ($AB\&BC$) of Phase 1, each of which is associated with a piece of fake news (AC) of Phase 2 (the bottom row) through a keyword listed as a hashtag.

retrieval (Murnane and Phelps 1995) of associated pairs and facilitate the associative inference. Xiong et al. (2022) also compared participants' recognition and perceived accuracy rating of associatively inferred fake news with and without hashtags, but did not find significant differences. Since Xiong et al. only evaluated one tweet format, we further investigated the impact of hashtags using different presentation modes.

Considering relatively few studies have been conducted on people's susceptibility to associatively inferred fake news and the inconclusive results on the false belief measure and the impact of hashtags, we address three research questions (**RQs**).

- **RQ1**: Given a piece of fake news, do participants recognize it more and give higher accuracy rating when it is associatively inferred from real news than otherwise (e.g., association only or irrelevant)?
- **RQ2**: Do participants' cognitive ability levels impact the effect of associative inference on their susceptibility to misinformation?
- **RQ3**: Do hashtags' presentation modes have an influence on the effect of associative inference on participants' susceptibility to misinformation?

We conducted three online human-subject experiments. Each experiment utilized a two-phase setup. In Experiment 1, we examined whether people's recognition rate and perceived accuracy rating of fake news (Phase 2) depend on how the fake news is associated with real news that they viewed previously (Phase 1) across three conditions: associative inference (*a.Inf*), association only (*a.Only*), and control (*CON*). Between the two phases, participants' cognitive ability level was evaluated to determine whether it moderates the effects of associative inference. We evaluated the effect of embedded hashtags in Experiment 2 using the same setting as Experiment 1. Experiment 3 was designed to replicate the results of Experiments 1 and 2.

In Experiments 1 and 3a, participants in the *a.Inf* condition, regardless of their cognitive ability levels, showed higher recognition rate and gave higher accuracy rating for the same fake news at Phase 2 (**RQ1** & **RQ2**). However, when keywords were presented as hashtags and embedded within tweets (Experiments 2 and 3b), the effect of associative inference was only evident for participants of higher cognitive ability (**RQ3**). Based on our findings, we end the paper by discussing theoretical and practical implications for understanding and mitigating misinformation on social media platforms.

Related Work

Misinformation has been defined as an umbrella term to include any information spreading on social media that is false or inaccurate (Wu et al. 2019). Closely related, Lazer et al. (2018) defined fake news as false or fabricated information written and published to mimic legitimate news media content in form. Given the similar, broad definitions of misinformation and fake news, we use those two terms interchangeably (Lewandowsky and Van Der Linden 2021).

Associative Inference

The associative inference is an adaptive process that allows people to link together related information acquired to make novel connections that they have not directly experienced (Zeithamova and Preston 2010). For example, Carpenter and Schacter (2017) conducted four experiments showing that if participants learned direct associations between two items (AB, e.g., a person [A] and a toy [B] in a room) and then learned direct associations that include one member of the previously studied pairs (BC, e.g., the toy [B] with a different person [C] in a room). The participants were susceptible to misattribute AB event with BC event or vice versa, suggesting the impact of associative inference (i.e., AC).

Lee et al. (2020) conducted a preliminary online study investigating the effect of associative inference on individuals' susceptibility to fake news. Across two phases, they examined participants' recognition and perceived accuracy of news (Phase 2) as a function of how those pieces of news are associated with real news that participants viewed before (Phase 1). The association was varied in three between-subjects conditions (i.e., associative inference, with association, and control) across two phases. Participants' cognitive ability was also measured. Lee et al. only obtained non-significant results, which showed that participants, especially those of higher cognitive ability, tended to give higher perceived accuracy ratings to fake news with associative inference than that without associative inference.

Using the same two-phase procedure, Xiong et al. (2022) conducted two online human-subjects experiments examining the effect of associative inference with a within-subjects design. Critically, each participant viewed three types of news articles (i.e., real, fake, and fake with associative inference) at Phase 2. In both experiments, the participants recognized more the fake news with associative inference than that without, indicating their susceptibility to associatively inferred fake news. However, the participants gave the lowest accuracy rating for fake news with associative inference than the other types. The authors further analyzed factors that the participants considered for the accuracy ratings of the fake news with associative inference. They proposed that one possible reason could be the *within-subjects* design: Each participant saw fake news with and without associative inference, thus her or his awareness of the associative inference was increased. While the increased awareness could have enhanced the relatively intentional control and effortful process of accuracy rating task (e.g., semantic judgment), it had limited impacts on the relatively intuitive and automatic recognition task (e.g., familiarity judgment).

Cognitive Ability

Prior studies evaluated participant's critical thinking ability with the Cognitive Reflection Test (CRT) (Frederick 2005) and found that participants who scored higher on CRT were also better at discerning fake and real news (Bago, Rand, and Pennycook 2020). The Wordsum has been used to assess people's cognitive ability in various fields of social science (Malhotra, Krosnick, and Haertel 2007). Recently in the fake news study, Murphy et al. (2019) evaluated participants' cognitive ability using the Wordsum (Wechsler 2008) and obtained similar findings as using the CRT.

Compared to the CRT, substantially fewer MTurk workers have been previously exposed to the CRT-2 questions (Thomson and Oppenheimer 2016). Lee et al. (2020) and Xiong et al. (2022) evaluated participants' cognitive ability with the CRT-2 and the Wordsum. Both studies revealed the impact of participants' cognitive ability on the perceived accuracy ratings, but in the opposite direction. While participants of higher cognitive ability gave lower perceived accuracy ratings for associatively inferred fake news in Xiong et al., Lee et al. only obtained a nonsignificant trend that participants of higher cognitive ability in the associative inference condition tended to give higher perceived accuracy ratings.

Effect of Hashtags

Hashtags are user-created keywords starting with the prefix pound symbol, #, to annotate, categorize and contextualize posts on social media (e.g., Twitter) (Huang, Thornton, and Efthimiadis 2010). Previous studies have investigated the impact of hashtags on people's susceptibility of misinformation on social media but revealed mixed results. Rho and Mazmanian (2019) examined the impact of political hashtags in posts on Meta via a human-subject online study. They found that participants who saw political posts with hashtags believed more compared to participants who viewed the same posts without hashtags. On the contrary, Xiong et al. (2022) presented keywords of the snippets of real and fake news as hashtags in tweets. The authors compared participants' recognition rates and accuracy ratings of fake news with hashtags and without hashtags conditions, but did not find any significant differences. While Rho and Mazmanian embedded bold hashtags in the title of news posts, Xiong et al. only presented the gray hashtags at the bottom of the tweets.

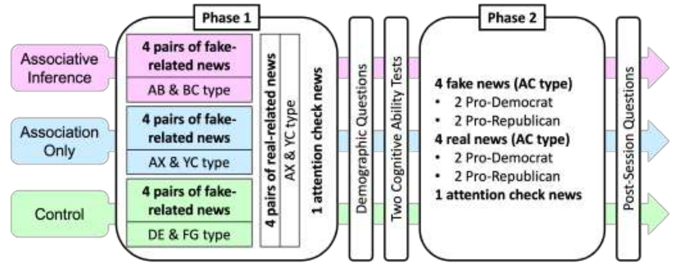


Figure 2: A flow chart of all experiments. "Phase 1" and "Phase 2" boxes show different types of news at each phase. Fake-related, real news at Phase 1 were different among the three conditions (*associative inference*, *association only*, *control*). Real-related, real news at Phase 1 and all news at Phase 2 were the same across the conditions. Participants answered demographic questions and completed cognitive ability tests between the two phases. We also asked post-session questions at the end.

Present Study

In the following, we present three online human-subject experiments investigating whether participants' recognition rate and perceived accuracy rating of fake news (Phase 2) depend on how the fake news is associated with real news that they viewed previously (Phase 1). In each experiment, there were three between-subjects conditions: associative inference, association only, and control (see Figure 2). Experiment 1 used a tweet format in which gray hashtags were placed at the bottom (see Figure 1). With the effect of associative inference evident in both recognition and perceived accuracy measures in Experiment 1, Experiment 2 applied another format in which blue-highlighted hashtags were embedded in tweet messages. Surprisingly, we did not obtain the effect of associative inference for either measure. Experiment 3 provided a further investigation of the impact of hashtag presentation on associative inference. We replicated the results of Experiments 1 and 2 in Experiments 3a and 3b, respectively.

Experiment 1

Experiment 1 was conducted to investigate **RQ1** and **RQ2**. We anticipated that participants in the associative inference condition would recognize more and give higher accuracy ratings for fake news than those in the other conditions. Awareness of associative inference was expected to be higher for participants with higher cognitive-ability level (Xiong et al. 2022), which should result in the effect of associative inference being more evident for participants with lower cognitive-ability level.

Method

Participants. For the perceived accuracy rating of fake news, a small effect size of the two-way interaction of *cognitive ability* × *condition*, $\eta_p^2 = .019$, was reported (Lee et al. 2020). Power analysis using G*Power 3.1 (Faul et al. 2007) suggested $n = 501$ participants to detect a small effect size

Table 1: Participants' demographic information in each experiment. Number in the bracket on the top row indicates the number of participants.

Item	Options	EXP1 (686)	EXP2 (718)	EXP3a (788)	EXP3b (740)
Gender	Male	52.8%	52.0%	51.4%	46.0%
	Female	47.1%	47.6%	48.2%	53.4%
	Other	0%	0.3%	0.3%	0.5%
	Prefer not to answer	0.1%	0.1%	0.1%	0.1%
Age	18-27	17.8%	17.3%	19.5%	19.7%
	28-37	42.1%	42.8%	43.0%	40.1%
	38-47	18.7%	22.1%	20.9%	20.8%
	48-57	12.1%	13.1%	9.5%	11.1%
	58 or older	9.3%	4.7%	6.7%	8.2%
	Prefer not to answer	0%	0%	0.3%	0%
Education	No high school	0.3%	0.3%	0.4%	0.3%
	High School	8.3%	6.1%	8.5%	8.2%
	College/Bachelor	72.9%	68.8%	68.9%	69.3%
	Professional degree/Masters/Ph.D.	18.5%	24.8%	22.0%	22.2%
	Prefer not to answer	0%	0%	0.3%	0%
Time on Soc. Media (Per Day)	Less than 1 hour	46.9%	31.3%	37.2%	43.1%
	Between 1 to 4 hours	44.6%	57.9%	55.3%	49.2%
	Longer than 4 hours	8.2%	10.4%	7.4%	7.4%
	Prefer not to answer	0.3%	0.3%	0.1%	0.3%
Political Stance	Liberal	48.3%	38.7%	44.0%	49.5%
	Moderate	24.5%	21.5%	23.9%	24.1%
	Conservative	27.3%	39.8%	32.1%	26.4%
	Prefer not to answer	0%	0%	0%	0%

(Cohen's f = 0.14) of interaction of cognitive ability (*lower*, *higher*) $\times$ condition (*a.Inf*, *a.Only*, *CON*) with a power of 0.8 [analysis of variances (ANOVA) test, α = .05].

Considering the relatively less control of online experiments on Amazon Mechanical Turk (MTurk), we recruited $1,200$ participants to make the design more powerful. Each MTurk worker were only allowed to participate in one of the experiments. The human intelligent tasks (HITs) were posted with the restrictions to workers who (1) were at least 18 years old; (2) previously completed more than 100 HITs and had a HIT approval rate of at least 95%; and (3) were located in the United States. Participants were excluded if they failed one of the two attention checks (Hauser and Schwarz 2016).[1] About 64.1% of the participants passed all attention check questions and completed the study. The median completion time of the experiment was about 12 min. We paid $1.50 for participants who completed the whole study in this and the rest experiments.

After removing duplicate responses and participants who selected "*Prefer not to answer*",[2] we included 686 participants' results for the data analysis at Phase 2 and 644 participants' answers were used for the data analysis at Phase 1. Participants' demographics are shown in Table 1.

Materials. A total of 42 different news were presented. Each news was a snippet from one real or fake news article. Thirty-eight pieces of the news were based on real news articles reported from major news media, including *washingtonpost.com*, *usatoday.com*, and *foxnews.com*. The remaining four pieces of news were based on the fake news that were debunked by the fact-checking website *snopes.com*.

Each news was designated by two letters (e.g., AB, YC or DE). Each letter, such as A, Y, and D, refers to a unique keyword (e.g., one public figure, an event, or an entity, see Figure 1 as an example) in the news. Each keyword was also listed after the pound symbols as hashtags in the news. To minimize the impact of source (Visentin, Pizzi, and Pichierri 2019), we applied an identical Twitter user ID and a blurred profile image for each piece of news. Stimuli in all experiments are shown in the Supplementary Figures S1-S24 [3].

Due to the main interest in associative inference, we started the news selection from the four pieces of fake news (i.e., AC) at Phase 2. Across the three conditions, the fake-news triads between the two phases are as follows.

- *a.Inf* ($AB\&BC \rightarrow AC$): the four pairs of fake-related, real news at Phase 1 were in an AB & BC type such that two keywords of each news overlapped through one common keyword B (e.g., AB: Mitch McConnell [A], Ping May [B]; BC: Ping May [B], cocaine [C]), affording an associative inference for the fake news at Phase 2 (e.g., Mitch McConnell [A], cocaine [C], see Figure 1).
- *a.Only* ($AX\&YC \rightarrow AC$): the four pairs of fake-related, real news at Phase 1 were in an AX & YC type, each of which had no overlapped keyword (e.g., Mitch McConnell [A], Climate Change [X], Authorities [Y], and Cocaine [C]). Thus, each piece of real news was only associated with the fake news at Phase 2.
- *CON* ($DE\&FG \rightarrow AC$): the four pairs of fake-related, real news at Phase 1 were in a DE & FG type. Each pair have neither an overlapped keyword nor was associated with the fake news at Phase 2.

The real-news triads were constructed in the same way as the fake-news triads in the *a.Only* condition. Since news at Phase 2 was politically related, the news stance was also controlled at each veracity level: half was pro-Republican and the other was pro-Democrat.

Procedure. After participants accepted the HIT on MTurk, they were directed to the online survey designed on Qualtrics. Participants who agreed with the consent form were randomly assigned to one of three conditions (*a.Inf*, *a.Only*, *CON*), the procedure of which was identical.

Phase 1. Eight pairs of real news were presented in a randomized order in each condition. Half of the pairs were related to real news at Phase 2 (i.e., real-related) and the other half were related to fake news at Phase 2 (i.e., fake-related). We also counterbalanced the presentation order of two pieces of news in each pair.

After participants viewed each piece of news, they were asked to answer, "*Have you ever seen this before?*" with four options (*Yes*; *Unsure*; *No*; and *Prefer not to answer*) at first. Such recognition task was used to reflect whether participants have been exposed to the news before our study. Then participants answered, "*How would you rate the accuracy of this news article?*" using a 5-point scale (1:*Very inaccurate*, 5: *Very accurate*). For the perceived accuracy ratings, we also provided "*Prefer not to answer*" as the 6th option.

[1] We provided instructions asking participants to choose one specified answer to pass the attention check for both recognition and accuracy rating tasks at each phase.

[2] The removed duplicate responses were from the same IP addresses. We applied the same criteria for all experiments. Due to space limit, we present the exclusion details of this and the following experiments in Table S1 of the Supplementary Material.

