Beyond Cognitive Ability: Susceptibility to Fake News Is Also Explained by Associative Inference

Sian Lee
The Pennsylvania State University
University Park, PA 16802, USA
szl43@psu.edu

Haeseung Seo
The Pennsylvania State University
University Park, PA 16802, USA
hxs378@psu.edu

Joshua P. Forrest
The Pennsylvania State University
University Park, PA 16802, USA
jzf5451@psu.edu

Dongwon Lee
The Pennsylvania State University
University Park, PA 16802, USA
dongwon@psu.edu

Jessica Strait
The Pennsylvania State University
University Park, PA 16802, USA
jls7571@psu.edu

Aiping Xiong
The Pennsylvania State University
University Park, PA 16802, USA
axx29@psu.edu

CHI '20 Extended Abstracts, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6819-3/20/04.
http://dx.doi.org/10.1145/3334480.3383077

Abstract
We conducted a preliminary online study (N=261) investigating whether people's susceptibility to fake news on social media depends on how fake news are *associated* with real news that they viewed previously, as well as individuals' cognitive ability. Across two phases, we varied the association in three between-subjects conditions, i.e., *associative inference*, *repetition*, and *irrelevant* (control). Our study results showed limited impact of association type on participants of low cognitive ability. In contrast, for participants of high cognitive ability, their discrimination of fake news from real news tended to be worse for the *associative inference* condition than for the other two conditions. Thus, our findings suggest that individuals of high cognitive ability are likely to be susceptible to form the belief of fake news, but differently from those of low cognitive ability.

Author Keywords
Fake News; Cognitive Ability; Associative Inference; False Memory

CCS Concepts
•**Human-centered computing** → User studies;

Introduction
Fake News refers to intentionally false stories or fabricated information written and published for various incentives

Figure 1: One example of associative inference type. Top two are the *AB & BC* real news presented in Phase 1, and the bottom one is the *AC* fake news presented in Phase 2.

such as political agenda or financial gains [7, 9]. In recent years, the proliferation of fake news on social media platforms has been identified as a major risk for individuals and society [20]. For instance, fake news has negatively influenced the elections in many nations [4], fostered people's bias [12], and promoted false beliefs about vaccines [6].

The issue of fake news has gained immense traction in the community of computer science and information science, which has led to a significant amount of development in machine learning models for detecting and mitigating fake news [8, 10, 19]. Besides the technical solutions, more studies have started to examine *cognitive factors* that may impact people's susceptibility to fake news, including the repetition effect that people increased their belief in repeated news headlines regardless of the legitimacy of the news [15]. Studies also showed that an individual's cognitive ability was highly correlated with his/her resistance to fake news [14, 16].

In this work, we investigate the effect of another cognitive factor, *associative inference* [5, 18], on individuals' information processing of news on social media platforms and to understand how it contributes to individuals' susceptibility to fake news. We conducted one study on Amazon Mechanical Turk (MTurk) with 261 participants, in which we examined participants' recognition and perceived accuracy of fake news as a function of how those pieces of fake news are *associated* with real news that they viewed before. Our guiding research questions (RQs) are:

RQ 1: To what extent will participants' recognition and perceived accuracy of fake news depend on how the fake news articles are associated with real news that they viewed previously, i.e., associative inference, repetition, or irrelevant (control)?

RQ 2: Will participants of high cognitive ability overcome the propensity to engage in associative inference more than participants of low cognitive ability?

Related Work

In this section, we first review prior work on associative inference, and then discuss cognitive mechanisms that contribute to individuals' susceptibility to fake news.

Background on Associative Inference. Human memory has been described as an optimization of information retrieval, which uses the statistics derived from past experience to estimate which knowledge will be currently relevant [1]. Besides allowing individuals to remember objects and events that they have actually experienced, individuals also show the ability to flexibly recombine prior details into a novel event [2, 17]. Nevertheless, such *associative inference* drawn from prior knowledge is not necessarily true. For example, Carpenter and Schacter [5] conducted four experiments showing that if participants learned direct associations between two items (AB, e.g., a person [A] with a toy [B] in a room) and then learned direct association that include one member of the previous studied pairs (BC, e.g., the toy [B] with a different person [C] in a room), they were susceptible to draw a false associative inference, AC.

