

Can algorithmic advice override preconceived notions and the believability of fake news?

BY DORIT NEVO AND BENJAMIN D. HORNE

How Topic Novelty Impacts the Effectiveness of News Veracity Interventions

AS OTHERS HAVE, we define fake news as “intentionally” and “verifiably” false news articles that mislead readers.¹ As such, the characterization, detection, and prevention of fake news has become a top priority for preventing the spread of *misinformation* (false or inaccurate information) and *disinformation* (false information that is intended to mislead). Significant efforts in algorithmic fake-news detection have led to the development of artificial intelligence (AI) tools that provide signals and advice to news consumers to assist with fake news detection, albeit with varying effectiveness. Research surrounding this topic has predominantly focused on the design of algorithms (for instance Baly et al.;³ Cruz et al.;⁷ Hosseinimotlagh and Papalexakis;¹¹ and Bozarth, Saraf, and Budak⁵), with other work examining surrounding issues such as the impact of advice (for example, Moravec et al.¹⁶) and potential negative implications (for instance, Pennycook et al.²¹).

In this work, we seek to provide insight on the effectiveness of AI advice in terms of reader acceptance. We specifically focus on news interventions in the form of statements pertaining to the accuracy and reliability of an article, which we term news veracity statements. Twitter, for example, began using new labels and warning messages on some “Tweets containing disputed or misleading information related to COVID-19.”²³ We further examine news interventions through a specific lens: that of the novelty of the news topic. Novelty refers to the extent to which incoming information is similar to prior knowledge.^{12,28} In this article, it refers to the extent to which news readers encounter unfamiliar news. Specifically, we ask: when a novel situation arises, will interventions in the form of statements on the veracity of articles be more effective than when those same interventions are used on news articles about more familiar situations? The theoretical reasoning for this research question comes from the concept of confirmation bias, which states that people tend to process information in a way that favors their previously held beliefs. We are, therefore, interested in testing a novel news scenario for which prior beliefs are weak. We find that interventions are significantly more effective in novel news situations, implying that their use in AI tools should be focused on novel situations. An important implication for future research is that work is needed to better understand the contingencies underlying the acceptance and effectiveness of AI news interventions.

» key insights

- **When it comes to fake news interventions, one size doesn't fit all and various contingencies should be accounted for.**
- **Novelty is one influential factor; news veracity interventions are significantly more effective in novel news situations.**
- **News readers may change their reasoning about the content they read when novel news situations are being assessed.**



Fake News Interventions

Much has been written about fake news and misinformation, and many have tried to devise approaches to better inform news consumers on the veracity of the information they read. These intervention approaches can be as complex as educating readers to deeply process the information, or they can be as simple as incorporating fake news flags and warning statements.^{2,24} Unfortunately, news readers may ignore or reject these simple news veracity interventions, and their effectiveness can change across users depending on many individual factors.¹⁰ While research has shown some success with news veracity interventions,^{10,14} the factors that may lead to effective interventions are still under-explored.

One inhibitor to the acceptance of advice—algorithmic or otherwise—on the veracity of news is the fact that people tend to believe what they already believe. As previously noted, this phenomenon is called *confirmation bias* and it stems from people being overconfident in the correctness of their knowledge.¹³ With confirmation bias, people tend to interpret new information in ways that support pre-existing views while ignoring information that challenges those views.¹⁵ This, in turn, can propagate belief in fake news²⁷ and a refusal to accept advice on news veracity.

Research has shown that confirmation bias is strong when prior beliefs or knowledge are strong¹⁹ and when one is highly confident in their decision-making ability.²² In the specific context of fake news, it has further been shown that prior exposure to fake news increases their perceived accuracy.²⁰ Given these findings, it is important to develop approaches for interventions before prior beliefs are set.

Returning to our AI focus and the literature on algorithmic advice, much of the literature has interpreted this gap

as a need to develop early interventions in the sense of the speed of identifying a news article as fake and, specifically, the methods to intervene against false news items before they spread widely. This interpretation has led to many works focused on building technology for fast and automatic detection of fake news—for instance, Vicario et al.,²⁷ Horne et al.,¹⁰ and Baly et al.³ While early interventions are clearly vital in stopping exposure to false information,²⁰ we argue that not only does prior exposure to a fake news story impact intervention effectiveness, but also exposure to the news topic. Hence, the topic novelty of the news story, which has not been studied in this context, is important to understand.