[3] http://tiny.cc/Leeetal2023supp

The perceived accuracy rating thus evaluated whether participants were able to identify fake news.

Between Two Phases. After Phase 1, participants filled out their demographic information, such as age, gender, and educational background. We also assessed participants' cognitive ability using the CRT-2 (Thomson and Oppenheimer 2016) and the Wordsum (Wechsler 2008). CRT-2 measured participants' tendency to override an incorrect "gut" response with four questions. For example, we asked participants, "*If you are running a race and you pass the person in the second place, what place are you in*?" While the intuitive answer is "*first*", the correct answer is "*second*." The Wordsum measured participants' intelligence scale of vocabulary with ten items. Within the test, we showed participants different words in capital letters (e.g., SPACE). Then, we asked participants to choose one word that comes closest to the meaning of the word in capital letters from five options (e.g., captain, school, noon, board, room). Options of "*do not know*" and "*prefer not to answer*" were also provided for both tests.

Phase 2. After the demographic questionnaire and cognitive ability tests, Phase 2 started, in which we presented eight pieces of news (half fake and half real) in a randomized order. Same as Phase 1, participants completed the recognition and perceived accuracy tasks for each piece of news.

Post-session Questionnaire. After Phase 2, we asked participants' interests in politics and their political ideology. Additionally, we randomly presented one piece of the four fake news at Phase 2 with the perceived accuracy rating that the participant gave and asked the participant to select all factors affecting his/her decision. Seven options (*Source*; *Writing style*; *Content*; *Web search results*[4]; *News presented in Stage 1*; *News that I saw before this study*; *Opinions from others*) were provided which were presented in a randomized order. We also provided "*Other (with text box)*" and "*Prefer not to answer*" as two extra options in the end. To further understand the effect of associative inference, for any participant who chose "News presented in Stage 1" (Stage 1 refers to Phase 1 in the paper), we asked the participant to explain how the "News presented in Stage 1" had influence on his/her decision with an open-ended question.

Data analysis

We used the same method to analyze the recognition and perceived accuracy results at Phases 1 and 2 for this and the following experiments.[5] In literature, participants were categorized as of high cognitive ability with eight or more items answered correctly for the Wordsum (Murphy et al. 2019). Participants, on average, got 56.2% of the four items answered correctly for the CRT-2 (Thomson and Oppenheimer 2016). Thus, we chose *ten* as a cutoff to categorize participants of a higher or lower cognitive-ability-test score for a total of 14 questions of the Wordsum and the CRT-2. We categorized 419 participants who gave ten or more correct answers as having a *higher* cognitive-ability-test score and the remaining 267 participants as having a *lower* score. There were 231, 217, and 238 participants in the *a.Inf*, *a.Only*, and *CON* conditions, respectively. The demographics were similar across conditions.

Phase 1. Each pair of real news was related to one piece of real or fake news in Phase 2. Thus, we divided news in Phase 1 into two types, *real-related* and *fake-related*. Recognition rates and perceived accuracy ratings were entered into 2 (*news type*: real-related, fake-related) $\times$ 3 (*condition*: *a.Inf*, *a.Only*, *CON*) $\times$ 2 (*cognitive-ability-test score*: lower, higher) mixed ANOVAs with a significance level of .05, respectively. Post-hoc tests with Bonferroni correction were performed, testing all pairwise comparisons with corrected p values for possible inflation. We report the effect size using η_p^2 (Keppel 1991), which is reported by SPSS.

Phase 2. Recognition rates and perceived accuracy ratings were entered into 2 (*news veracity*: real, fake) $\times$ 3 (*condition*: *a.Inf*, *a.Only*, *CON*) $\times$ 2 (*cognitive-ability-test score*: lower, higher) mixed ANOVAs with a significance level of .05. Post-hoc analyses were conducted in the similar way to the results of Phase 1. Although we obtained mean values on the perceived accuracy ratings of all news items for each participant in each condition at Phase 2, possible distribution existed among the multiple ratings of the same news item. Thus, we further performed Liner Mixed Effect Regression (LMER) on perceived accuracy ratings at Phase 2 for each experiment with the lme4 package in R (Bates et al. 2011). LMER allows controlling for the random effect for participants without data aggregation (Brauer and Curtin 2018). We intentionally used the same model as ANOVAs for LMERs except that we included random intercepts for participants and news trials. Thus, we could directly compare the results of ANOVAs and LMERs. Following ANOVA results, we report LMER results in square bracket without the effect size (Pennycook et al. 2021), since it is under debate for a single, agreed standardized effect size of LMER (Rights and Sterba 2018). We report the degree of freedom with Satterthwaite approximation. In general, ANOVAs and LMERs show similar results. Tables S25 - S28 show the complete results of LMER.

Thematic Analysis. We did a thematic analysis (Braun and Clarke 2006) for the answers of open-ended questions. The first two co-authors of this article and two undergraduate students performed the thematic analysis independently at first. Then they discussed the results and finalized the thematic analysis together. Supplementary Table S3 shows the percentage of each factor affecting participants' perceived accuracy rating decision.

Results

The means for the conditions are shown in Figure 3 and the ANOVA values are shown in Table 2 [6]

[4]Since there was no strict way to block participants from web search during an online survey, we added "web search results" as one option to understand how many participants used web search during the study. We asked the question at the end of our survey, 6.85% of the participants selected the option.

[5]All independent variables and dependent measures in our experiments are described in Supplementary Table S4.

[6]We provide the complete post-hoc analysis results in the Supplementary Material.

Table 2: Summary table for the statistical result in Experiments 1 and 2. Note. df=degrees of freedom. Bold font indicates statistical significance (*p*< .05).

Phase	Effect	Experiment 1								Experiment 2							
		Recognition				Accuracy				Recognition				Accuracy			
		df	F	*p*	η_p^2	*df*	F	*p*	η_p^2	*df*	F	*p*	η_p^2	*df*	F	*p*	η_p^2
P1	news type	**1, 638**	**34.18**	**0.000**	**0.051**	1, 638	0.56	0.453	0.001	**1, 647**	**16.17**	**0.000**	**0.024**	1, 647	1.91	0.167	0.003
	news type * condition	**2, 638**	**4.37**	**0.013**	**0.013**	**2, 638**	**14.27**	**0.000**	**0.043**	2, 647	1.95	0.143	0.006	**2, 647**	**8.58**	**0.000**	**0.026**
	news type * cognitive ability	**1, 638**	**7.16**	**0.008**	**0.011**	1, 638	1.59	0.208	0.002	**1, 647**	**5.10**	**0.024**	**0.008**	1, 647	0.07	0.792	0.000
	news type * condition * cognitive ability	2, 638	0.29	0.750	0.001	2, 638	0.95	0.388	0.003	2, 647	0.06	0.945	0.000	2, 647	1.41	0.246	0.004
	condition	2, 638	2.14	0.118	0.007	2, 638	1.96	0.141	0.006	2, 647	0.51	0.600	0.002	2, 647	1.81	0.164	0.006
	cognitive ability	**1, 638**	**54.01**	**0.000**	**0.078**	1, 638	2.44	0.119	0.004	**1, 647**	**95.99**	**0.000**	**0.129**	**1, 647**	**17.46**	**0.000**	**0.026**
	condition * cognitive ability	2, 638	0.23	0.797	0.001	2, 638	1.50	0.223	0.005	2, 647	0.65	0.523	0.002	2, 647	0.10	0.903	0.000
P2	news veracity	**1, 680**	**5.08**	**0.025**	**0.007**	**1, 680**	**524.71**	**0.000**	**0.436**	**1 ,712**	**10.28**	**0.001**	**0.014**	**1 ,712**	**370.29**	**0.000**	**0.342**
	news veracity * condition	**2, 680**	**3.35**	**0.036**	**0.010**	**2, 680**	**5.98**	**0.003**	**0.017**	2 ,712	1.07	0.342	0.003	2 ,712	1.49	0.227	0.004
	news veracity * cognitive ability	1, 680	0.57	0.451	0.001	**1, 680**	**43.38**	**0.000**	**0.06**	1 ,712	2.24	0.135	0.003	**1 ,712**	**119.87**	**0.000**	**0.144**
	news veracity * condition * cognitive ability	2, 680	0.30	0.738	0.001	2, 680	0.90	0.406	0.003	2 ,712	0.02	0.982	0.000	2 ,712	2.12	0.121	0.006
	condition	2, 680	0.36	0.698	0.001	2, 680	1.19	0.304	0.003	2 ,712	0.93	0.396	0.003	2 ,712	2.18	0.113	0.006
	cognitive ability	**1, 680**	**34.89**	**0.000**	**0.049**	**1, 680**	**29.14**	**0.000**	**0.041**	**1 ,712**	**114.67**	**0.000**	**0.139**	**1 ,712**	**138.03**	**0.000**	**0.162**
	condition * cognitive ability	2, 680	0.92	0.399	0.003	2, 680	1.00	0.369	0.003	2 ,712	0.25	0.780	0.001	2 ,712	0.35	0.704	0.001

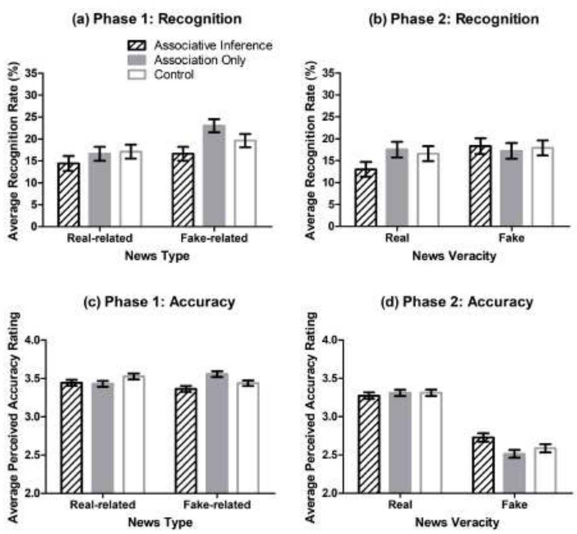


Figure 3: Average recognition rate (top row) and perceived accuracy rating (bottom row) with standard errors across the three conditions for each news type at Phase 1 (left column) and each news veracity at Phase 2 (right column) of Experiment 1.

Phase 1. Figures 3a and 3c depict the results of Phase 1, which serve as baseline to understand possible differences across the conditions *before* our key manipulation. Interestingly, participants recognized more fake-related news (19.7%) than real-related news (16.0%), suggesting that fake news might have been based on real news that is more familiar to news consumers. Such a higher recognition rate of the fake-related news was more evident for participants of a higher cognitive-ability-test score (fake-related vs. real-related: 14.3% vs 8.9%) than those of a lower cognitive-ability-test score (fake-related vs. real-related: 25.2% vs. 23.1%).

The two-way interaction of *news type* (2: real-related, fake-related) × *condition* (3: *a.Inf*, *a.Only*, *CON*) was significant for both recognition rates and the perceived accuracy ratings. For the real-related news, neither recognition rate nor perceived accuracy measure showed any statistically significant difference across the conditions (see Figures 3a and 3c). In contrast, for the fake-related news, the differences across conditions were evident for both recognition rates and perceived accuracy ratings. Specifically, participants in the *a.Only* condition recognized more the fake-related news than those in the *a.Inf* condition ($p_{adj} = .004$), and gave higher accuracy rating in general ($p_{adjs} \leq .023$). Such results are essential, which indicate that any higher susceptibility to associatively inferred fake news at Phase 2 is not due to participants' higher familiarity or more belief in the real-related news.

Phase 2. We presented the same fake and real news to the participants in each condition. Thus, any differences across the conditions would indicate the impacts of key manipulations across Phases 1 and 2.

Participants recognized more fake news (17.8%) than real news (15.7%), but such a main effect of *news veracity* was *condition* dependent ($F_{(2,680)} = 3.35, p = .036, \eta_p^2 = .010$). Post-hoc analysis revealed that only participants in the *a.Inf* condition recognized more the fake news (18.3%) than the real news (13.0%, $F_{(1,229)} = 9.36, p = .002, \eta_p^2 = .039$). Thus, the effect of associative inference was evident for the familiarity judgment, which is consistent with prior work (Xiong et al. 2022).

Turning to the accuracy measure (See Figure 3d), participants gave higher ratings for the real news (3.30) than for the fake news (2.61), indicating that they could differentiate the fake news from the real ones. Same as the recognition measure, the two-way interaction of *news veracity* × *condition* was also significant ($F_{(2,680)} = 5.98, p = .003, \eta_p^2 = .017$, [$F_{(2,4790)} = 7.17, p = .001$]). Across the three conditions, participants' accuracy ratings were similar for the real news, but showed differences for the fake news (*a.Inf* vs. *a.Only* vs. *CON*: 2.73 vs. 2.51 vs. 2.59, $F_{(2,680)} = 4.08, p = .017, \eta_p^2 = .012$). Post-hoc analysis revealed that, for the same fake news, the average accuracy rating in the *a.Inf* condition was significantly higher than that of the *a.Only* condition ($p_{adj} = .005$), and showed a non-significant trend to be higher than that of the *CON* condition ($p_{adj} = .059$).

Finding 1: *Using a between-subject design, we found the effect of associative inference for both recognition and perceived accuracy measures (RQ1).*

The main effect of cognitive-ability-test score was significant for both recognition rates and perceived accuracy ratings. Participants of a lower cognitive-ability-test score recognized more news (22.1%) and gave higher accuracy ratings (3.08) than those of a higher cognitive-ability-test score (recognition: 11.5%; accuracy: 2.83). The two-way interaction of *news veracity* × *cognitive-ability-test score* was also significant for the perceived accuracy measure. Post-hoc analysis revealed that participants gave similar accuracy ratings for the real news regardless of their cognitive-ability levels. In contrast, for the fake news, participants of a higher cognitive-ability-test score gave lower accuracy rating (2.39) than participants of a lower cognitive-ability-test score (2.83). Thus, participants with higher cognitive-ability level were less susceptible to misinformation than those with lower cognitive-ability level in general (Pennycook and Rand 2019). However, participants' cognitive-ability level showed no impact on the obtained effects of associative inference.

Thematic Analysis. Among the 686 participants, 76 (11.1%) of them indicated that their accuracy ratings were impacted by "News presented in Stage 1." We further disregarded 24 meaningless answers for the open-ended question, such as "yes" or "nice." For the remaining 52 meaningful answers, most of them were in the *a.Inf* condition (73.1%), but only 11.5% and 15.4% were in the *a.Only* and *CON* conditions, respectively. Such results highlight that more participants in the *a.Inf* condition explicitly made their accuracy decision of the fake news based on "News presented in Stage 1" than those in the other conditions.

We identified three major themes: 1) *Association between Two Phases:* Among the 52 answers, 46.2% (24) of the participants indicated that when they gave accuracy ratings for fake news in Phase 2, they noticed its connection with news headlines in Phase 1. For example, one participant answered, *"Because it made me think that if it were repeated more than once there was some possibility it was accurate."* 2) *Gist-based Recall:* 28.8% (15) of the participants made the accuracy rating decision based on their prior knowledge of or belief in the news. For instance, one participant explained, *"I remember seeing it on a Facebook post, and this is something that I believe can be very true."* 3) *Verbatim Recall:* About 23.1% (12) participants noticed the gap between the news in two phases. They recalled the details of news in Phase 1 and detected the distorted or exaggerated parts in Phase 2. For example, one participant answered, *"There were a few modified stage 1 messages that were reworded or combined for stage 2..."*

We also examined participants' cognitive-ability-test score distribution in each theme. For the theme of association between two phases, 67% of the participants were of a higher cognitive-ability-test score, showing a similar distribution that includes all participants. For the verbatim recall theme, 92% of the participants were of a higher cognitive-ability-test score, but the percentage was only 40% for the theme of gist-based recall (schema).