With the effect of associative inference obtained in aforementioned studies, we are interested in understanding how it may impact people's information processing of news on Twitter. Specifically, hashtags have become a common tagging method to associate tweet messages [3]. Therefore, we conjecture that when an individual reads online news in tweet format, if she has been exposed to two real news associated with one common element/hashtag (AB & BC), then she is more likely to recognize and/or believe in a false associative inference (AC).

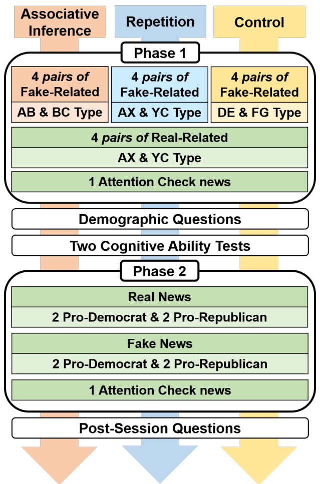


Figure 2: A flow chart shows the three between-subjects conditions, i.e., associative inference, repetition, and control (irrelevant), in the experiment.

Repetition Effect. Pennycook et al. [15] conducted online studies examining the influence of repetition on people's perceived accuracy of fake news. In their Experiments 2 and 3, participants evaluated different pieces of news in multiple stages. In stage 1, participants were asked to indicate whether they would share news headlines (half fake and half real) on social media. After a distraction in stage 2, participants rated their familiarity and perceived accuracy of real and fake news headlines in stage 3 (one half from stage 1 and the other half from a new set of headlines). Results showed that repeated headlines were rated as more "real" than novel headlines regardless of headlines' legitimacy. The increased perceived accuracy obtained with a single exposure lasted even after a week. Pennycook et al. concluded that prior exposure increased participants' perceived accuracy of fake news. Their findings suggest that individuals rely on memory or recognition-based heuristics when they make decisions about the legitimacy of news articles. Compared to repetition, an extra inference process is included in associative inference. Thus, we expect associative inference to have a stronger effect on people's susceptibility to fake news than repetition.

Cognitive Ability. A few studies have shown that individuals' cognitive ability predicts their susceptibility to fake news [14, 16]. For example, Pennycook and Rand evaluated participant's critical thinking ability with the CRT test [11]. They found that participants of low critical thinking ability were more likely to believe in fake news than participants of high critical thinking ability [16]. Murphy and her colleagues evaluated participants' cognitive ability using the Wordsum test [22], and obtained evidence suggesting that participants of high cognitive ability can overcome bias from political stance congruence [14]. Instead of the political stance congruence, we are interested in knowing whether the propensity to engage in associative inference can be overcome by individuals of high cognitive ability.

Method

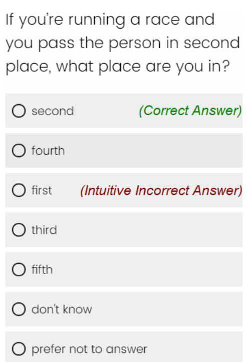
We conducted one online experiment, in which we varied the relations between real news that participants viewed initially in Phase 1 and fake news that they consumed afterwards in Phase 2 in three conditions: *associative inference*, *repetition*, or *irrelevant* (control). In addition to assess the effect of associative inference, we also examined its interaction with participants' cognitive ability.

The online experiment was designed using Qualtrics and was conducted on Amazon MTurk. All participants: (1) were at least 18 years old; (2) had completed more than 100 human intelligence tasks (HITs); (3) had at least 95% HIT approval rate; and (4) were located in the United States. The study was approved by the institutional review board (IRB) office at the authors' institution.

Materials and procedure. There were a total of 42 different news, 38 pieces of which were based on real news gathered from major news medias, including *washingtonpost.com, nytimes.com, usatoday.com, and foxnews.com*. The remaining four articles were based on fake news, which were verified by the fact-check website *snopes.com*. All news were presented in the tweet format, i.e., a snippet from real or fake news articles, with keywords (e.g., A/B/C or X/Y) listed as tweet hashtags below it. To control for potential impact from source, we also blurred the user name of each tweet (see Figure 1).