This is the gap this study seeks to fill: When a novel situation arises, on which there are no prior beliefs, will veracity statements presented with news articles be more effective than when those same statements are provided for news articles about more familiar situations and events? We hypothesize that in such scenarios, the existence of confirmation bias might be weaker, thereby lowering the resistance of news readers to news veracity interventions.

The outbreak of COVID-19 offers an opportunity to test this idea. The World Health Organization (WHO) classified the outbreak as an international public health emergency in January 2020. Soon after, the WHO described information on COVID-19 as an “infodemic” due to “an over-abundance of information—some accurate and some not.”^a The high frequency of new, evolving, and sometimes conflicting information reported during the outbreak created a cognitive challenge to news consumers, who were now tasked with making

a <https://bit.ly/33kPrCf>

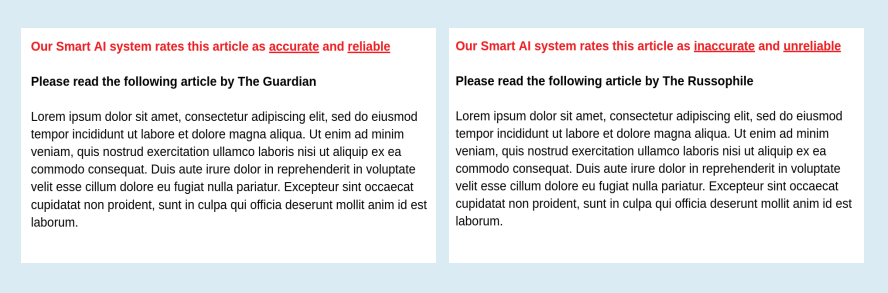
sense of information with which they did not have much prior experience.

Information overload, common in modern news environments (the current outbreak being just one example), is made worse by the uncertainty of information during an evolving crisis. Researchers in the field of crisis informatics have noted that “information voids” may exist in evolving crisis events, in cases where facts may change as new information is found, or where information is simply unknown during the ongoing event.²⁵ These voids can be challenging, particularly for consumers with a high need for cognitive closure,⁸ but they can also present an opportunity to employ effective information interventions. Unlike many news events, which may be tied to an ideological position or related to information the consumer already has experience with, reports on events such as COVID-19 are completely new to the reader, especially at the time the research was conducted. It is within this brief window of opportunity that, we argue, advice on news veracity can be most useful and help news consumers accurately assess news.

Based on the above, this study focuses on two specific questions: (1) Does the topic novelty of news impact the effectiveness of news veracity statements? (2) Does the topic novelty of news impact the consumer’s reasoning behind belief decisions? We address these questions under specific boundary conditions. First, we acknowledge that extremely novel news can pose challenges to algorithmic assessment due to limited knowledge on the topic. Our assumption is that, as was the case with the COVID-19 news that we analyzed, sufficient information had been gained to allow for veracity assessments but that this information had not yet solidified in public views. Second, we focus on scientifically verifiable news as we describe in our study section below. This allows us to accurately determine a ground truth that is more rooted in fact than opinion. Third, we only study the case of an accurate AI. In other words, our AI does not make mistakes in terms of news veracity assessments. These boundary conditions trade off an increase in internal validity for a slight decrease in external validity. We return to them when we discuss our limitations and future research plans.

To answer the above research ques-

Figure 1. Example of AI interventions in our study.



tions and to draw insights on the effectiveness of algorithmic interventions—particularly those using news veracity statements—during this window of opportunity, we used a human-subject experiment across varying conditions, changing the novelty of news articles and the availability of veracity statements. Our research method is explained below, followed by a discussion of our findings.

Method

Study design. In this work, we argue that AI interventions, in the form of news veracity statements, would be more effective in novel news situations than in familiar news situations. To test this claim about the effectiveness of AI advice in novel situations, we set up an experiment on Amazon Mechanical Turk (AMT), where we asked individuals to read a single news article and then comment on whether they believed the article and why. We chose AMT because prior work shows that it provides a good representation of the U.S. population with high generalizability.^{4,6} We conducted the experiment using a 2x2 factorial design.