<u>**Finding 2:**</u> *Out of our expectation, individuals' cognitive-ability level showed no impact on the effects of associative inference (RQ2). Responses to the post-session questions revealed that few participants were aware of the association or inference.*

Experiment 2

Hashtags are typically highlighted in blue and embedded within the tweet messages. Thus, we recruited extra participants in Experiment 2 and evaluated whether the effects of associative inference obtained in Experiment 1 (**RQ1** & **RQ2**) can be generalized to or accentuated with blue-highlighted hashtags embedded in tweets (**RQ3**).

Participants, Stimuli, and Procedure

We published 1,200 tasks on MTurk. Applying the same criteria as Experiment 1, we included 718 participants' results for the data analysis in Phase 2. There were 255 participants who were categorized as with a *higher* cognitive-ability-test score and the remaining 463 participants were categorized as with a *lower* cognitive-ability-test score. The median completion time was about 14 min. We analyzed 653 participants' results in Phase 1.

We conducted Experiment 2 using the same setting as Experiment 1 except as noted. In Experiment 2, we presented hashtags in blue and embedded the hashtags within the tweets to assess whether such a presentation would impact the effect of associative inference and participants' susceptibility to fake news. Since the meaningful answer rate was less than 10%, we did not include the open-ended question for this and the following experiments.

Results

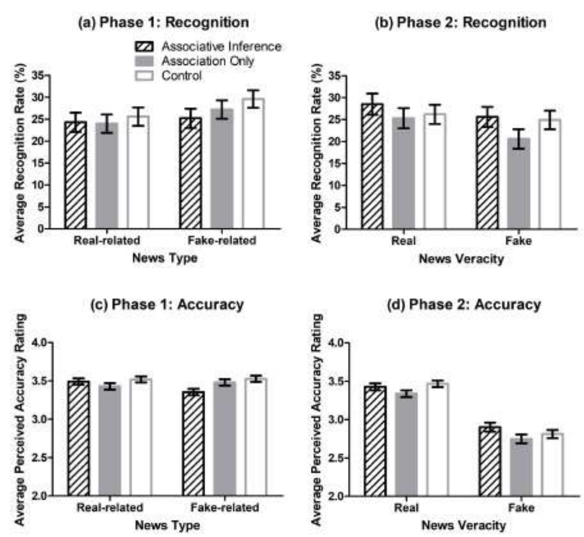


Figure 4: Average recognition rate (top row) and perceived accuracy rating (bottom row) with standard errors across the three conditions for each news type at Phase 1 (left column) and each news veracity at Phase 2 (right column) of Experiment 2.

Phase 1. The results are shown in Figures 4a and 4c, which replicate the main findings of Experiment 1. As expected, participants recognized more the fake-related news (27.3%) than the real-related news (24.7%), but such a recognition gap did not show any significant differences across conditions. For the perceived accuracy rating, only the two-way interaction of *news type* × *condition* was significant. Same as Experiment 1, post-hoc analysis showed that differences across the conditions were only significant for the fake-related news but not for the real-related news. Specifically, participants in the *a.Inf* condition (3.35) gave lower perceived accuracy rating than those in the *CON* condition (3.53, $p_{adj} = .011$). The main effect of cognitive-ability-test score was significant for both measures. Participants of a lower cognitive-ability-test score recognized more news and gave higher accuracy ratings than those of a higher score in general. The two-way interaction of *cognitive-ability-test score* × *news type* was only significant for the recognition measure. Specifically, a non-significant trend of higher recognition rate of fake-related news than real-related news was more evident for participants of a higher cognitive-ability-test score than those of a lower score.

Phase 2. As illustrated in Figures 4b and 4d, the main effect of *news veracity* was significant for both recognition and perceived accuracy measures. The same as Experiment 1, participants could differentiate the real news (3.41) from the fake news (2.82). In contrast, they recognized more the real news (26.7%) than the fake news (23.7%) in Experiment 2. The main effect of *cognitive-ability-test score* was also significant for both measures. Participants of a lower cognitive-ability-test score recognized more (38.2%) and gave higher accuracy ratings (3.41) than those of a higher score (recognition: 12.2%; accuracy: 2.82). Moreover, the two-way interaction of *news veracity* × *cognitive-ability-test score* was significant for the accuracy rating: Participants of a higher cognitive-ability-test score could distinguish the real news (3.29) from the fake news (2.36), but participants of a lower score rated news as real in general (real: 3.54; fake: 3.28).

Different from Experiment 1, *condition* did not show any significant impacts on either measure. One possible explanation is offered by the encoding of associated pairs in human memory (Craik et al. 1996). Prior works have supported the notion that the effectiveness of encoding can be determined by the presented materials (Brown and Craik 2000). Participants might have primarily attended to the blue-highlighted hashtags in tweets, but encoded the rest parts of the tweets shallowly.

Highlighting relevant parts in the texts has often been considered an effective encoding process tool to aid memory (Lorch 1989). Yet, preexisting highlighting could interfere with reading comprehension, resulting in decreased performance (Hunt and Lamb 2001). Peterson (1992) examined college students' learning of a history chapter. Results showed that students who highlighted while reading performed worse on the tests of comprehension wherein they needed to make inferences that required connecting different ideas across the text. Instead of serving an encoding or review function, embedding blue-highlighted hashtags might have been counterproductive for inferential tasks.

Finding 3: *After embedding blue-highlighted hashtags within tweet messages, we obtained the results of Phase 1 similar to those of Experiment 1. Yet, the non-significant interactions with condition at Phase 2 indicate that the blue-highlighted hashtags embedded within tweet messages might have reduced the associative inference (RQ3).*

Experiment 3

We conducted two sub-experiments to further understand the impacts of hashtags presentation on associative inference (**RQ3**). Experiment 3a was identical to Experiment 1 except that the hashtags at the bottom of the tweets were highlighted in blue (see Figure 5a). We evaluated the impact of blue-highlighted hashtags on the effect of associative inference. Experiment 3b was the same as Experiment 2 except that the hashtags were presented at the bottom of the tweets again (see Figure 5b), which was to test whether presenting hashtags again was critical to obtain the effect of associative inference.

Participants, Stimuli, and Procedure

We published 2,400 HITs on MTurk. We included 1,528 participants' results for the data analysis at Phase 2. Among them, 51.6% (788) were in Experiment 3a and 48.4% (740) were in Experiment 3b. The demographics of each experiment are shown in Table 1). We analyzed 1,427 participants' results of Phase 1, with 51.3% (732) in Experiment 3a and 48.7% (695) in Experiment 3b. The median completion time was about 13 min for both experiments.

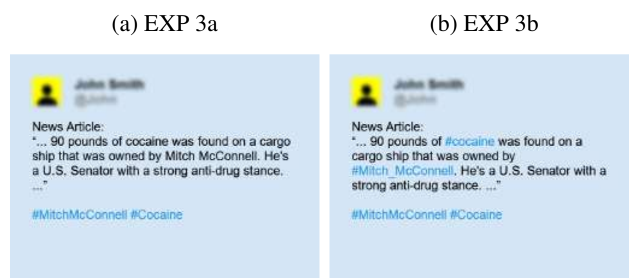


Figure 5: An example of news stimulus in Experiments 3a (left panel) and 3b (right panel).

Results of Experiments 3a & 3b

We analyzed the results of Experiments 3a and 3b using the same methods as the prior experiments. In Experiment 3a, 423 and 365 participants were categorized as having a *higher* and *lower* cognitive-ability-test score, respectively. Similarly, 434 and 306 participants were categorized as having a *higher* and a *lower* cognitive-ability-test score in Experiment 3b. The means for the conditions are shown in Figures 6 and 7. The ANOVA values are shown in Table 2 [7]

Experiment 3a. Experiment 3a replicated all the major findings of Experiment 1. As shown in Figures 6a and 6c, participants recognized more the fake-related news (20.1%) than the real-related news (16.2%) at **Phase 1**. The effect

[7]Complete descriptive statistics for both measures of Experiments 3a and 3b are shown in Supplementary Table S7 and S8, respectively. We also provide the complete posthoc analysis in the Supplementary Material.

Table 3: Summary table for the statistical result in Experiment 3a and 3b. Note. df=degrees of freedom. Bold font indicates statistical significance (*p*< .05).

Phase	Effect	Experiment 3a								Experiment 3b							
		Recognition				Accuracy				Recognition				Accuracy			
		df	F	*p*	η_p^2	*df*	F	*p*	η_p^2	*df*	F	*p*	η_p^2	*df*	F	*p*	η_p^2
P1	news type	**1, 726**	**51.37**	**0.000**	**0.066**	1, 726	0.34	0.560	0.000	**1, 689**	**12.44**	**0.000**	**0.018**	**1, 689**	**15.87**	**0.000**	**0.023**
	news type * condition	2, 726	2.91	0.055	0.008	**2, 726**	**13.71**	**0.000**	**0.036**	**2, 689**	**3.29**	**0.038**	**0.009**	**2, 689**	**12.42**	**0.000**	**0.035**
	news type * cognitive ability	1, 726	2.13	0.145	0.003	1, 726	1.14	0.286	0.002	**1, 689**	**15.34**	**0.000**	**0.022**	1, 689	0.11	0.738	0.000
	news type * condition * cognitive ability	2, 726	0.94	0.392	0.003	2, 726	0.63	0.534	0.002	2, 689	0.30	0.740	0.001	2, 689	0.34	0.709	0.001
	condition	2, 726	0.24	0.784	0.001	2, 726	0.35	0.702	0.001	2, 689	0.04	0.960	0.000	2, 689	0.27	0.765	0.001
	cognitive ability	**1, 726**	**59.46**	**0.000**	**0.076**	1, 726	2.24	0.135	0.003	**1, 689**	**76.80**	**0.000**	**0.100**	1, 689	0.07	0.786	0.000
	condition * cognitive ability	**2, 726**	**3.30**	**0.037**	**0.009**	2, 726	1.13	0.325	0.003	2, 689	0.78	0.459	0.002	**2, 689**	**4.75**	**0.009**	**0.014**
P2	news veracity	1, 782	3.61	0.058	0.005	**1, 782**	**698.48**	**0.000**	**0.472**	**1, 734**	**10.98**	**0.001**	**0.015**	**1, 734**	**624.81**	**0.000**	**0.460**
	news veracity * condition	**2, 782**	**8.08**	**0.000**	**0.020**	**2, 782**	**7.69**	**0.000**	**0.019**	2, 734	1.11	0.330	0.003	2, 734	0.47	0.623	0.001
	news veracity * cognitive ability	1, 782	0.92	0.338	0.001	**1, 782**	**35.56**	**0.000**	**0.043**	1, 734	2.12	0.146	0.003	**1, 734**	**55.51**	**0.000**	**0.070**
	news veracity * condition * cognitive ability	2, 782	0.61	0.544	0.002	2, 782	0.57	0.564	0.001	**2, 734**	**3.23**	**0.040**	**0.009**	2, 734	2.43	0.089	0.007
	condition	2, 782	0.89	0.409	0.002	**2, 782**	**6.65**	**0.001**	**0.017**	2, 734	1.40	0.248	0.004	**2, 734**	**4.02**	**0.018**	**0.011**
	cognitive ability	**1, 782**	**74.60**	**0.000**	**0.087**	**1, 782**	**40.67**	**0.000**	**0.049**	**1, 734**	**68.15**	**0.000**	**0.085**	**1, 734**	**56.47**	**0.000**	**0.071**
	condition * cognitive ability	2, 782	2.57	0.077	0.007	2, 782	1.03	0.356	0.003	2, 734	1.42	0.243	0.004	2, 734	1.02	0.361	0.003

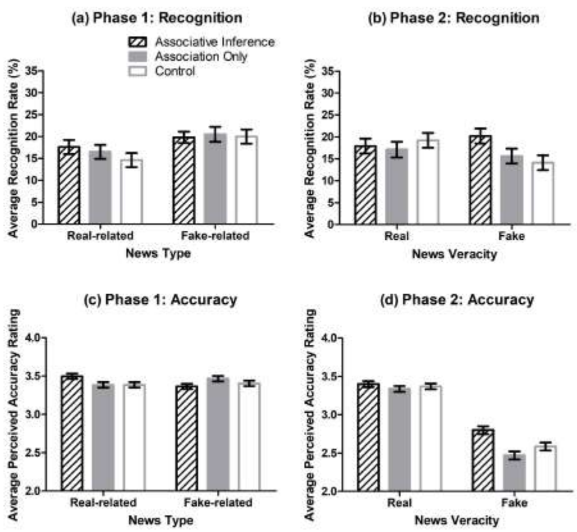


Figure 6: Average recognition rate (top row) and perceived accuracy rating (bottom row) with standard errors across the three conditions for each news type at Phase 1 (left column) and each news veracity at Phase 2 (right column) of Experiment 3a.

of *news type* showed a non-significant trend to be different across *conditions*. The main effect of *news type* was not significant for the accuracy measure, but the two-way interaction of *news type* × *condition* was significant. While participants in the *a.Inf* condition gave higher accuracy for the real-related news than the fake-related news, those in the other conditions revealed an opposite pattern (see Figure 6c).

When moving to **Phase 2** (see Figures 6b and 6d), the two-way interaction of *news veracity* × *condition* was significant for both recognition ($F_{(2,782)} = 8.08, p < .001, \eta_p^2 = .020$) and accuracy ($F_{(2,782)} = 7.69, p < .001, \eta_p^2 = .019$, [$F_{(2,5503)} = 10.37, p < .001$]) measures. Results of post-hoc analysis showed that condition was only significant for the fake news recognition ($F_{(2,782)} = 3.63, p = .027, \eta_p^2 = .009$). Specifically, participants in the *a.Inf* condition recognized more fake news (20.2%) than those in the *CON* condition (14.1%, $p_{adj} = .030$). Similar results were revealed for the post-hoc analysis of accuracy measure. Condition was only significant for the fake news ($F_{(2,782)} = 10.24, p < .001, \eta_p^2 = .026$). Specifically, across the three conditions, participants in the *a.Inf* condition gave the highest accuracy rating (2.80, $p_{adjs} \leq .011$).

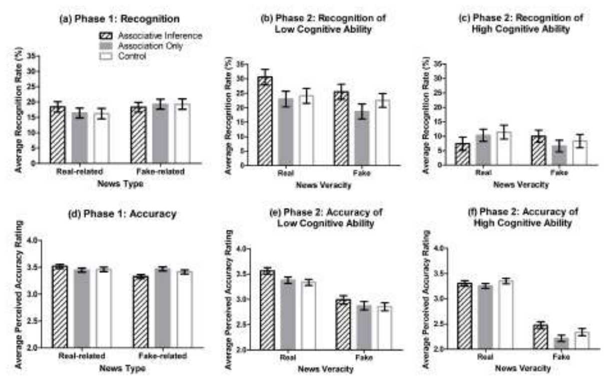


Figure 7: Average recognition rate (top row) and perceived accuracy rating (bottom row) with standard errors across the three conditions for each news type at Phase 1 (left column) and each news veracity of lower cognitive ability (center column) and higher cognitive ability (right column) at Phase 2 of Experiment 3b.