Figure 2 shows the flowchart of the study. Participants were randomly assigned into one of the three conditions. There were two phases in each condition. After an informed consent, Phase 1 started. Eight pairs of real news were pre-

(a) CRT2 measured participants' tendency to override an incorrect "gut" response with four questions. For example:



(b) Wordsum tested participants' intelligence scale of vocabulary with 10 items. Within this test, we showed participants different words in capital letters. Then, we asked participants to choose one word that comes closest to the meaning of the word in capital letters from five options. For example:

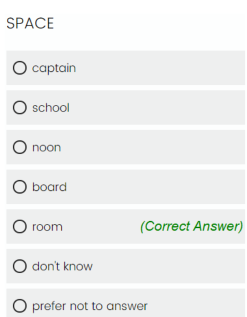


Figure 3: Descriptions of the two cognitive ability tests with an example for each.

sented in a randomized order. For each piece of news, participants were asked to view the tweet first and then answer whether they had heard about the news (i.e., "Yes", "Unsure", or "No"). Then, participants judged the accuracy of the news on a 5-point Likert scale (1 means "Very inaccurate" and 5 means "Very accurate"). For each question, we also included "Prefer not to answer" as an extra option.

After Phase 1, participants answered their demographic information, such as age, gender, and education background. Participants also completed two cognitive ability tests–CRT 2 [21] and Wordsum [14, 22] (see details in Figure 3).

After the distraction of demographic questions and cognitive ability tests, Phase 2 started. Participants answered the recognition and perceived accuracy questions for eight extra pieces of news–half fake and half real. Both fake and real news in Phase 2 were politically related. There were two pro-Republican and two pro-Democrat news for fake and real news, respectively. Each pair of real news in Phase 1 was related to one piece of the eight news in Phase 2. Thus, half of the pairs in Phase 1 were related to real news in Phase 2 (real-related), and the other half were related to fake news in Phase 2 (fake-related). All participants saw the same four real-related news pairs in Phase 1 and the same eight piece of news in Phase 2.

Furthermore, we varied the relation between the fake-related news pairs in Phase 1 and the fake news in Phase 2 across three conditions. Specifically, in the *associative inference* condition, each pair of fake-related news in Phase 1 was associated in an AB & BC type, such that the AB tweet overlapped with the BC tweet through the keyword B. Moreover, the fake news in Phase 2 was in the AC type, affording an inference with both news in Phase 1. For the *repetition* condition, each pair of fake-related news at Phase 1 was in the AX & YC type, which had no association (i.e., no common keyword between AX and YC). However, one keyword from each news article was repeated in the fake news of Phase 2, i.e., AC. For the *irrelevant* condition, news in Phase 1 were in the DE & FG type, which had neither association nor repetition either within or between phases.

In each phase, one piece of real news was used to check participants' attention [13]. For the attention checking news, we gave specific instructions for the answers of both recognition and accuracy rating questions. If one participant did not choose the specified correct answers for either attention check news article, the survey was terminated and that participant's results were excluded from data analysis.

Results

We recruited 300 MTurk workers on November 15, 2019. After removing three incomplete results, four results due to duplicate IP addresses, and another 32 participants who answered "Prefer not to answer" to at least one piece of news in Phase 2, there were 88, 85 and 88 participants for the *associative inference* (AB&BC), *repetition* (AX&YC), and *control* (DE&FG) conditions, respectively. About 59.0% of the participants were male. 74.4% of them were between 18 to 37 years old, 19.9% of them were between 38 to 57, and 5.7% were older than 58 years. 89.6% of the participants were college students or had a bachelor or higher degrees. The demographic distributions were similar across the three conditions. We paid 1.5 US dollars for participants who completed the study.

Each participant's selection of "Yes" for the recognition questions was calculated for fake news and real news of specified conditions in each phase, respectively. For the perceived accuracy rating, average ratings of each participant for fake news and real news were measured similarly. Participants' cognitive ability were categorized as *high* or

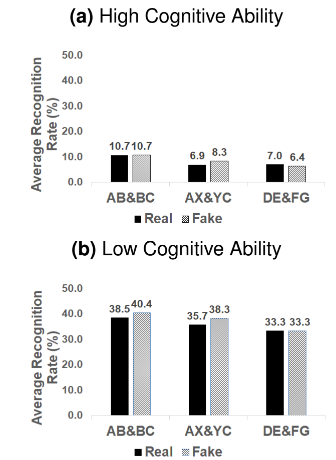


Figure 4: Average recognition rate (%) of each condition in Phase 2.