The first factor, the *novelty of news*, had two levels. Familiar news articles presented both pro and con views on climate change and vaccinations (both topics have been widely covered for an extended time). Novel news articles presented information on COVID-19, a novel topic with few connections to previous news reports. The second factor, *AI intervention*, also had two levels. In the No AI condition, only the news article was presented to the respondent, while the AI Intervention condition presented one of two statements at the top of the news article: “Our smart AI system rates this article as accurate and reliable” or “Our smart AI system rates this article as inaccurate and unreliable.” This intervention is shown along with specific article examples in Figure 1. We specifically focused on advice that is framed as AI-given advice for our intervention: AI offers great speed and scalability, and there is ample work in the literature on building AI systems for news veracity detection. Each participant was assigned to one cell in our factorial design and read one randomly chosen article.

Table 1. Articles used in this study.

Source	Title	Ground Truth
Familiar News		
<i>Natural News</i>	“Climate Change Hoax Has Literally Convinced a Member of Congress that ‘the World Is Going to End in 12 Years’”	Not Credible
<i>Freedom Bunker</i>	“Fight Illness with This Ancient Immune Booster”	Not Credible
<i>Jew World Order</i>	“Greenpeace Founder: Global Warming Is a Hoax Pushed by Corrupt Scientists ‘Hooked on Government Grants’”	Not Credible
<i>The Gateway Pundit</i>	“NOAA Ruins Assertions by Unhinged Democrats that Global Warming Has Caused Increase in Hurricane Activity”	Not Credible
<i>Natural News</i>	“World Health Organization Declares Anti-Vax Movement to Be a Top ‘Global Health Threat’ Just like the Climate Change Hoax...”	Not Credible
<i>BBC</i>	“‘Completely Avoidable’ Measles Outbreak Hits 25-Year High in US”	Credible
<i>NPR</i>	“Climate Change Was the Engine That Powered Hurricane Maria’s Devastating Rains”	Credible
<i>Chicago-Sun Times</i>	“Kentucky Governor Exposed His Kids to Chickenpox instead of Getting Vaccine”	Credible
<i>NPR</i>	“New U.S. Measles Cases Break 25-Year-Old Record, Health Officials Say”	Credible
<i>Fortune</i>	“U.S. Carbon Emissions Soared in 2018. Here’s Why”	Credible
Novel News		
<i>The New York Times</i>	“Open Windows. Don’t Share Food. Here’s the U.S. Government’s Coronavirus Advice.”	Credible
<i>Reuters</i>	“World Faces Chronic Shortage of Coronavirus Protective Equipment: WHO”	Credible
<i>Breitbart</i>	“Hillary Clinton Falsely Claims to Jimmy Fallon That Trump Called Coronavirus Outbreak a ‘Hoax’”	Not Credible
<i>Natural News</i>	“Vitamin C Infusions Being Studied in China as Possible Treatment for Coronavirus-related Pneumonia”	Not Credible
<i>The Russophile</i>	“Coronavirus Hoax: Fake Virus Pandemic Fabricated to Cover-Up Global Outbreak of 5G Syndrome”	Not Credible
<i>The Russophile</i>	“Coronavirus Special Report: Worldwide Outbreaks of 5G Syndrome and 5G Flu Driving Pandemic”	Not Credible
<i>The Liberty Daily</i>	“Coronavirus: Chinese Espionage Behind Wuhan Bioweapon?”	Not Credible

Ground truth determination. In this study, ground truth was needed to determine whether an article should be classified as fake news or not and to provide the appropriate intervention. We further use this ground truth to measure the extent to which our respondents were able to correctly identify fake news articles. To determine the ground truth of articles, we used a two-phased approach, in which we first selected a source and then an article from that source. Specifically, third-party journalistic organizations, such as NewsGuard or Media Bias/Fact Check (MBFC), were used to label sources on various factors (similar to the labeling in Nørregaard et al.¹⁸) NewsGuard has journalists rate news sources based on nine weighted points, including whether the source repeatedly publishes false content, gathers

and presents information responsibly, handles the difference between news and opinion responsibly, and more. MBFC uses a similar process, but provides more news source categories, such as news sources with specific political biases or news sources that often push pseudoscience. Using these organizations, we selected a large number of sources and then searched current articles within those sources. We explain this process below.