Experiment 3b. Results of Experiment 3b replicated the results of Experiment 2: At **Phase 1**, (1) the main effect of *news type*, *cognitive-ability-test score*, and the two-way interaction of *news type* × *cognitive-ability-test score* were significant for the recognition measure; (2) the two-way interaction of *news type* × *condition* was significant for the accuracy measure; **At Phase 2**, (3) the main effect of *news veracity* and *cognitive-ability-test score* were significant for both the recognition and the accuracy measures; and (4) the two-way interaction of *news veracity* × *cognitive-ability-test score* was only significant for the accuracy measure (See Table 3)

In addition, at **Phase 1**, the two-way interaction of *news type* × *condition* was significant for the recognition measure, showing participants recognized more the fake-related news than the real-related news in the *a.Only* condition and *CON* condition, but not in the *a.Inf* condition ($F < 1$) (see Figure 7a). Furthermore, participants gave slightly higher accuracy rating for the real-related news (3.47) than the fake-related news (3.40).

At **Phase 2**, in addition to the replications, the three-way interaction of *news veracity* × *condition* × *cognitive-ability-test score* was significant ($F_{(2,734)} = 3.23, p = .040, \eta_p^2 = .009$) for the recognition measure (see Figures 7b and 7c). Post-hoc analyses revealed that the two-way interaction of *news veracity* × *condition* was only significant for the participants of a higher cognitive-ability-test score ($F_{(2,431)} = 5.06, p = .007, \eta_p^2 = .023$). Specifically, participants of a higher cognitive-ability-test score in the *a.Inf* condition revealed a non-significant trend of larger recognition rate for the fake news (10.0%) than for the real news (7.4%), but those in *CON* (fake vs. real = 8.3% vs. 11.4%) and *a.Only* (fake vs. real = 6.6% vs. 10.3%) conditions revealed an opposite pattern. For the participants of a lower cognitive-ability-test score, the two-way interaction of *condition* × *news veracity* was not significant, $F < 1$.

The effect of condition was also significant for the perceived accuracy measure (see Figures 7e and 7f). Specifically, participants in the *a.Inf* condition (3.08) gave higher accuracy rating than those in *a.Only* condition (2.93, $p_{adj} = .019$). Also, there was a non-significant three-way interaction of *news veracity* × *condition* × *cognitive-ability-test score* ($F_{(2,734)} = 2.43, p = .089, \eta_p^2 = .007$), which was significant in LMER [$F_{(2,5168)} = 3.28, p = .038$].

To check the effect of blue-highlighted hashtags (Exp.1 vs. Exp.3a) and repeated hashtags (Exp.2 vs. Exp.3b), we conducted additional ANOVAs by adding 'experiment' as another between-subject factor for the combined dataset of Exp.1 and 3a, and Exp. 2 and 3b, respectively. Here, we selectively report the three-way interaction involving the 'experiment' factor. The three-way interaction of experiment × news veracity × condition was not significant for both recognition and accuracy measures ($F_s \leq 1.68$), indicating limited impacts of blue-highlighted or repeated hashtags when they were presented at the bottom of tweets.

Finding 4: *Experiments 3a and 3b replicated the results of Experiments 1 and 2, respectively. Regardless of experiments, results of Phase 1 kept revealing an appropriate baseline for investigating associative inference at Phase 2. Moreover, participants of higher cognitive ability level might have been less impacted by the embedded hashtags and still showed the increased susceptibility to associatively inferred misinformation.*

General Discussion

Findings from our three experiments consistently showed that associative inference could be one cognitive driver of misinformation susceptibility. To this end, we argue that adaptive memory processes (e.g., associative inference) can become maladaptive (e.g., increasing individuals' misremembering and misbelief in fake news), highlighting the essentiality of bringing insights from cognitive psychology into the understanding and mitigation of misinformation on social media platforms. We discuss the theoretical and practical implications of the study.

Associative Inference Increases Recognition and Perceived Accuracy Rating. The most clear-cut findings are that associative inference increases individuals' susceptibility in both recognition (i.e., familiarity) and perceived accuracy measures. Thus, our experiments provide novel evidence showing increased *misbelief* in associatively inferred fake news, which was not obtained in prior studies (Xiong et al. 2022). Xiong et al. explained that their within-subjects design might have made participants, especially those of higher cognitive ability level, more aware of the associative inference across phases. Our findings corroborate their account. Using a between-subjects design, we found that the increased misbelief in associatively inferred fake news was not dependent on participants' cognitive-ability levels. Therefore, without relative comparison with other conditions (e.g., association only), individuals, even those of higher cognitive ability level, could miss the associative inference across phases.

The effect sizes of associative inference may indicate a small to moderate effect, but the observed magnitude is generally in line with the average effect size in the context of misinformation (Lutzke et al. 2019; Murphy et al. 2019). Small effects can be still worth noting (Lakens 2013), especially considering the issue of misinformation spread out on social media that can be aggregated across individuals (Gelman 2018) and over time (Funder and Ozer 2019).

Impact of Hashtags' Presentation. Comparing to the color coding, embedding hashtags in the tweet messages turned out to be more influential for the effect of associative inference. Highlighted hashtags in a tweet message draw a reader's attention. Since the hashtags were highlighted, participants of higher cognitive ability might seek to understand why, which required thinking about the meaning of the overall tweet and how its different pieces relate to one another, indicating analytical thinking (Worthen et al. 2006). In contrast, participants of lower cognitive ability might pay attention to the "pop-out" hashtags mainly, and were too "lazy" to go through the meaning of the overall tweets. Therefore, individuals who are better equipped with adaptive ability to combine existing knowledge in response to novel circumstances can be more vulnerable to misinformation. This is a post-hoc explanation, and future research needs to more thoroughly replicate and investigate the difference.

Associatively Inferred Misinformation Mitigation. The difference between the two designs (i.e., within-subjects and between-subjects) is particularly important for generating guidelines to mitigate associatively inferred misinformation. Existing misinformation correction often focuses on correcting inaccurate content and providing accurate information (Barrera et al. 2020) or information source (Bode and Vraga 2018). Yet, the gap between our study and Xiong et al. points out an approach of *process-based* correction. For instance, a correction that highlights the association and possible inference between existing real news and misinformation might help online news consumers become aware of the attempt to deception. Such process-based correction might offer protection against different types of misinformation that leverage associative inference.

Empirical Analysis of Associatively Inferred Misinformation. Moreover, developing effective misinformation mitigation requires understanding how associatively inferred fake news occurred *in practice*. So far, relatively few experimental studies (Lee et al. 2020; Xiong et al. 2022) examined political misinformation using a specific type (i.e., AB&BC $\rightarrow$ BC), which may not be representative.

Building upon predicting susceptibility to misinformation among social media users (Shen et al. 2019; Teng et al. 2022), future research can conduct an empirical analysis by systematically measuring afforded associative inferences between misinformation on social media and real news consumed by social media users to predict susceptible users to the misinformation. By predicting the most vulnerable group of users, we can set a priority to mitigate the spread of misinformation on social media.

Limitations. Our study has a few limitations. First, all tweets used in our study covered politics mainly. Further investigation on other topics, e.g., health information, can improve the external validity of our study. In addition, we recruited MTurk workers who tended to be younger, better educated, and have better digital literacy (Guess and Munger 2022). Thus, our results may represent a population having more concerns on fake news than the broader U.S. public. Moreover, our study only examined the impact of associative inference in the short term. Future work on extending the gap between two phases can help reveal whether the effect of associative inference will hold in the long term.

Conclusion

Our study shows that one basis for people's susceptibility to misinformation is adaptive, constructive processes of memory. Fake news can be designed to appeal cognitive preference, maybe more than real news. While individuals of lower cognitive-ability level are more susceptible to misinformation in general, our results indicate that those of higher cognitive-ability level are more susceptible to associatively inferred misinformation. Different from the idea that misinformation is attractive to lazy people, our study implies that misinformation, at least some of which fits with our cognitive mechanisms, can make more adaptive individuals also become susceptible.

Ethical Statement

We have taken careful steps to ensure research ethics. First, we worked with our institutional review board (IRB) office and obtained their approval before running the online experiments. Second, implied consent was obtained for each participant before all experiments. Third, the study presented little to no risk compared to those encountered in people's everyday online activities. Moreover, the results from our study can benefit researchers and practitioners to build better tools to mitigate misinformation from human aspects. We believe the benefit outweighs the potential risk.

Acknowledgements

This research was supported in part by Penn State under PSU SSRI Seed Grant and the National Science Foundation under awards No. 1820609 and No. 1915801.

References

Allcott, H.; and Gentzkow, M. 2017. Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2): 211–236.

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

Barrera, O.; Guriev, S.; Henry, E.; and Zhuravskaya, E. 2020. Facts, alternative facts, and fact checking in times of post-truth politics. *Journal of Public Economics*, 182: 104123.

Bates, D.; Maechler, M.; Bolker, B.; Walker, S.; Christensen, R. H. B.; Singmann, H.; Dai, B.; Scheipl, F.; and Grothendieck, G. 2011. Package 'lme4'. *Linear mixed-effects models using S4 classes. R package version*, 1(6).

Begg, I. M.; Anas, A.; and Farinacci, S. 1992. Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General*, 121(4): 446–458.

Bode, L.; and Vraga, E. K. 2018. See something, say something: Correction of global health misinformation on social media. *Health Communication*, 33(9): 1131–1140.

Bonilla, Y.; and Rosa, J. 2015. # Ferguson: Digital protest, hashtag ethnography, and the racial politics of social media in the United States. *American Ethnologist*, 42(1): 4–17.

Brauer, M.; and Curtin, J. J. 2018. Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, 23(3): 389–411.

Braun, V.; and Clarke, V. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2): 77–101.

Brown, S. C.; and Craik, F. I. 2000. Encoding and retrieval of information. *The Oxford handbook of memory*, 93–107.

Bruns, A.; and Burgess, J. 2011. The use of Twitter hashtags in the formation of ad hoc publics. In *Proceedings of the 6th European consortium for political research (ECPR) general conference 2011*, 1–9.

Carpenter, A. C.; and Schacter, D. L. 2017. Flexible retrieval: When true inferences produce false memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(3): 335–349.

Craik, F. I.; Govoni, R.; Naveh-Benjamin, M.; and Anderson, N. D. 1996. The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, 125(2): 159–180.

Faul, F.; Erdfelder, E.; Lang, A.-G.; and Buchner, A. 2007. G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2): 175–191.

Foster, J. L.; Huthwaite, T.; Yesberg, J. A.; Garry, M.; and Loftus, E. F. 2012. Repetition, not number of sources, increases both susceptibility to misinformation and confidence in the accuracy of eyewitnesses. *Acta Psychologica*, 139(2): 320–326.

Frederick, S. 2005. Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19(4): 25–42.

Funder, D. C.; and Ozer, D. J. 2019. Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, 2(2): 156–168.

Gelman, A. 2018. Discussion of effects of growth mindset: Let's not demand unrealistic effect sizes. *Statistical Modeling, Causal Inference, and Social Science*.

Guess, A. M.; and Munger, K. 2022. Digital literacy and online political behavior. *Political Science Research and Methods*, 1–19.

Hauser, D. J.; and Schwarz, N. 2016. Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods*, 48(1): 400–407.

Huang, J.; Thornton, K. M.; and Efthimiadis, E. N. 2010. Conversational tagging in twitter. In *Proceedings of the 21st ACM conference on Hypertext and hypermedia*, 173–178.

Hunt, R. R.; and Lamb, C. A. 2001. What causes the isolation effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(6): 1359–1366.

Keppel, G. 1991. *Design and analysis: A researcher's handbook*. NJ: Prentice-Hall, Inc.

Lakens, D. 2013. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4.

Lazer, D. M.; Baum, M. A.; Benkler, Y.; Berinsky, A. J.; Greenhill, K. M.; Menczer, F.; Metzger, M. J.; Nyhan, B.; Pennycook, G.; Rothschild, D.; et al. 2018. The science of fake news. *Science*, 359(6380): 1094–1096.

Lee, S.; Forrest, J. P.; Strait, J.; Seo, H.; Lee, D.; and Xiong, A. 2020. Beyond Cognitive Ability: Susceptibility to Fake News Is Also Explained by Associative Inference. In *Extended Abstracts of the 2020 ACM CHI Conf.*, 320:1–8.

Lewandowsky, S.; and Van Der Linden, S. 2021. Countering misinformation and fake news through inoculation and prebunking. *European Review of Social Psychology*, 32(2): 348–384.

Loftus, E. F. 2005. Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & memory*, 12(4): 361–366.

Loomba, S.; de Figueiredo, A.; Piatek, S. J.; de Graaf, K.; and Larson, H. J. 2021. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nature Human Behaviour*, 5(3): 337–348.

Lorch, R. F. 1989. Text-signaling devices and their effects on reading and memory processes. *Educational psychology review*, 1(3): 209–234.

Lutzke, L.; Drummond, C.; Slovic, P.; and Árvai, J. 2019. Priming critical thinking: Simple interventions limit the influence of fake news about climate change on Facebook. *Global Environmental Change*, 58: 101964.

Malhotra, N.; Krosnick, J. A.; and Haertel, E. 2007. The psychometric properties of the GSS Wordsum vocabulary test. *GSS Methodological Report*, 11: 1–63.

Matsa, K. E.; and Shearer, E. 2018. News use across social media platforms 2018. *Pew Research Center*.

Murnane, K.; and Phelps, M. P. 1995. Effects of changes in relative cue strength on context-dependent recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(1): 158–172.

Murphy, G.; Loftus, E. F.; Grady, R. H.; Levine, L. J.; and Greene, C. M. 2019. False Memories for Fake News During Ireland's Abortion Referendum. *Psychological Science*, 30(10): 1449–1459.

Pennycook, G.; Cannon, T. D.; and Rand, D. G. 2018. Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, 147(12): 1865–1880.

Pennycook, G.; Epstein, Z.; Mosleh, M.; Arechar, A. A.; Eckles, D.; and Rand, D. G. 2021. Shifting attention to accuracy can reduce misinformation online. *Nature*, 592(7855): 590–595.

Pennycook, G.; and Rand, D. G. 2019. Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, 188: 39–50.

Perlow, J. 2020. Coronavirus misinformation spreading fast: Fake news on COVID-19 shared far more than CDC, WHO reports. https://www.zdnet.com/article/coronavirus-misinformation-is-increasing-newsguard-finds/.

Peterson, S. E. 1992. The cognitive functions of underlining as a study technique. *Literacy Research and Instruction*, 31(2): 49–56.

Pillai, R. M.; and Fazio, L. K. 2021. The effects of repeating false and misleading information on belief. *Wiley Interdisciplinary Reviews: Cognitive Science*, 12(6): e1573.

Rho, E. H. R.; and Mazmanian, M. 2019. Hashtag Burnout? A Control Experiment Investigating How Political Hashtags Shape Reactions to News Content. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW): 1–25.