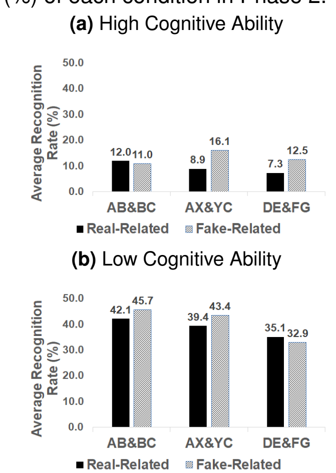


Figure 5: Average recognition rate (%) of each condition in Phase 1.

low based on their results of both CRT 2 and Wordsum tests. Specifically, for a total of 14 questions, 116 participants who got 10 or more correct answers were assigned to the group of *high* cognitive ability while the remaining 125 participants were categorized as the group of *low* cognitive ability.

For both recognition and accuracy rating measures, we report the analysis results of Phase 2 to address **RQs** first, then present the results of Phase 1. Among the 261 participants, 20 of them answered "Prefer not to answer" in Phase 1. We excluded the results from those participants from data analysis for Phase 1.

Recognition. Recognition rates of Phase 2 (see Figure 4) were entered into 2 (news legitimacy: *real*, *fake*) $\times$ 3 (association type: *AB&BC*, *AX&YC*, *DE&FG*) $\times$ 2 (cognitive ability: *high low*) mixed analysis of variances (ANOVAs) with a significance level of .05. Post-hoc tests with Bonferroni correction were performed, testing all pairwise comparisons with corrected *p* values for possible inflation.

Participants' recognition rates were similar regardless of news legitimacy (*real* vs. *fake*: 22.0% vs. 22.9%) or conditions (*AB&BC* vs. *AX&YC* vs. *DE&FG*: 25.1% vs. 22.3% vs. 20.0%), $Fs < 1.0$. However, participants of *low* cognitive ability recognized more news articles (36.6%) than participants of *high* cognitive ability (8.3%), $F_{(1,255)} = 61.08, p < .001, \eta_p^2 = .193$. No other terms were significant or approached significance, $Fs < 1.0$.

Recognition rates of Phase 1 (see Figure 5) were entered into 2 (news type: *real-related*, *fake-related*) $\times$ 3 (association type: *AB&BC*, *AX&YC*, *DE&FG*) $\times$ 2 (cognitive ability: *high*, *low*) ANOVAs. Post-hoc analyses were performed similarly as Phase 2.

Same as Phase 2, participants of *low* cognitive ability recognized more news (39.8%) than participants of *high* cognitive ability (11.3%), $F_{(1,235)} = 62.32, p < .001, \eta_p^2 = .21$. Also, participants recognized more fake-related news (26.9%) than real-related news (24.1%), $F_{(1,235)} = 6.88, p = .009, \eta_p^2 = .028$. The three-way interaction of news type $\times$ cognitive level $\times$ association type approached significance, $F_{(2,235)} = 2.74, p = .066, \eta_p^2 = .023$. Post-hoc analysis revealed that the effect of news type were similar across three association types for participants of *low* cognitive ability, $F_{(2,113)} = 1.75, p = .178, \eta_p^2 = .030$. However, for participants of *high* cognitive ability, news type $\times$ association type approached significance, $F_{(2,122)} = 2.90, p = .059, \eta_p^2 = .045$. Specifically, for participants in the *AB & BC* condition, recognition rates (real-related vs. fake-related: 12.0% vs. 11.0%) were similar, $F_{(1,48)} = 1.15, p = .704, \eta_p^2 = .003$. However, news type was significant for the *AX & YC* condition (real-related vs. fake-related: 8.9% vs. 16.1%), $F_{(1,34)} = 6.63, p = .015, \eta_p^2 = .163$ and for the *DE & FG* condition (real-related vs. fake-related: 7.3% vs. 12.5%) , $F_{(1,40)} = 6.14, p = .018, \eta_p^2 = .133$, respectively.