News article selection. With a large number of sources selected, we searched for articles on our selected topics (vaccination, climate change, and COVID-19). During the article selection for vaccination and climate change, we purposely balanced between pro and con articles on each topic. We further ensured that all articles were current to reduce prior exposure to each article.

Table 2. Two-way ANOVA on agreement with ground truth.

Source	F Value	Pr > F
Novelty	0.14	0.7087
AI Intervention	8.12	0.0045
Interaction	6.55	0.0107

Next, we narrowed down the selected articles to only include those which we could confirm as true or false using journalistic and fact-checking organizations at the article level, such as Snopes, FactCheck.org, and AP Fact Check.

At the end of the process, we obtained a set of 10 articles in the familiar news condition and seven articles in the novel news condition, as shown in Table 1. Both news contexts contained articles from similar time frames in 2020 and both contexts contained articles labeled false and true. Specifically, our study ran from December 2019 to May 2020; articles in our familiar news condition were published between March 2019 and December 2019, while articles in our novel news condition were published between late February 2020 and early March 2020.

Participants. The study had 636 participants across the conditions. Each participant was paid \$0.50 for reading one article and answering two questions. We only allowed respondents from the U.S. (given the articles selected) with a HIT approval rate of at least 99%. A small-scale pilot study was performed to ensure that our chosen HIT approval rate and payment were sufficient for receiving quality responses. The responses from the pilot study were added to the main pool of responses. In

terms of demographics, respondents were between the ages of 25-33 (33.1%), 35-44 (31.9%), and 45 or older (33.5%). Just over 56% of respondents identified as males, and 58% of respondents held a Bachelor’s degree or higher. Looking at news consumption habits, respondents indicated they predominantly consumed news through websites (44.7%), social media (30.9%), TV (19.5%), and other sources (4.9%). Many respondents consumed news daily (48.9%), with another large group (34%) consuming news multiple times per day. The remaining respondents indicated they consumed news weekly (13.8%) or less (3.3%). Finally, 6.8% of respondents often shared news on social media, 55.2% sometimes shared news, and 35.5% never shared news on social media. We began collecting data in November 2019, focusing on the everyday news condition. We collected data on the emerging news condition in April 2020.

Summary of study design. As previously mentioned, we employed a 2x2 design, in which we studied two news conditions (familiar news and novel news) and two AI conditions (No AI and AI Intervention). The familiar news condition contained news articles on vaccination and climate change, both of which had been widely reported prior to this study, while the novel news condition contained news articles on COVID-19. The No AI condition provided just the article, while the AI Intervention condition displayed one of two statements at the top of the article: either “Our smart AI system rates this article as accurate and reliable” or “Our smart AI system rates this article as inac-

curate and unreliable.” Each participant read one randomly chosen article and answered two questions: 1. Do you believe the information in this news article? (Answered on a five-point scale from “Definitely yes” to “Definitely not”), and 2. Why do you or do you not believe the information in this news article? (Answered using open-ended text).