Rights, J. D.; and Sterba, S. K. 2018. A framework of R-squared measures for single-level and multilevel regression mixture models. *Psychological methods*, 23(3): 434–457.

Rosenberg, S. 2016. EU Referendum: What Does Russia Gain from Brexit? *BBC News*.

Seo, H.; Xiong, A.; and Lee, D. 2019. Trust It or Not: Effects of Machine-Learning Warnings in Helping Individuals Mitigate Misinformation. In *Proceedings of the 10th ACM Conference on Web Science*, 265–274.

Shen, T. J.; Cowell, R.; Gupta, A.; Le, T.; Yadav, A.; and Lee, D. 2019. How gullible are you? Predicting susceptibility to fake news. In *Proceedings of the 10th ACM Conference on Web Science*, 287–288.

Teng, X.; Lin, Y.-R.; Chung, W.-T.; Li, A.; and Kovashka, A. 2022. Characterizing User Susceptibility to COVID-19 Misinformation on Twitter. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 16, 1005–1016.

Thomson, K. S.; and Oppenheimer, D. M. 2016. Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making*, 11(1): 99–113.

Visentin, M.; Pizzi, G.; and Pichierri, M. 2019. Fake news, real problems for brands: The impact of content truthfulness and source credibility on consumers' behavioral intentions toward the advertised brands. *Journal of Interactive Marketing*, 45: 99–112.

Vosoughi, S.; Roy, D.; and Aral, S. 2018. The spread of true and false news online. *Science*, 359(6380): 1146–1151.

Wechsler, D. 2008. Wechsler adult intelligence scale–Fourth Edition (WAIS–IV). *San Antonio, TX: NCS Pearson*, 22(498): 816–827.

Worthen, R. R. H. J. B.; et al. 2006. *Distinctiveness and memory*. Oxford University Press.

Wu, L.; Morstatter, F.; Carley, K. M.; and Liu, H. 2019. Misinformation in social media: definition, manipulation, and detection. *ACM SIGKDD Explorations Newsletter*, 21(2): 80–90.

Xiong, A.; Lee, S.; Seo, H.; and Lee, D. 2022. Effects of associative inference on individuals' susceptibility to misinformation. *Journal of Experimental Psychology: Applied.*

Zeithamova, D.; and Preston, A. R. 2010. Flexible memories: differential roles for medial temporal lobe and prefrontal cortex in cross-episode binding. *Journal of Neuroscience*, 30(44): 14676–14684.

Stimuli of EXP1

(a) Real 1 related AX news **(b)** Real 1 related YC news **(c)** Real 2 related AX news **(d)** Real 2 related YC news

(e) Real 3 related AX news **(f)** Real 3 related YC news **(g)** Real 4 related AX news **(h)** Real 4 related YC news

Fig. S1. Snippets of news used in Phase 1 which are related with **real** news in Phase 2.

(a) Fake 1 related AB news **(b)** Fake 1 related BC news **(c)** Fake 2 related AB news **(d)** Fake 2 related BC news

(e) Fake 3 related AB news **(f)** Fake 3 related BC news **(g)** Fake 4 related AB news **(h)** Fake 4 related BC news

Fig. S2. Snippets of news used in ***a.Inf*** condition at Phase 1 which are related with **fake** news in Phase 2.

(a) Fake 1 related AX news **(b)** Fake 1 related YC news **(c)** Fake 2 related AX news **(d)** Fake 2 related YC news **(e)** Fake 3 related AX news **(f)** Fake 3 related YC news **(g)** Fake 4 related AX news **(h)** Fake 4 related YC news

Fig. S3. Snippets of news used in ***a.Only*** condition at Phase 1 which are related with **fake** news in Phase 2.

(a) Fake 1 related DE news **(b)** Fake 1 related FG news **(c)** Fake 2 related DE news **(d)** Fake 2 related FG news **(e)** Fake 3 related DE news **(f)** Fake 3 related FG news **(g)** Fake 4 related DE news **(h)** Fake 4 related FG news

Fig. S4. Snippets of news used in ***CON*** condition at Phase 1 which are related with **fake** news in Phase 2.

(a) Real 1 news

(b) Real 2 news

(c) Real 3 news

(d) Real 4 news

(e) Fake 1 news

(f) Fake 2 news

(g) Fake 3 news

(h) Fake 4 news

Fig. S5. Snippets of news used in Phase 2. Each row represents different *veracity* (real, fake) level.

Fig. S6. Snippets of news used for participant's attention check in Phase 1 (left) and 2 (right).

Stimuli of EXP2

(a) Real 1 related AX news

(b) Real 1 related YC news

(c) Real 2 related AX news

(d) Real 2 related YC news

(e) Real 3 related AX news

(f) Real 3 related YC news

(g) Real 4 related AX news

(h) Real 4 related YC news

Fig. S7. Snippets of news used in Phase 1 which are related with **real** news in Phase 2.

(a) Fake 1 related AB news

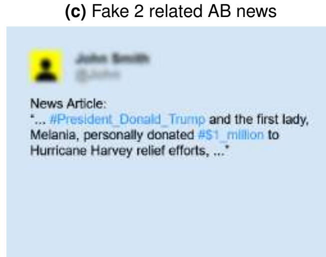
(b) Fake 1 related BC news

(c) Fake 2 related AB news

(d) Fake 2 related BC news

(e) Fake 3 related AB news

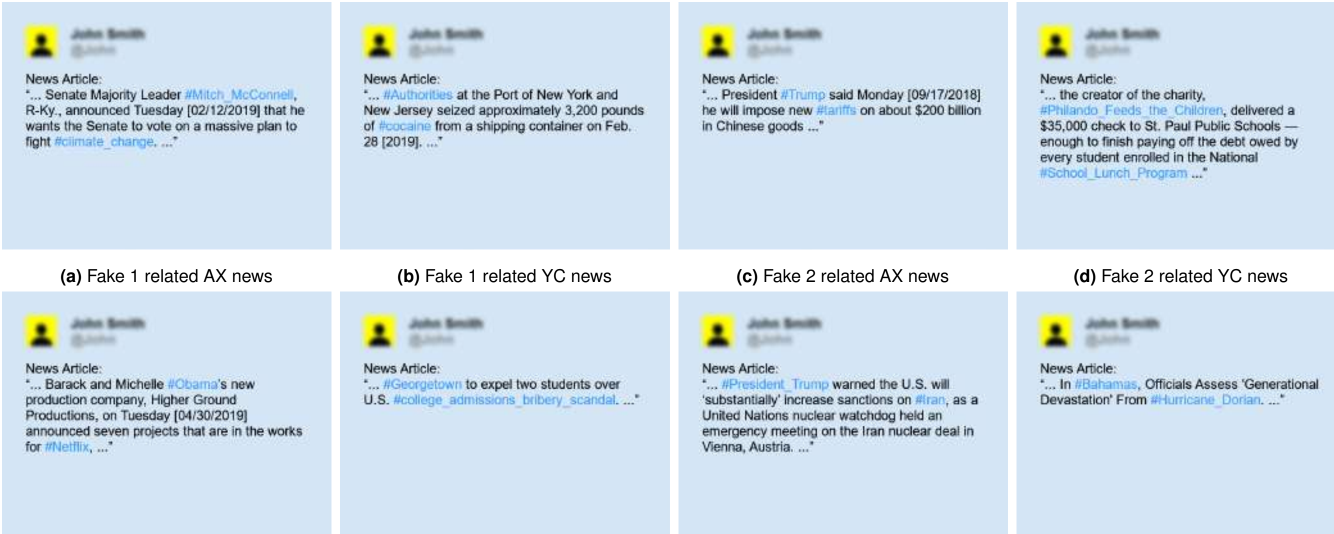
(f) Fake 3 related BC news



(g) Fake 4 related AB news

(h) Fake 4 related BC news

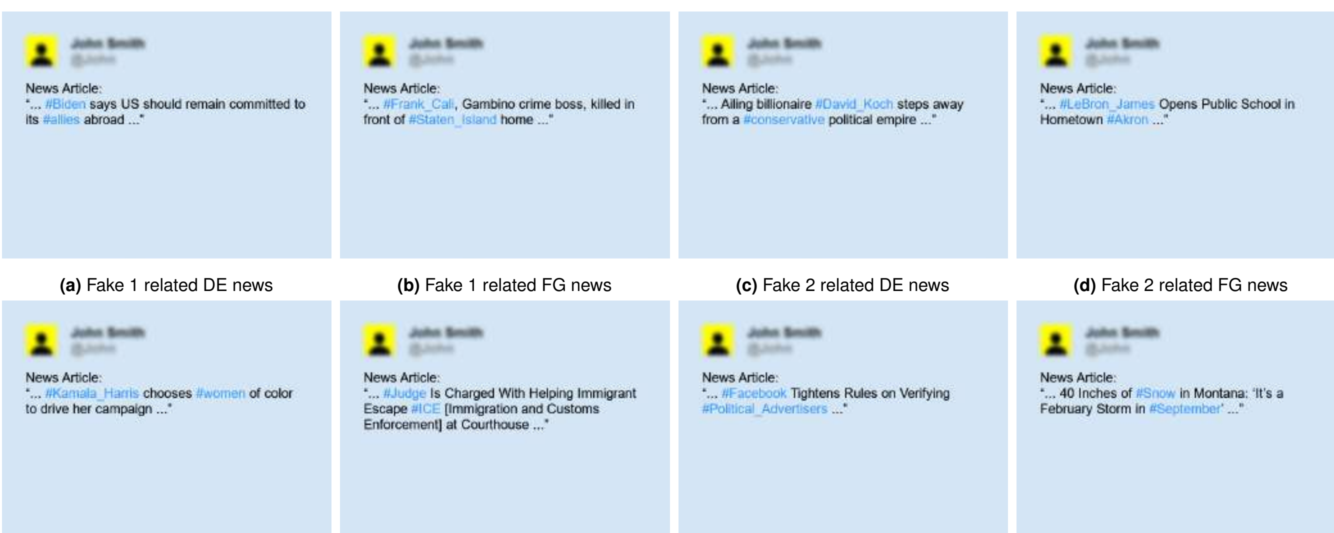
Fig. S8. Snippets of news used in ***a.Inf*** condition at Phase 1 which are related with **fake** news in Phase 2.



(a) Fake 1 related AX news **(b)** Fake 1 related YC news **(c)** Fake 2 related AX news **(d)** Fake 2 related YC news

(e) Fake 3 related AX news **(f)** Fake 3 related YC news **(g)** Fake 4 related AX news **(h)** Fake 4 related YC news

Fig. S9. Snippets of news used in ***a.Only*** condition at Phase 1 which are related with **fake** news in Phase 2.



(a) Fake 1 related DE news **(b)** Fake 1 related FG news **(c)** Fake 2 related DE news **(d)** Fake 2 related FG news

(e) Fake 3 related DE news **(f)** Fake 3 related FG news **(g)** Fake 4 related DE news **(h)** Fake 4 related FG news

Fig. S10. Snippets of news used in ***CON*** condition at Phase 1 which are related with **fake** news in Phase 2.

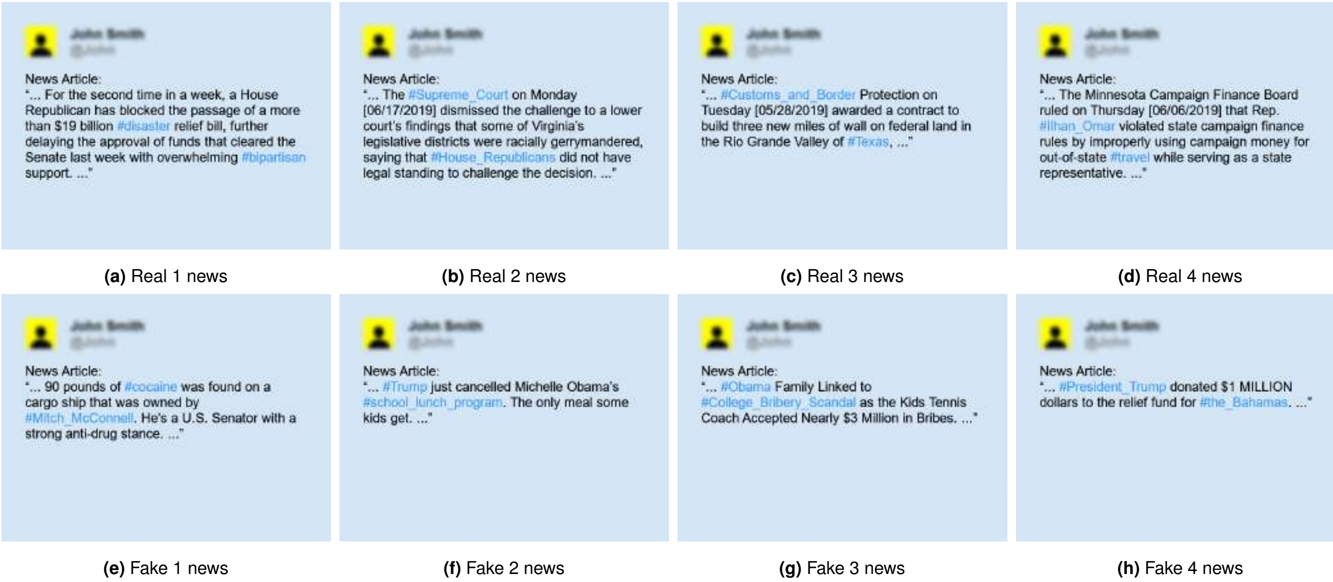


Fig. S11. Snippets of news used in Phase 2. Each row represents different *veracity* (real, fake) level.

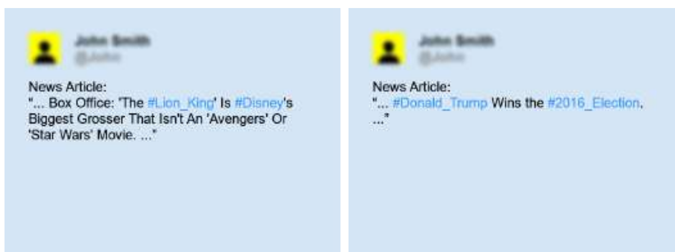


Fig. S12. Snippets of news used for participant's attention check in Phase 1 (left) and 2 (right).

Stimuli of EXP3a

(a) Real 1 related AX news

(b) Real 1 related YC news

(c) Real 2 related AX news

(d) Real 2 related YC news

(e) Real 3 related AX news

(f) Real 3 related YC news

(g) Real 4 related AX news

(h) Real 4 related YC news

Fig. S13. Snippets of news used in Phase 1 which are related with **real** news in Phase 2.