Accuracy Rating. Average accuracy ratings of Phase 2 (see Figure 6) were analyzed similarly as the recognition results. Participants gave higher accuracy ratings for the real news (3.32) than for the fake news (2.78), $F_{(1,255)} = 142.52, p < .001, \eta_p^2 = .359$. However, the average ratings for each condition showed no significant difference, (*AB&BC* vs. *AX&YC* vs. *DE&FG*: 3.13 vs. 3.06 vs. 2.95), $F_{(2,255)} = 1.69, p = .186, \eta_p^2 = .013$. Similar to the recognition result, participants of *low* cognitive ability gave higher ratings for news articles (3.34) than participants of *high* cognitive ability (2.76), $F_{(1,255)} = 50.63, p < .001, \eta_p^2 = .166$. The two-way interaction of news legitimacy $\times$ cognitive ability level was also significant, $F_{(1,255)} = 25.33, p < .001, \eta_p^2 = .09$. While participants of *low* cognitive abil-

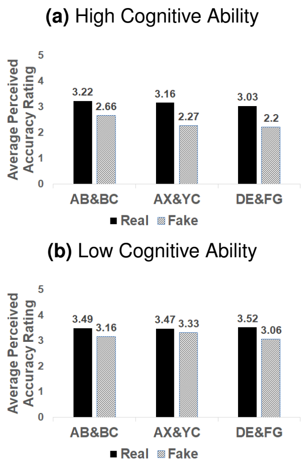


Figure 6: Average perceived accuracy rating of each condition in Phase 2.

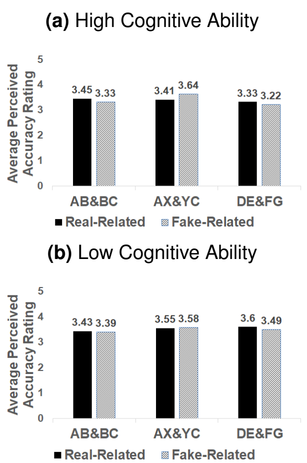


Figure 7: Average perceived accuracy rating of each condition in Phase 1.

ity gave high rating regardless of news legitimacy (real vs. fake: 3.49 vs. 3.18), participants of *high* cognitive ability showed reasonable discrimination of fake news from real news (real vs. fake: 3.14 vs. 2.38).

Critically, the better discrimination for participants of *high* cognitive ability tended to be qualified by association type, $F_{(2,255)} = 2.73, p = .067, \eta_p^2 = .021$. Post-hoc analysis revealed that across conditions, while average ratings for the real news articles were similar, $F_{(2,125)} = 1.28, p = .281, \eta_p^2 = .020$ (*AB&BC* vs. *AX&YC* vs. *DE&FG*: 3.22 vs. 3.16 vs. 3.03), average ratings for the fake news articles differed from each other, $F_{(2,125)} = 6.56, p = .002, \eta_p^2 = .095$ (*AB&BC* vs. *AX&YC* vs. *DE&FG*: 2.66 vs. 2.27 vs. 2.20). Specifically, perceived accuracy ratings of the *AB&BC* condition were higher than those of the other two conditions, $p_{adjs} \leq .008$, revealing the effect of associative inference. We repeated two keywords within the fake news in the *AX&YC* condition. However, the difference between *AX&YC* and *DE&FG* conditions was not significant, $p_{adj} = .622$, implying that the repetition effect [15] may rely on repeating the whole news article.

Accuracy rating results of Phase 1 (see Figure 7) were analyzed similarly as the recognition results. No terms were significant or approach significance, $Fs \leq 2.14$, except the two-way interaction of news type $\times$ association type, $F_{(2,235)} = 6.46, p = .002, \eta_p^2 = .052$. Post-hoc analysis showed that for *real-related* news, the average ratings were similar among three conditions (*AB & BC* vs. *AX & YC* vs. *DE & FG*: 3.44 vs. 3.48 vs. 3.47), $F < 1.0$. However, for *fake-related* news, the average ratings were different among three conditions, $F_{(2,235)} = 3.76, p = .025, \eta_p^2 = .031$. Specifically, participants' average rating in the *AX & YC* condition (3.61) was higher than that of *AB & BC* (3.36), $p_{adj} = .019$, and *DE & FG* (3.36), $p_{adj} = .016$, respectively.

Thus, those results indicated that larger accuracy rating of AB & BC fake news in Phase 2 were not due to participants' familiarity or bias to the associated real news in Phase 1.

Discussion

In this work, we took the first step to investigate the impact of associative inference and its interaction with cognitive ability on individuals' susceptibility to fake news. Our findings showed that while participants of high cognitive ability can differentiate fake news from real news, participants of low cognitive ability recognized more news articles and rated news articles as real in general. Thus, we obtained the results consistent with prior work [14, 16].