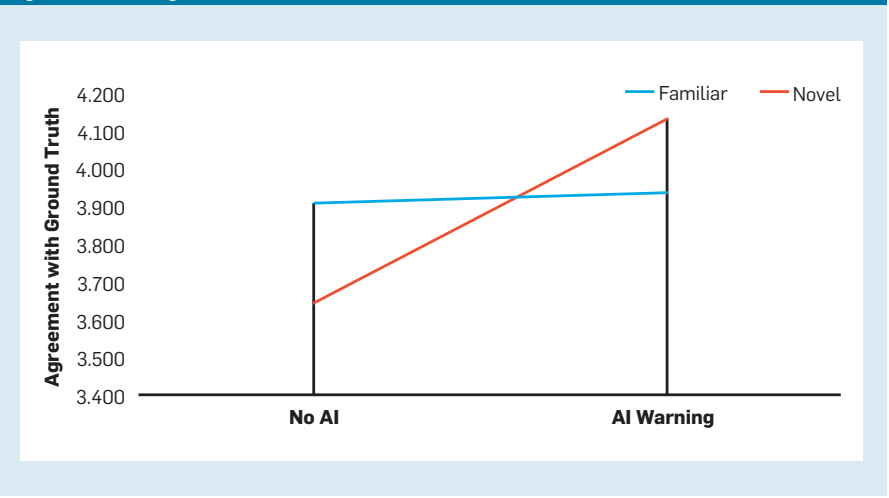
Findings

News veracity interventions are significantly more effective in novel news situations. Table 2 shows the results of a two-way ANOVA using the five-point response scale as our dependent variable and the 2x2 factorial design as the independent variables. To capture agreement with the ground truth, which is the dependent variable in this study, we first reverse-coded our five-point response scale for the false ground truth articles. Specifically, if the ground truth was “false” (that is, the article was deemed as not credible) and a respondent answered “1” (that is, I definitely do not believe the article), we reverse-coded this response as “5” to reflect full agreement with the ground truth. Consequently, the newly coded dependent variable in this analysis measures agreement with the ground truth on a five-point scale, with “5” representing “fully agree” and “1” being “fully disagree.” This dependent variable was then used in the two-way ANOVA, with novelty and AI intervention as the two independent factors.

Our results show a significant interaction effect, which reflects the effectiveness of the AI intervention under the novel news condition but not under the familiar news condition. This effect is also shown in Figure 2.

We found that when participants were given a news article from the familiar news context, the veracity statement had no significant impact on the probability that a participant correctly identified false or true news articles. In general, we note that for familiar news, participants did well in correctly identifying news veracity regardless of the AI condition they were assigned. However, when participants were given a news article from the novel news context, the news veracity statement had a significant impact on the participants’ ability to correctly identify the news article’s veracity.


Figure 2. Zooming into the interaction effect.




We followed up with pairwise comparisons of the proportion of respondents who agreed with the ground truth. Specifically, this analysis compared the proportion of participants in each condition who correctly identified false or true news articles, where the answers “Definitely yes” and “Probably yes” were combined as the participant believing an article and “Definitely not” and “Probably not” combined as the participant not believing an article. The midpoint of our scale, “might or might not,” was left as is and not used in any analysis. We found no significant difference between the two AI conditions for the familiar news group ($z=0.11$, $\text{sig}=0.36$), but saw a significant effect for the novel news group ($z=3.75$, $\text{sig}=0.00$). Looking at the other effect, when we compared the proportion of agreement with the ground truth in the familiar news group versus the novel news group without the advice of the AI, respondents in the familiar news groups had a significantly higher agreement proportion with the ground truth ($z=2.19$, $\text{sig}=0.01$). When we compared these two groups under the AI intervention condition, the effect was reversed; respondents in the novel news group performed significantly better than the familiar news group ($z=2.27$, $\text{sig}=0.01$).

Summarizing the above, our first insight is that the effect of news veracity statements (presented in this study as generated by the AI) is contingent upon the novelty of the news, having a greater impact in novel news situations.

A change in decision reasoning when novel news situations are being assessed. In addition to assessing the participant’s quantitative responses to our survey, based on the ground truth of the article, we also qualitatively assessed their reasoning through open-ended comments. To this end, authors read the comments provided and individually developed codes pertaining to the justification used—based on subjective beliefs, external knowledge, writing style of the article, and so on). After each coder reached saturation in terms of emerging codes, the authors met to discuss and develop a unified set of codes. The two authors and a research assistant then coded all responses based on this list, with disagreements resolved in follow-up discussions. For the analysis below, we used this list of justification



Many have tried to devise approaches to better inform news consumers on the veracity of the information they read.



codes to draw examples from the text on how decision reasoning varied between the conditions.

A qualitative investigation of the comments demonstrates that confirmation bias and prior opinions clearly influence belief decisions in the familiar news context (a reminder: the familiar news context had current news on vaccinations and climate change). For example, one respondent, who believed an article which we flagged as false, noted:

“Global warming and climate change are lies. These lies have been promoted since the 1800’s. NONE of their predictions have panned out. NONE of their predictions will ever pan out. What happened to the ozone hole crisis? What about the melting ice caps? They’ve grown back to be thicker than ever in recorded history. How about those rising oceans that have been predicted over and over and over and over again? Shouldn’t there be actual evidence of claims made since the 1800’s?”

This respondent is using his or her belief about global warming to form an opinion about the article. These types of responses were heavily present throughout the familiar news results, for articles on both vaccination and climate change. Reliance on prior beliefs, knowledge, and experience was also present when the participants correctly marked articles. For example, these respondents provided the following subjective reasoning for why they believed an article that was labeled as true:

“It seems to fit well with other news that I have heard.”