(a) Fake 1 related AB news

(b) Fake 1 related BC news

(c) Fake 2 related AB news

(d) Fake 2 related BC news

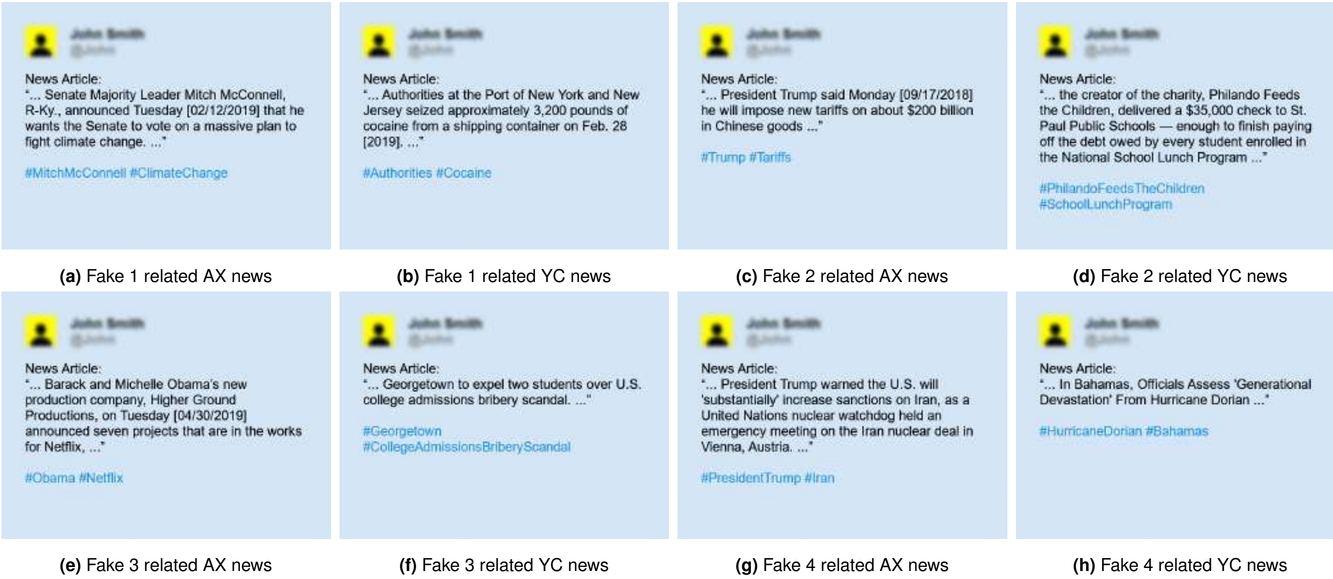
(e) Fake 3 related AB news

(f) Fake 3 related BC news

(g) Fake 4 related AB news

(h) Fake 4 related BC news

Fig. S14. Snippets of news used in ***a.Inf*** condition at Phase 1 which are related with **fake** news in Phase 2.



(a) Fake 1 related AX news **(b)** Fake 1 related YC news **(c)** Fake 2 related AX news **(d)** Fake 2 related YC news

(e) Fake 3 related AX news **(f)** Fake 3 related YC news **(g)** Fake 4 related AX news **(h)** Fake 4 related YC news

Fig. S15. Snippets of news used in ***a.Only*** condition at Phase 1 which are related with **fake** news in Phase 2.

(a) Fake 1 related DE news **(b)** Fake 1 related FG news **(c)** Fake 2 related DE news **(d)** Fake 2 related FG news

(e) Fake 3 related DE news **(f)** Fake 3 related FG news **(g)** Fake 4 related DE news **(h)** Fake 4 related FG news

Fig. S16. Snippets of news used in ***CON*** condition at Phase 1 which are related with **fake** news in Phase 2.

(a) Real 1 news

(b) Real 2 news

(c) Real 3 news

(d) Real 4 news

(e) Fake 1 news

(f) Fake 2 news

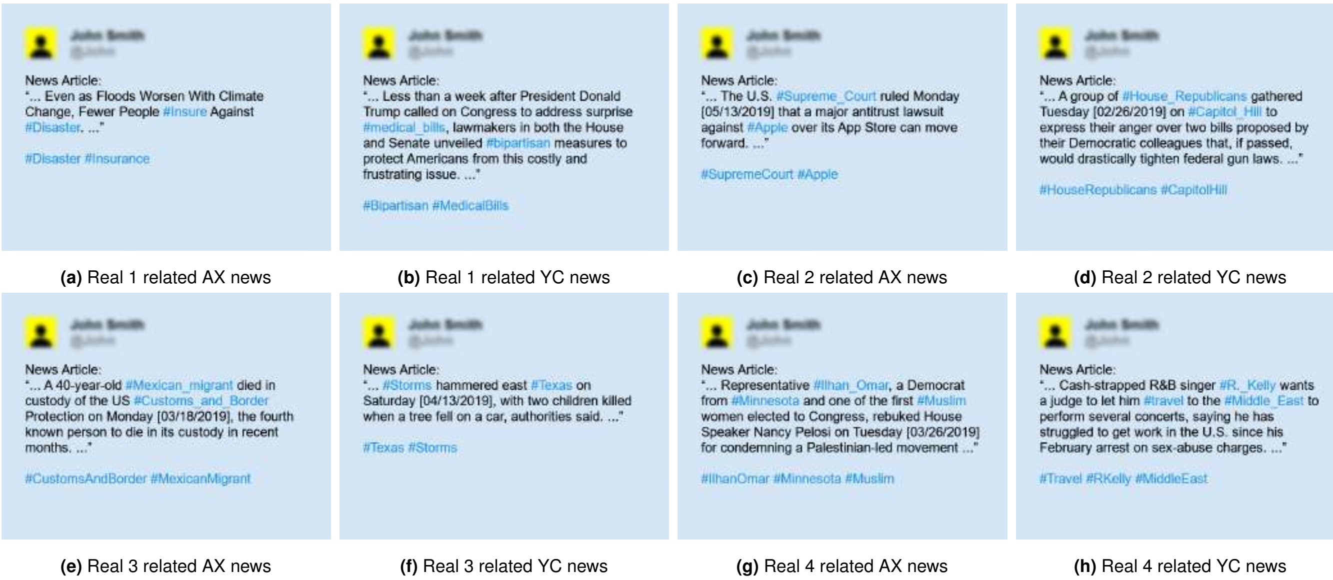
(g) Fake 3 news

(h) Fake 4 news

Fig. S17. Snippets of news used in Phase 2. Each row represents different *veracity* (real, fake) level.

Fig. S18. Snippets of news used for participant's attention check in Phase 1 (left) and 2 (right).

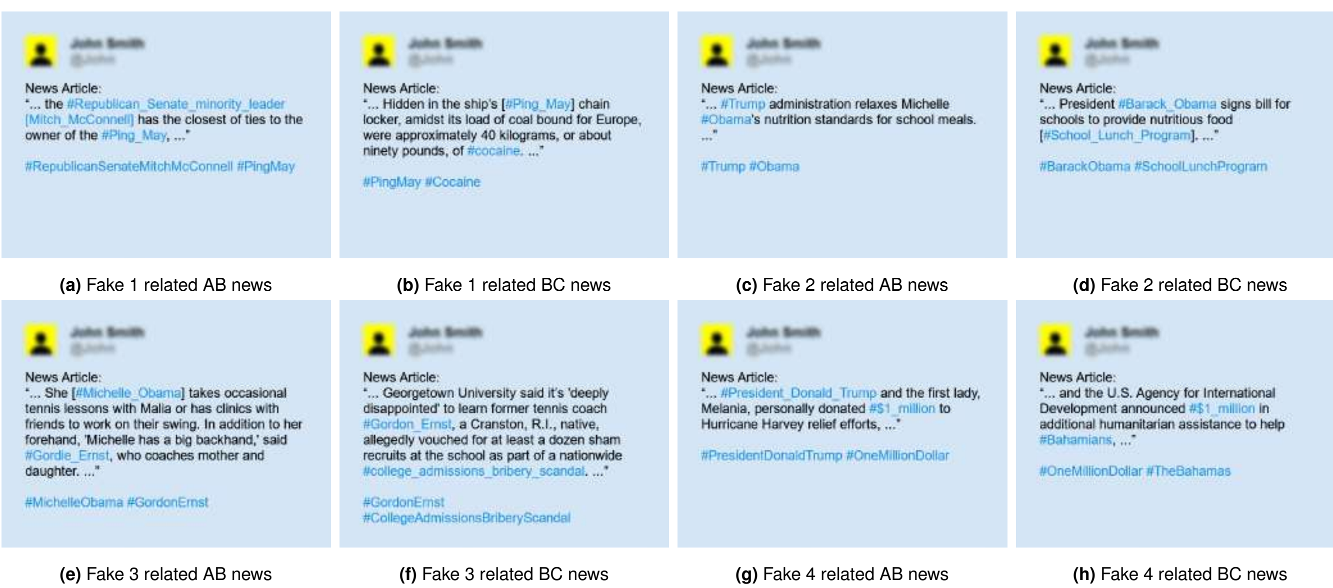
4 Stimuli of EXP3b



(a) Real 1 related AX news **(b)** Real 1 related YC news **(c)** Real 2 related AX news **(d)** Real 2 related YC news

(e) Real 3 related AX news **(f)** Real 3 related YC news **(g)** Real 4 related AX news **(h)** Real 4 related YC news

Fig. S19. Snippets of news used in Phase 1 which are related with **real** news in Phase 2.



(a) Fake 1 related AB news **(b)** Fake 1 related BC news **(c)** Fake 2 related AB news **(d)** Fake 2 related BC news

(e) Fake 3 related AB news **(f)** Fake 3 related BC news **(g)** Fake 4 related AB news **(h)** Fake 4 related BC news

Fig. S20. Snippets of news used in ***a.Inf*** condition at Phase 1 which are related with **fake** news in Phase 2.

(a) Fake 1 related AX news

(b) Fake 1 related YC news

(c) Fake 2 related AX news

(d) Fake 2 related YC news

(e) Fake 3 related AX news

(f) Fake 3 related YC news

(g) Fake 4 related AX news

(h) Fake 4 related YC news

Fig. S21. Snippets of news used in ***a.Only*** condition at Phase 1 which are related with **fake** news in Phase 2.

(a) Fake 1 related DE news

(b) Fake 1 related FG news

(c) Fake 2 related DE news

(d) Fake 2 related FG news

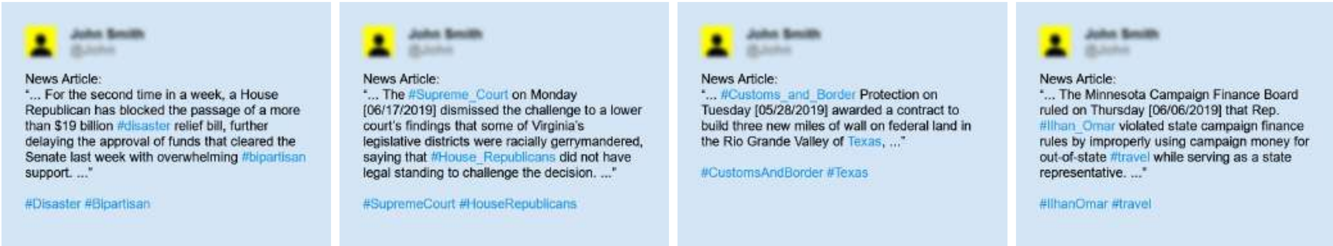
(e) Fake 3 related DE news

(f) Fake 3 related FG news

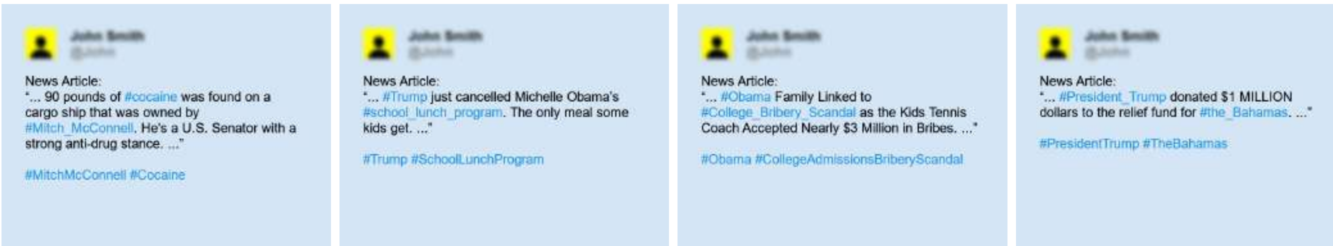
(g) Fake 4 related DE news

(h) Fake 4 related FG news

Fig. S22. Snippets of news used in ***CON*** condition at Phase 1 which are related with **fake** news in Phase 2.



(a) Real 1 news **(b)** Real 2 news **(c)** Real 3 news **(d)** Real 4 news



(e) Fake 1 news **(f)** Fake 2 news **(g)** Fake 3 news **(h)** Fake 4 news

Fig. S23. Snippets of news used in Phase 2. Each row represents different *veracity* (real, fake) level.

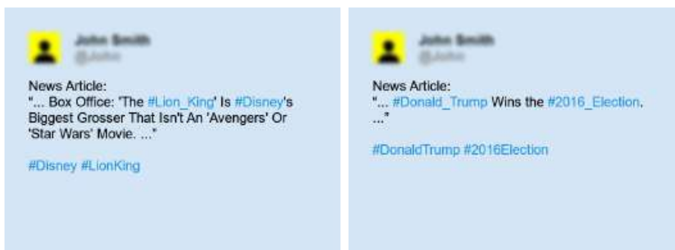


Fig. S24. Snippets of news used for participant's attention check in Phase 1 (left) and 2 (right).

Tables for Demographics

Table S1. Summary of participants exclusion criteria. Negative values mean exclusion of participants. We also removed participants who selected "Prefer Not to Answer (PNA)" for any question in Phase 2 or Phase 1.

	EXP 1	EXP 2	EXP 3
Recruitment	**1200**	**1200**	**2400**
Failed to pass attention checks	*-431*	*-388*	*-694*
Duplicate IPs	*-4*	*-8*	*-5*
PNA for Phase 2	*-79*	*-94*	*-173*
Final Sample in Phase 2	**686**	**718**	**1528** **EXP3a: 788** **EXP3b: 740**
PNA for Phase 1	*-42*	*-65*	*-101*
Final Sample in Phase 1	**644**	**653**	**1472** **EXP3a: 732** **EXP3b: 695**

Table S2. Participants' demographic information in each condition for each experiment. Number in the bracket indicates the number of participants.