Critically, instead of overcoming the propensity of associative inference, we found that participants of high cognitive ability tended to be more susceptible to fake news when there were associative inferences between fake news and previously viewed real news than when there were no such inferences. Thus, our work contributes to research that identified high cognitive ability as a predictor of resistance to fake news [14, 16], but further suggests that individuals of high cognitive ability could be susceptible to fake news affording associative inference with prior real news.

In future work, we aim to repeat this study with a larger sample size. Moreover, we seek to find answers to explain why individuals' of high and low cognitive abilities show different susceptibility to fake news with associative inference.

Acknowledgements

We thank the anonymous reviewers for their constructive comments and the study participants for their input. This work was in part supported by the PSU SSRI seed grant and the NSF awards #1742702, #1820609, and #1915801.

REFERENCES

[1] John R Anderson and Robert Milson. 1989. Human memory: An adaptive perspective. *Psychological Review* 96, 4 (1989), 703–719.

[2] John D Bransford and Jeffery J Franks. 1971. The abstraction of linguistic ideas. *Cognitive Psychology* 2, 4 (1971), 331–350.

[3] Axel Bruns and Jean E Burgess. 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*.

[4] Carole Cadwalladr. 2017. The great British Brexit robbery: how our democracy was hijacked. *The Guardian* 7 (2017).

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

[6] Vincenzo Carrieri, Leonardo Madio, and Francesco Principe. 2019. Vaccine hesitancy and (fake) news: Quasi-experimental evidence from Italy. *Health Economics* 28 (2019), 1377–1382.

[7] Michaela Cavanagh. 2018. Climate change: 'Fake news', real fallout. (2018). `https://goo.gl/tCbwYq` Accessed: 2019-01-10.

[8] Limeng Cui, Suhang Wang, and Dongwon Lee. 2019. SAME: Sentiment-Aware Multi-Modal Embedding for Detecting Fake News. In *IEEE/ACM Int'l Conf. on Social Networks Analysis and Mining (ASONAM)*.

[9] Craig Silverman et al. 2016. Hyperpartisan Facebook pages are publishing false and misleading information at an alarming rate. (2016). `https://goo.gl/6pWtTT`

[10] Mehrdad Farajtabar, Jiachen Yang, Xiaojing Ye, Huan Xu, Rakshit Trivedi, Elias Khalil, Shuang Li, Le Song, and Hongyuan Zha. 2017. Fake news mitigation via point process based intervention. In *Proceedings of Machine Learning Research*, Vol. 70. JMLR, 1097–1106.

[11] Shane Frederick. 2005. Cognitive reflection and decision making. *Journal of Economic Perspectives* 19, 4 (2005), 25–42.

[12] Andrew Guess, Brendan Nyhan, and Jason Reifler. 2018. Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign. *European Research Council* 9 (2018).

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

[14] Gillian Murphy, Elizabeth F Loftus, Rebecca Hofstein Grady, Linda J Levine, and Ciara M Greene. 2019. False Memories for Fake News During Ireland's Abortion Referendum. *Psychological Science* 30, 10 (2019), 1449–1459.

[15] Gordon Pennycook, Tyrone Cannon, and David G Rand. 2018. Prior Exposure Increases Perceived Accuracy of Fake News. *Journal of Experimental Psychology: General* 147, 12 (2018), 1865–1880.

[16] Gordon Pennycook and David G Rand. 2018. Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition* 188 (2018), 39–50.

[17] Henry L Roediger and Kathleen B McDermott. 1995. Creating false memories: Remembering words not

presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21, 4 (1995), 803–814.

[18] Henry L Roediger III and Kathleen B McDermott. 2000. Tricks of memory. *Current Directions in Psychological Science* 9, 4 (2000), 123–127.

[19] Kai Shu, Limeng Cui, Suhang Wang, Dongwon Lee, and Huan Liu. 2019. dEFEND: Explainable Fake News Detection. In *25th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining (KDD)*.

[20] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017. Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter* 19, 1 (2017), 22–36.

[21] Keela S Thomson and Daniel M Oppenheimer. 2016. Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making* 11, 1 (2016), 99–113.

[22] David Wechsler. 2008. *Wechsler adult intelligence scale–Fourth Edition* (WAIS–IV). San Antonio, TX: Pearson.