“I have read many news stories about outbreaks of chicken pox, measles, etc., stemming from unvaccinated individuals. My mother told me that when my older brother got mumps, her pediatrician told her to put my older sister in the room with him and expose her to the mumps so she could get it at the same time instead of later on down the road. I guess that was a popular practice in those days (40s and early 50s) but have not heard much about it being a common practice currently. But it does not surprise me and I find it plausible that the governor would do just that.”

On the other hand, in the novel news context, we found that participants used more neutral justifications rooted in the writing style, provision of supporting evidence, source credibility, and, to a much lesser extent,

reliance on prior knowledge and beliefs. Some examples of these responses include:

“Everything in the article is well written and makes sense. Does not seem biased.”

“The article makes many claims but offers no evidence of these claims. The frequent unnecessary capital letters are a big hint, and so is the lack of sources or an author’s first and last name. It is written poorly, and is meant to scare, not meant to inform. It just has an obvious fake tone to it, and makes ridiculous claims.”

“The language used and style of writing give the impression the article was not written by a highly educated or scientific person. Although there are some technical words used, the overall impression of the article is ‘amateur.’”

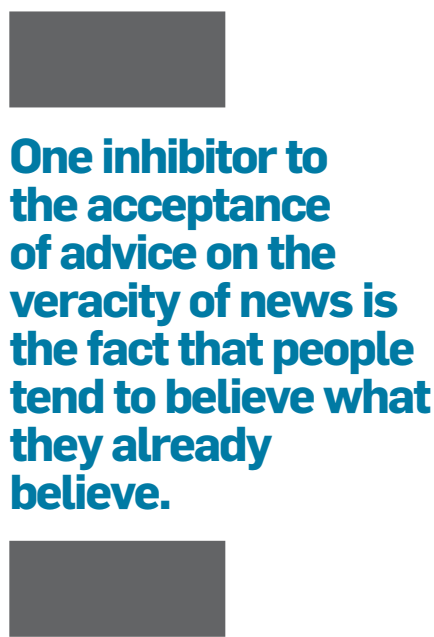
Finally, even when subjective justification was used, it was not as elaborate as the justifications that were provided in the familiar news condition.

The implicit effect of the AI. When looking just at the AI intervention condition, we found that more respondents used the AI veracity statement in their justifications in the novel news context than in the familiar news context. Using a simple keyword search, we found that 8% of participants mentioned “AI” in the familiar news condition, while more than 15% of participants mentioned “AI” in the novel news condition. An example of this type of response is:

“The smart AI system doesn’t believe it. Also I strongly doubt vitamin c would be effective at preventing or curing COVID-19.”

Beyond this direct admittance to using the AI, however, our data shows that it had an implicit effect on the justification used. Specifically, a higher proportion of respondents mentioned the word “accuracy” in their justification statement in the AI condition than in the No AI condition (17.5% vs. 3.7%, $z=4.508$ sig. <0.01). This implies that the AI has an implicit effect in guiding respondents’ justification of their news belief.

Overall, we found that the decision justification was different under different conditions. The reliance on prior knowledge and beliefs, which is indicative of confirmation bias, was more strongly present in the familiar news condition, regardless of the AI intervention. In the novel news condition, there was less evidence of such justification and more reliance on objective features of the article.



One inhibitor to the acceptance of advice on the veracity of news is the fact that people tend to believe what they already believe.

Limitations and Future Work

As one considers these core findings, it is important to note the limitations of our study and to highlight some alternative explanations to these results, specifically due to potentially confounding effects in our topic novelty test. While the topic of COVID-19 is certainly a novel one, unrelated to previous topics, it is also a crisis event with high uncertainty, including lack of expert agreement and lack of evidence for many open questions. More uniquely, COVID-19 is a crisis event where uncertainty has not been resolved for an extended period. Long-term uncertainty and lack of expert agreement on the topic likely also play a role in reducing the strength of prior opinions and knowledge. Hence, this balance of news consumers having gaps in their settled knowledge about a situation and that same situation having fact-checked information to be used in an intervention is likely a delicate one. Future work can address this gap by experimenting with other topics that are considered novel, thereby establishing that the effect is not uniquely due to the uncertainty surrounding COVID-19. Alternatively, future work can strive to establish a clear definition of when a topic can be considered novel and other factors that impact the novelty to news consumers, allowing for a more granular explanation of the found effect.