Item	**Options**	**EXP 1 (686)**			**EXP 2 (718)**			**EXP 3a (788)**			**EXP3b (740)**		
		a.Inf (231)	**a.Only (217)**	**CON (238)**	**a.Inf (231)**	**a.Only (231)**	**CON (256)**	**a.Inf (265)**	**a.Only (258)**	**CON (265)**	**a.Inf (244)**	**a.Only (261)**	**CON (235)**
Gender	Male	47.6%	58.0%	52.9%	54.5%	53.7%	48.0%	50.6%	48.8%	54.7%	45.5%	45.2%	47.2%
	Female	52.4%	41.5%	47.1%	45.0%	46.3%	51.2%	48.7%	50.8%	45.3%	54.1%	54.4%	51.5%
	Other	0%	0%	0%	0.5%	0%	0.4%	0.8%	0%	0%	0.4%	0.4%	0.9%
	Prefer not to answer	0%	0.5%	0%	0%	0%	0.4%	0%	0.4%	0%	0%	0%	0.4%
Age	18-27	16.5%	16.6%	20.2%	13.9%	21.2%	16.8%	18.9%	17.8%	21.9%	20.1%	17.6%	21.7%
	28-37	46.3%	41.9%	38.2%	48.1%	43.7%	37.1%	46.0%	41.5%	41.5%	39.3%	41.0%	40.0%
	38-47	17.7%	17.1%	21.0%	18.2%	21.2%	26.6%	17.7%	20.9%	24.2%	21.7%	20.7%	20.0%
	48-57	11.7%	13.8%	10.9%	13.9%	10.4%	14.8%	11.3%	10.5%	6.8%	12.7%	12.6%	7.7%
	58 or order	7.8%	10.6%	9.7%	6.1%	3.5%	4.7%	6.0%	8.5%	5.7%	6.1%	8.0%	10.6%
	Prefer not to answer	0%	0%	0%	0%	0%	0%	0%	0.8%	0%	0%	0%	0%
Education	No high school	0.9%	0%	0%	0.4%	0%	0.4%	0.4%	0.8%	0%	0.4%	0%	0.4%
	High school	7.4%	7.8%	9.7%	6.5%	5.2%	6.6%	9.8%	5.8%	9.8%	7.0%	8.8%	8.9%
	College/Bachelor	71.9%	74.2%	72.7%	68.4%	68.4%	69.5%	69.4%	69.8%	67.5%	72.1%	65.5%	70.6%
	Professional degree/Masters/Ph.D.	19.9%	18.0%	17.6%	24.7%	26.4%	23.4%	20.4%	22.9%	22.6%	20.5%	25.7%	20.0%
	Prefer not to answer	0%	0%	0%	0%	0%	0%	0%	0.8%	0%	0%	0%	0%
Time on Soc. Media (Per Day)	Less than 1 hour	51.1%	43.4%	46.2%	28.5%	30.3%	34.8%	38.1%	36.8%	36.6%	43.9%	46.4%	38.7%
	Between 1 to 4 hours	40.7%	46.1%	47.1%	61.5%	59.3%	53.5%	53.6%	57.8%	54.7%	48.8%	47.1%	51.9%
	Longer than 4 hours	8.2%	10.6%	6.3%	10.0%	10.0%	11.3%	8.3%	5.0%	8.7%	7.0%	6.1%	9.4%
	Prefer not to answer	0%	0.5%	0.4%	0%	0.4%	0.4%	0%	0.4%	0%	0.4%	0.4%	0%
Political Stance	Liberal	52.4%	46.5%	45.8%	41.1%	41.1%	34.4%	42.6%	45.0%	44.5%	54.1%	47.1%	47.7%
	Moderate	24.2%	24.0%	25.2%	19.0%	22.9%	22.3%	25.7%	22.1%	23.8%	23.4%	23.8%	25.1%
	Conservative	23.4%	29.5%	29.0%	39.8%	35.9%	43.4%	31.7%	32.9%	31.7%	22.5%	29.1%	27.2%
	Prefer not to answer	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Table S3. Factors affecting participants' perceived accuracy rating decision for one of the four fake news in Phase 2, EXP1. In EXP1, We randomly presented one piece of the four fake news in Phase 2 with the perceived accuracy rating that the participant gave and asked the participant to select all factors affecting his/her decision. Since multiple choice allowed, percentages (second column) were calculated based on the total number of participants (686).

Options	% based on total participants (686)
Source	17.3%
Writing style	19.4%
Content	60.2%
Web search results	6.9%
News presented in Phase 1	11.1%
News that I saw before this study	31.8%
Opinions from others	14.6%
Other	11.7%
Prefer not to answer	1.5%

Summary Tables

Table S4. Independent and dependent measures of each phase for all Experiments.

Phase	Independent Variables		
1	News Type	real-related	A binary categorical variable indicates different type of real news that participants viewed in Phase 1.
		fake-related	
2	Veracity	real	A binary categorical variable indicates whether participants saw a real or fake news in Phase 2.
		fake	
1 & 2	Condition	Associative Inference (a.Inf)	Three-level between subject factor indicating manipulation of fake-related news in Phase 1. Each snippet of news in Phase 2 corresponded to one pair of real news in Phase 1. Critically, we manipulated the correspondence between the two phases across three conditions.
		Association Only (a.Only)	
		Control (CON)	
	Cognitive-ability-test Score	low	Based on participants' cognitive ability tests results with 14 questions. Participants who got 10 or more correct answers were categorized as of high cognitive ability. Otherwise, participants were categorized as of low cognitive ability.
		high	
Phase	**Dependent Variables**		
1 & 2	Recognition Rate	N/A	Selection rate of "Yes" option for each news type in Phase 1 or each veracity level of news in Phase 2 of each condition were calculated for each participant.
	Perceived Accuracy Rating		Average rating for each news type in Phase 1 or each veracity level of news in Phase 2 of each condition were calculated for each participant.

Table S5. Recognition and perceived accuracy results of each condition in each phase for Experiment 1. Sub. # means the number of subjects.

Cognitive-ability-test Score	Condition	Phase 1					Phase 2				
		Sub. #	Recognition		Perceived Accuracy		Sub. #	Recognition		Perceived Accuracy	
			Real -related	Fake -related	Real -related	Fake -related		Real	Fake	Real	Fake
High Group	**Associative Inference**	140	7.6%	11.1%	3.35	3.33	148	8.6%	14.0%	3.20	2.51
	Association Only	130	9.8%	18.6%	3.45	3.57	132	12.1%	13.4%	3.33	2.34
	Control	133	9.3%	13.3%	3.47	3.40	139	9.5%	11.0%	3.28	2.31
Low Group	**Associative Inference**	71	21.1%	22.2%	3.53	3.40	83	17.5%	22.6%	3.35	2.95
	Association Only	81	23.5%	27.5%	3.41	3.54	85	22.9%	20.9%	3.29	2.69
	Control	89	24.9%	25.8%	3.58	3.47	99	23.7%	24.7%	3.34	2.86

Table S6. Recognition and perceived accuracy results of each condition in each phase for Experiment 2. Sub. # means the number of subjects.

Cognitive-ability-test Score	Condition	Phase 1					Phase 2				
		Sub. #	Recognition		Perceived Accuracy		Sub. #	Recognition		Perceived Accuracy	
			Real -related	Fake -related	Real -related	Fake -related		Real	Fake	Real	Fake
High Group	**Associative Inference**	73	10.3%	12.5%	3.43	3.26	77	15.3%	10.7%	3.30	2.50
	Association Only	84	11.8%	16.7%	3.31	3.41	86	12.8%	7.0%	3.17	2.28
	Control	87	14.9%	20.1%	3.41	3.43	92	14.9%	12.2%	3.38	2.30
Low Group	**Associative Inference**	138	38.4%	37.9%	3.56	3.45	154	41.7%	40.4%	3.56	3.31
	Association Only	129	36.3%	37.7%	3.55	3.55	145	37.8%	34.3%	3.50	3.22
	Control	142	36.4%	39.0%	3.63	3.63	164	37.5%	37.5%	3.56	3.32

Table S7. Recognition and perceived accuracy results of each condition in each phase for Experiment 3a. Sub. # means the number of subjects.

Cognitive-ability-test Score	Condition	Phase 1					Phase 2				
		Sub. #	Recognition		Perceived Accuracy		Sub. #	Recognition		Perceived Accuracy	
			Real -related	Fake -related	Real -related	Fake -related		Real	Fake	Real	Fake
High Group	**Associative Inference**	121	8.0%	10.0%	3.46	3.28	125	7.6%	9.6%	3.30	2.57
	Association Only	141	9.2%	14.2%	3.35	3.44	149	9.2%	7.7%	3.26	2.21
	Control	140	9.6%	16.7%	3.40	3.40	149	14.8%	7.7%	3.37	2.39
Low Group	**Associative Inference**	124	27.2%	29.6%	3.54	3.45	140	28.2%	30.9%	3.50	3.03
	Association Only	101	23.8%	26.9%	3.41	3.49	109	25.0%	23.4%	3.41	2.72
	Control	105	19.5%	23.3%	3.37	3.41	116	23.7%	20.5%	3.37	2.78

Table S8. Recognition and perceived accuracy results of each condition in each phase for Experiment 3b. Sub. # means the number of subjects.

Cognitive-ability-test Score	Condition	Phase 1					Phase 2				
		Sub. #	Recognition		Perceived Accuracy		Sub. #	Recognition		Perceived Accuracy	
			Real -related	Fake -related	Real -related	Fake -related		Real	Fake	Real	Fake
High Group	Associative Inference	136	8.8%	10.6%	3.54	3.33	145	7.4%	10.0%	3.31	2.48
	Association Only	155	6.5%	11.4%	3.35	3.39	162	10.3%	6.6%	3.25	2.22
	Control	119	8.4%	14.3%	3.51	3.48	127	11.4%	8.3%	3.35	2.34
Low Group	Associative Inference	94	27.9%	26.1%	3.50	3.32	99	30.6%	25.5%	3.56	2.99
	Association Only	92	26.4%	27.3%	3.54	3.55	99	23.0%	18.7%	3.38	2.88
	Control	99	24.0%	24.2%	3.41	3.34	108	24.1%	22.5%	3.34	2.85

Table S9. Summary table for the recognition statistical result in Phase 1 of Experiment 1. Indention of an effect refers to a post-hoc analysis result.

Effect	*df*	F	*p*	η_p^2
news type	**1, 638**	**34.18**	**0.000**	**0.051**
news type * cognitive-ability-test score	**1, 638**	**7.16**	**0.008**	**0.011**
low: news type	1, 238	3.29	0.071	0.014
high: news type	**1, 400**	**55.90**	**0.000**	**0.123**
news type * condition	**2, 638**	**4.37**	**0.013**	**0.013**
real-related: condition	2, 638	0.81	0.447	0.003
fake-related: condition	**2, 638**	**4.15**	**0.016**	**0.013**
a.Inf vs. a.Only	**NA**	**NA**	**0.004**	**NA**
a.Inf vs. con	NA	NA	0.175	NA
a.Only vs. con	NA	NA	0.113	NA
news type * cognitive-ability-test score * condition	2, 638	0.29	0.750	0.001
cognitive-ability-test score	**1, 638**	**54.01**	**0.000**	**0.078**
condition	2, 638	2.14	0.118	0.007
cognitive-ability-test score * condition	2, 638	0.23	0.797	0.001

Table S10. Summary table for the perceived accuracy statistical result in Phase 1 of Experiment 1. Indention of an effect refers to a post-hoc analysis result.

Effect	*df*	F	*p*	η_p^2
news type	1, 638	0.56	0.453	0.001
news type * cognitive-ability-test score	1, 638	1.59	0.208	0.002
news type * condition	**2, 638**	**14.27**	**0.000**	**0.043**
real-related: condition	2, 638	1.90	0.151	0.006
fake-related: condition	**2, 638**	**6.58**	**0.001**	**0.020**
a.Inf vs. a.Only	**NA**	**NA**	**0.000**	**NA**
a.Inf vs. con	NA	NA	0.161	NA
a.Only vs. con	**NA**	**NA**	**0.023**	**NA**
news type * cognitive-ability-test score * condition	2, 638	0.95	0.388	0.003
cognitive-ability-test score	1, 638	2.44	0.119	0.004
condition	2, 638	1.96	0.141	0.006
cognitive-ability-test score * condition	2, 638	1.50	0.223	0.005

Table S11. Summary table for the recognition statistical result in Phase 2 of Experiment 1. Indention of an effect refers to a post-hoc analysis result.

Effect	*df*	F	*p*	η_p^2
veracity	**1, 680**	**5.08**	**0.025**	**0.007**
veracity * condition	**2, 680**	**3.35**	**0.036**	**0.010**
a.Inf: veracity	**1, 229**	**9.36**	**0.002**	**0.039**
a.Only: veracity	1, 215	0.05	0.820	0.000
con: veracity	1, 236	0.80	0.372	0.003
veracity * cognitive-ability-test score	1, 680	0.57	0.451	0.001
veracity * condition * cognitive-ability-test score	2, 680	0.30	0.738	0.001
condition	2, 680	0.36	0.698	0.001
cognitive-ability-test score	**1, 680**	**34.89**	**0.000**	**0.049**
condition * cognitive-ability-test score	2, 680	0.92	0.399	0.003

Table S12. Summary table for the perceived accuracy statistical result in Phase 2 of Experiment 1. Indention of an effect refers to a post-hoc analysis result.

Effect	*df*	F	*p*	η_p^2
veracity	**1, 680**	**524.71**	**0.000**	**0.436**
veracity * condition	**2, 680**	**5.98**	**0.003**	**0.017**
real: condition	2, 680	0.26	0.771	0.001
fake: condition	**2, 680**	**4.08**	**0.017**	**0.012**
a.Inf vs. a.Only	**NA**	**NA**	**0.005**	**NA**
a.Inf vs. con	NA	NA	0.059	NA
con vs. a.Only	NA	NA	0.326	NA
veracity * cognitive-ability-test score	**1, 680**	**43.38**	**0.000**	**0.060**
real: cognitive ability	1, 680	1.15	0.284	0.002
fake: cognitive ability	**1, 680**	**53.39**	**0.000**	**0.073**
veracity * condition * cognitive-ability-test score	2, 680	0.90	0.406	0.003
condition	2, 680	1.19	0.304	0.003
cognitive-ability-test score	**1, 680**	**29.14**	**0.000**	**0.041**
condition * cognitive-ability-test score	2, 680	1.00	0.369	0.003

Table S13. Summary table for the recognition statistical result in Phase 1 of Experiment 2. Indention of an effect refers to a post-hoc analysis result.

Effect	*df*	F	*p*	η_p^2
news type	**1, 647**	**16.17**	**0.000**	**0.024**
news type * condition	2, 647	1.95	0.143	0.006
news type * cognitive-ability-test score	**1, 647**	**5.10**	**0.024**	**0.008**
low: news type	1, 406	1.99	0.159	0.005
high: news type	1, 241	0.50	0.481	0.002
real-related: cognitive-ability-test score	**1, 647**	**99.13**	**0.000**	**0.133**
fake-related: cognitive-ability-test score	**1, 647**	**79.67**	**0.000**	**0.110**
news type * condition * cognitive-ability-test score	2, 647	0.06	0.945	0.000
condition	2, 647	0.51	0.600	0.002
cognitive-ability-test score	**1, 647**	**95.99**	**0.000**	**0.129**
condition * cognitive-ability-test score	2, 647	0.65	0.523	0.002

Table S14. Summary table for the perceived accuracy statistical result in Phase 1 of Experiment 2. Indention of an effect refers to a post-hoc analysis result.

Effect	*df*	F	*p*	η_p^2
news type	1, 647	1.91	0.167	0.003
news type * condition	**2, 647**	**8.58**	**0.000**	**0.026**
real-related: condition	2, 647	1.27	0.280	0.004
fake-related: condition	**2, 647**	**4.44**	**0.012**	**0.014**
a.Inf vs. a.Only	NA	NA	0.111	NA
a.Inf vs. con	**NA**	**NA**	**0.011**	**NA**
a.Only vs. con	NA	NA	1.000	NA
a.Inf: news type	**1, 209**	**13.23**	**0.000**	**0.060**
a.Only: news type	1, 211	3.28	0.072	0.015
con: news type	1, 227	0.05	0.816	0.000
news type * cognitive-ability-test score	1, 647	0.07	0.792	0.000
news type * condition * cognitive-ability-test score	2, 647	1.41	0.246	0.004
condition	2, 647	1.81	0.164	0.006
cognitive-ability-test score	**1, 647**	**17.46**	**0.000**	**0.026**
condition * cognitive-ability-test score	2, 647	0.10	0.903	0.000

Table S15. Summary table for the recognition statistical result in Phase 2 of Experiment 2. Indention of an effect refers to a post-hoc analysis result.

Effect	*df*	F	*p*	η_p^2
veracity	**1 ,712**	**10.28**	**0.001**	**0.014**
veracity * condition	2 ,712	1.07	0.342	0.003
veracity * cognitive-ability-test score	1 ,712	2.24	0.135	0.003
veracity * condition * cognitive-ability-test score	2 ,712	0.02	0.982	0.000
condition	2 ,712	0.93	0.396	0.003
cognitive-ability-test score	**1 ,712**	**114.67**	**0.000**	**0.139**
condition * cognitive-ability-test score	2 ,712	0.25	0.780	0.001

Table S16. Summary table for the perceived accuracy statistical result in Phase 2 of Experiment 2. Indention of an effect refers to a post-hoc analysis result.

Effect	*df*	F	*p*	η_p^2
veracity	**1 ,712**	**370.29**	**0.000**	**0.342**
veracity * condition	2 ,712	1.49	0.227	0.004
veracity * cognitive-ability-test score	**1 ,712**	**119.87**	**0.000**	**0.144**
real: cognitive ability	**1, 712**	**23.95**	**0.000**	**0.033**
fake: cognitive-ability-test score	**1, 712**	**201.23**	**0.000**	**0.220**
low: veracity	**1, 460**	**47.66**	**0.000**	**0.094**
high: veracity	**1, 252**	**364.99**	**0.000**	**0.592**
veracity * condition * cognitive-ability-test score	2 ,712	2.12	0.121	0.006
condition	2 ,712	2.18	0.113	0.006
cognitive-ability-test score	**1 ,712**	**138.03**	**0.000**	**0.162**
condition * cognitive-ability-test score	2 ,712	0.35	0.704	0.001

Table S17. Summary table for the recognition statistical result in Phase 1 of Experiment 3a. Indention of an effect refers to a post-hoc analysis result.

Effect	*df*	F	*p*	η_p^2
news type	**1, 726**	**51.37**	**0.000**	**0.066**
news type * condition	2, 726	2.91	0.055	0.008
news type * cognitive-ability-test score	1, 726	2.13	0.145	0.003
news type * condition * cognitive-ability-test score	2, 726	0.94	0.392	0.003
condition	2, 726	0.24	0.784	0.001
cognitive-ability-test score	**1, 726**	**59.46**	**0.000**	**0.076**
condition * cognitive-ability-test score	**2, 726**	**3.30**	**0.037**	**0.009**
low: condition	2, 327	1.51	0.223	0.009
high: condition	2, 399	2.01	0.135	0.010
a.Inf: cognitive-ability-test score	**1, 243**	**36.22**	**0.000**	**0.130**
a.Only: cognitive-ability-test score	**1, 240**	**17.76**	**0.000**	**0.069**
con: cognitive-ability-test score	**1, 243**	**8.71**	**0.003**	**0.035**

Table S18. Summary table for the perceived accuracy result in Phase 1 of Experiment 3a. Indention of an effect refers to a post-hoc analysis result.

Effect	*df*	F	*p*	η_p^2
news type	1, 726	0.34	0.560	0.000
news type * condition	**2, 726**	**13.71**	**0.000**	**0.036**
real-related: condition	**2, 726**	**3.41**	**0.034**	**0.009**
a.Inf vs. a.Only	NA	NA	0.072	NA
a.Inf vs. con	NA	NA	0.074	NA
a.Only vs. con	NA	NA	1.000	NA
fake-related: condition	2, 726	2.02	0.133	0.006
news type * cognitive-ability-test score	1, 726	1.14	0.286	0.002
news type * condition * cognitive-ability-test score	2, 726	0.63	0.534	0.002
condition	2, 726	0.35	0.702	0.001
cognitive-ability-test score	1, 726	2.24	0.135	0.003
condition * cognitive-ability-test score	2, 726	1.13	0.325	0.003

Table S19. Summary table for the recognition statistical result in Phase 2 of Experiment 3a. Indention of an effect refers to a post-hoc analysis result.

Effect	*df*	F	*p*	η_p^2
veracity	1, 782	3.61	0.058	0.005
veracity * condition	**2, 782**	**8.08**	**0.000**	**0.020**
real: condition	2, 782	0.38	0.687	0.001
fake: condition	**2, 782**	**3.63**	**0.027**	**0.009**
a.Inf vs. a.Only	NA	NA	0.155	NA
a.Inf vs. con	**NA**	**NA**	**0.030**	**NA**
a.Only vs. con	NA	NA	1.000	NA
a.Inf: veracity	1, 263	2.51	0.115	0.009
a.Only: veracity	1, 256	1.94	0.165	0.008
con: veracity	**1, 263**	**14.71**	**0.000**	**0.053**
veracity * cognitive-ability-test score	1, 782	0.92	0.338	0.001
veracity * condition * cognitive-ability-test score	2, 782	0.61	0.544	0.002
condition	2, 782	0.89	0.409	0.002
cognitive-ability-test score	**1, 782**	**74.60**	**0.000**	**0.087**
condition * cognitive-ability-test score	2, 782	2.57	0.077	0.007

Table S20. Summary table for the perceived accuracy result in Phase 2 of Experiment 3a. Indention of an effect refers to a post-hoc analysis result.

Effect	*df*	F	*p*	η_p^2
veracity	**1, 782**	**698.48**	**0.000**	**0.472**
veracity * condition	**2, 782**	**7.69**	**0.000**	**0.019**
real: condition	2, 782	0.69	0.504	0.002
fake: condition	**2, 782**	**10.24**	**0.000**	**0.026**
a.Inf vs. a.Only	**NA**	**NA**	**0.000**	**NA**
a.Inf vs. con	**NA**	**NA**	**0.011**	**NA**
a.Only vs. con	NA	NA	0.356	NA
veracity * cognitive-ability-test score	**1, 782**	**35.56**	**0.000**	**0.043**
real: cognitive ability	**1, 782**	**6.72**	**0.010**	**0.009**
fake: cognitive ability	**1, 782**	**56.23**	**0.000**	**0.067**
low: veracity	**1, 362**	**156.21**	**0.000**	**0.301**
high: veracity	**1, 420**	**720.50**	**0.000**	**0.632**
veracity * condition * cognitive-ability-test score	2, 782	0.57	0.564	0.001
condition	**2, 782**	**6.65**	**0.001**	**0.017**
cognitive-ability-test score	**1, 782**	**40.67**	**0.000**	**0.049**
condition * cognitive-ability-test score	2, 782	1.03	0.356	0.003

Table S21. Summary table for the recognition result in Phase 1 of Experiment 3b. Indention of an effect refers to a post-hoc analysis result.

Effect	*df*	F	*p*	η_p^2
news type	**1, 689**	**12.44**	**0.000**	**0.018**
news type * condition	**2, 689**	**3.29**	**0.038**	**0.009**
real-related: condition	**2, 689**	0.51	0.600	0.001
fake-related: condition	2, 689	0.12	0.885	0.000
a.Inf: news type	1, 228	0.00	0.952	0.000
a.Only: news type	**1, 245**	**9.83**	**0.002**	**0.039**
con: news type	**1, 216**	**8.74**	**0.003**	**0.039**
news type * cognitive-ability-test score	**1, 689**	**15.34**	**0.000**	**0.022**
low: news type	1, 282	0.06	0.800	0.000
high: news type	**1, 407**	**33.89**	**0.000**	**0.077**
news type * condition * cognitive-ability-test score	2, 689	0.30	0.740	0.001
condition	2, 689	0.04	0.960	0.000
cognitive-ability-test score	**1, 689**	**76.80**	**0.000**	**0.100**
condition * cognitive-ability-test score	2, 689	0.78	0.459	0.002

Table S22. Summary table for the perceived accuracy result in Phase 1 of Experiment 3b. Indention of an effect refers to a post-hoc analysis result.

Effect	*df*	F	*p*	η_p^2
news type	**1, 689**	**15.87**	**0.000**	**0.023**
news type * condition	**2, 689**	**12.42**	**0.000**	**0.035**
real-related: condition	**2, 689**	0.89	0.410	0.003
fake-related: condition	**2, 689**	**3.55**	**0.029**	**0.010**
a.Inf vs. a.Only	**NA**	**NA**	**0.025**	**NA**
a.Inf vs. con	NA	NA	0.336	NA
a.Only vs. con	NA	NA	0.922	NA
news type * cognitive-ability-test score	1, 689	0.11	0.738	0.000
news type * condition * cognitive-ability-test score	2, 689	0.34	0.709	0.001
condition	2, 689	0.27	0.765	0.001
cognitive-ability-test score	1, 689	0.07	0.786	0.000
condition * cognitive-ability-test score	**2, 689**	**4.75**	**0.009**	**0.014**
low: condition	2, 282	2.04	0.131	0.014
high: condition	2, 407	2.72	0.067	0.013
a.Inf vs. a.Only	NA	NA	0.684	NA
a.Inf vs. con	NA	NA	0.781	NA
a.Only vs. con	NA	NA	0.062	NA
a.Inf: cognitive-ability-test score	1, 228	0.12	0.727	0.001
a.Only: cognitive-ability-test score	**1, 245**	**6.70**	**0.010**	**0.027**
con: cognitive-ability-test score	1, 216	2.59	0.109	0.012

Table S23. Summary table for the recognition result in Phase 2 of Experiment 3b. Indention of an effect refers to a post-hoc analysis result.

Effect	*df*	F	*p*	η_p^2
veracity	**1, 734**	**10.98**	**0.001**	**0.015**
veracity * condition	2, 734	1.11	0.330	0.003
veracity * cognitive-ability-test score	**1, 734**	2.12	0.146	0.003
veracity * condition * cognitive-ability-test score	**2, 734**	**3.23**	**0.040**	**0.009**
a.Inf: veracity * cognitive-ability-test score	**1, 242**	**7.17**	**0.008**	**0.029**
low: veracity	1, 98	3.67	0.058	0.036
high: veracity	1, 144	2.89	0.092	0.020
real: cognitive ability	**1, 242**	**39.69**	**0.000**	**0.141**
fake: cognitive ability	**1, 242**	**19.79**	**0.000**	**0.076**
a.Only: veracity * cognitive-ability-test score	1, 259	0.05	0.821	0.000
con: veracity * cognitive-ability-test score	1, 233	0.38	0.539	0.002
low: veracity * condition	2, 303	0.66	0.519	0.004
high: veracity * condition	**2, 431**	**5.06**	**0.007**	**0.023**
real:condition	2, 431	1.52	0.219	0.007
fake:condition	2, 431	1.49	0.227	0.007
a.Inf: veracity	1, 144	2.89	0.092	0.020
a.Only: veracity	**1, 161**	**5.94**	**0.016**	**0.036**
con: veracity	1, 126	3.74	0.055	0.029
condition	2, 734	1.40	0.248	0.004
cognitive-ability-test score	**1, 734**	**68.15**	**0.000**	**0.085**
condition * cognitive-ability-test score	2, 734	1.42	0.243	0.004

Table S24. Summary table for the perceived accuracy result in Phase 2 of Experiment 3b. Indention of an effect refers to a post-hoc analysis result.

Effect	*df*	F	*p*	η_p^2
veracity	**1, 734**	**624.81**	**0.000**	**0.460**
veracity * condition	2, 734	0.47	0.623	0.001
veracity * cognitive-ability-test score	**1, 734**	**55.51**	**0.000**	**0.070**
low: veracity	**1, 303**	**104.93**	**0.000**	**0.257**
high: veracity	**1, 431**	**771.71**	**0.000**	**0.642**
real: cognitive ability	**1, 739**	**6.81**	**0.009**	**0.009**
fake: cognitive ability	**1, 739**	**85.43**	**0.000**	**0.104**
veracity * condition * cognitive-ability-test score	2, 734	2.43	0.089	0.007
low: veracity * condition	2, 303	0.30	0.741	0.002
high: veracity * condition	**2, 431**	**3.72**	**0.025**	**0.017**
real: condition	2, 431	1.05	0.351	0.005
fake: condition	**2, 431**	**5.70**	**0.004**	**0.026**
a.Inf vs. a.Only	**NA**	**NA**	**0.002**	**NA**
a.Inf vs. con	NA	NA	0.282	NA
a.Only vs. con	NA	NA	0.376	NA
condition	**2, 734**	**4.02**	**0.018**	**0.011**
cognitive-ability-test score	**1, 734**	**56.47**	**0.000**	**0.071**
condition * cognitive-ability-test score	2, 734	1.02	0.361	0.003

Table S25. Summary table of Liner Mixed Effect Regression for the perceived accuracy result in Phase 2 of Experiment 1.

Effect	*df*	F	*p*
veracity	**1, 6**	**8.04**	**0.030**
veracity * condition	**2, 4790**	**7.17**	**0.001**
veracity * cognitive-ability-test score	**1, 4790**	**51.98**	**0.000**
veracity * condition * cognitive-ability-test score	2, 4790	1.08	0.339
condition	2, 680	1.19	0.304
cognitive-ability-test score	**1, 680**	**29.14**	**0.000**
condition * cognitive-ability-test score	2, 680	1.00	0.369

Table S26. Summary table of Liner Mixed Effect Regression for the perceived accuracy result in Phase 2 of Experiment 2.

Effect	*df*	F	*p*
veracity	**1, 6**	**25.11**	**0.002**
veracity * condition	2, 5014	1.96	0.141
veracity * cognitive-ability-test score	**1, 5014**	**158.24**	**0.000**
veracity * condition * cognitive-ability-test score	2, 5014	2.80	0.061
condition	2, 712	2.18	0.113
cognitive-ability-test score	**1, 712**	**138.03**	**0.000**
condition * cognitive-ability-test score	2, 712	0.35	0.704

Table S27. Summary table of Liner Mixed Effect Regression for the perceived accuracy result in Phase 2 of Experiment 3a.

Effect	*df*	F	*p*
veracity	**1, 6**	**14.20**	**0.009**
veracity * condition	**2, 5503**	**10.37**	**0.000**
veracity * cognitive-ability-test score	**1, 5503**	**48.03**	**0.000**
veracity * condition * cognitive-ability-test score	2, 5503	0.78	0.459
condition	**2, 782**	**6.65**	**0.001**
cognitive-ability-test score	**1, 782**	**40.64**	**0.000**
condition * cognitive-ability-test score	2, 782	1.04	0.355

Table S28. Summary table of Liner Mixed Effect Regression for the perceived accuracy result in Phase 2 of Experiment 3b.

Effect	*df*	F	*p*
veracity	**1, 6**	**12.18**	**0.013**
veracity * condition	2, 5168	0.64	0.527
veracity * cognitive-ability-test score	**1, 5168**	**75.03**	**0.000**
veracity * condition * cognitive-ability-test score	**2, 5168**	**3.28**	**0.038**
condition	**2, 734**	**4.02**	**0.018**
cognitive-ability-test score	**1, 734**	**56.47**	**0.000**
condition * cognitive-ability-test score	2, 734	1.02	0.361