Another limitation concerns our choice of topics for this study. The three topics we selected share some common attributes, such as having their veracity anchored in science, being related to human health and wellbeing, and often being politicized. It is possible that these topics confound our results by impacting confirmation bias and opinion formation. At the same time, having this topic similarity allows us to directly compare the effect of novelty on the acceptance of the AI advice under similar conditions. Future work can explore similar studies using different news events.

Additionally, it is worth noting that not all resistance to information that contradicts one’s prior beliefs is due to confirmation bias and sometimes that resistance is good.²⁶ In our qualitative analysis, we also found instances where prior beliefs were used for correct assessments, in addition to those that were used for incorrect assessments.

Implications

The implications of this study are that AI interventions in the form of news veracity statements, which are quite common in current AI recommender tools, are not effective under all conditions. Our study resulted in two important insights for the design of algorithmic advice for fake news detection. First, one size does not fit all, meaning that there are contingencies external to the design of the recommendation tools that play a role in the effectiveness of algorithmic advice. In this work, we showed that topic novelty affects news readers' openness to accepting advice. Zooming in on how novelty affects news readers' assessment of veracity, we show that people vary the justification they use in believing specific articles under different novelty conditions. When they lack the subjective ability to assess the article, they resort to more objective heuristics, such as source credibility, writing style, or supporting evidence, to mention a few. Given that the AI commonly provides such objective assessment, we show that it is more effective in novel news conditions.

Focusing on the lack of impact of the news veracity statement in the familiar news condition, we note that our insights are important for future work in this area. Specifically, we show that different approaches and designs should be considered in different settings. In this specific study, we showed that statements about the accuracy and reliability of the articles (which are common in current literature) are not effective in familiar news settings, likely due to strong confirmation bias. Consequently, either different ways to present the statements or different statements altogether should be provided in such cases, or even a completely different design of recommendation tools. While we investigated topic novelty in this paper, we imagine there are other contingencies that should be explored to better understand how to target interventions to attain greater impact.

The implications of this work can be thought of both in terms of the effect that we find as well as the effect that we are not able to find. First, we found that in novel news situations, readers are more open to accepting algorithmic advice, in the form of news veracity statements, and this advice

helps them to justify their news assessment. For implementers of AI recommendation tools—for example, news websites and social media websites—this implies the need to target such AI interventions on novel news situations. Second, what we did not find was an effect of these interventions in familiar news conditions. This implies that it might not be economically wise to those same organizations to invest effort in providing such interventions if they are not effective. Further, an important research implication is to understand what interventions might be effective in a familiar news context. A longitudinal study that examines, for example, how a “relationship” might develop between a news consumer and an AI recommendation tool might shed light on how we can acknowledge and address confirmation bias in the design of tools.

From a theory perspective, future research efforts should focus on identifying contingencies other than the novelty of news and should collectively develop a theoretical foundation for the effectiveness of AI tools. Future design efforts should then focus on developing interventions that can overcome such contingencies. C

References

1. Allcott, H. and Gentzkow, M. Social media and fake news in the 2016 election. *Journal of Economic Perspectives* 31, 2 (2017), 211-36.
2. Auberry, K. Increasing students' ability to identify fake news through information literacy education and content management systems. *The Reference Librarian* 59, 4 (2018) 179-187.
3. Baly, R., Karadzhev, G., Alexandrov, D., Glass, J., and Nakov, P. Predicting factuality of reporting and bias of news media sources (2018). arXiv preprint arXiv:1810.01765.
4. Berinsky, A.J., Huber, G.A., and Lenz, G.S. Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis* 20, 3 (2012), 351-368.
5. Bozarth, L., Saraf, A., and Budak, C. Higher ground? How groundtruth labeling impacts our understanding of fake news about the 2016 U.S. presidential nominees. In *Proceedings of the Intern. AAAI Conf. on Web and Social Media* 14 (May 2020), 48-59.
6. Coppock, A. Generalizing from survey experiments conducted on Mechanical Turk: A replication approach. *Political Science Research and Methods* 7, 3 (2019), 613-628.
7. Cruz, A., Rocha, G., Sousa-Silva, R., and Cardoso, H.L. Team Fernando-Pessa at SemEval-2019 Task 4: Back to basics in hyperpartisan news detection. In *Proceedings of the 13th Intern. Workshop on Semantic Evaluation* (June 2019), 999-1003.
8. Hiel, A.V. and Mervielde, I. The need for closure and the spontaneous use of complex and simple cognitive structures. *The Journal of Social Psychology* 143, 5 (2003), 559-568.
9. Horne, B.D., Nørregaard, J., and Adali, S. Robust fake news detection over time and attack. *ACM Transactions on Intelligent Systems and Technology (TIST)* 11, 1 (2019), 1-23.
10. Horne, B.D., Nevo, D., O'Donovan, J., Cho, J.H., and

Adali, S. Rating reliability and bias in news articles: Does AI assistance help everyone? In *Proceedings of the International AAAI Conference on Web and Social Media* 13 (July 2019), 247-256.

11. Hosseinimotlagh, S. and Papalexakis, E.E. Unsupervised content-based identification of fake news articles with tensor decomposition ensembles. In *Proceedings of the Workshop on Misinformation and Misbehavior Mining on the Web (MIS2)*, 2018.
12. Karkali, M., Rousseau, F., Ntoulas, A., and Vazirgiannis, M. Efficient online novelty detection in news streams. In *Intern. Conf. on Web Information Systems Engineering* (2013), 57-71. Springer, Berlin, Heidelberg.
13. Koriat, A., Lichtenstein, S., and Fischhoff, B. Reasons for confidence. *J. of Experimental Psychology: Human Learning and Memory* 6, 2 (1980), 107.
14. Lutzke, L., Drummond, C., Slovic, P., and Arvai, J. Priming critical thinking: Simple interventions limit the influence of fake news about climate change on Facebook. *Global Environmental Change* 58 (2019), 101964.
15. Minas, R.K., Potter, R.F., Dennis, A.R., Bartelt, V., and Bae, S. Putting on the thinking cap: Using NeuroIS to understand information processing biases in virtual teams. *J. of Management Information Systems* 30, 4 (2014), 49-82.
16. Moravec, P., Minas, R., and Dennis, A.R. Fake news on social media: People believe what they want to believe when it makes no sense at all. Kelley School of Business research paper, 2018 (18-87).
17. Nickerson, R.S. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology* 2, 2 (1998), 175-220.
18. Nørregaard, J., Horne, B.D., and Adali, S. NELA-GT-2018: A large multi-labelled news dataset for the study of misinformation in news articles. In *Proceedings of the Intern. AAAI Conf. on Web and Social Media* 13 (July 2019), 630-638.
19. Park, J., Konana, P., Gu, B., Kumar, A., and Raghunathan, R. Information valuation and confirmation bias in virtual communities: Evidence from stock message boards. *Information Systems Research* 24, 4 (2013), 1050-1067.
20. Pennycook, G., Cannon, T.D., and Rand, D.G. Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General* 147, 12 (2018), 1865.
21. Pennycook, G., Bear, A., Collins, E.T., and Rand, D.G. The implied truth effect: Attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. *Management Science* (2020).
22. Rollwage, M., Loosen, A., Hauser, T.U., Moran, R., Dolan, R.J., and Fleming, S.M. Confidence drives a neural confirmation bias. *Nature Communications* 11, 1 (2020), 1-11.
23. Roth, Y. and Pickles, N. Updating our approach to misleading information. Twitter Blog (May 11, 2020).
24. Sharma, K., Qian, F., Jiang, H., Ruchansky, N., Zhang, M., and Liu, Y. Combating fake news: A survey on identification and mitigation techniques. *ACM Transactions on Intelligent Systems and Technology (TIST)* 10, 3 (2019), 1-42.
25. Starbird, K. How a crisis researcher makes sense of COVID-19 misinformation. *OneZero/Medium* (March 2020), <http://onezero.medium.com/reflecting-on-the-covid-19-infodemic-as-a-crisis-informatics-researcher-ce0656fa4d0a>.
26. Tappin, B.M., Pennycook, G., and Rand, D.G. Bayesian or biased? Analytic thinking and political belief updating. *Cognition* 204 (2020), 104375.
27. Vicario, M.D., Quattrociocchi, W., Scala, A., and Zollo, F. Polarization and fake news: Early warning of potential misinformation targets. *ACM Transactions on the Web (TWEB)* 13, 2 (2019), 1-22.
28. Vosoughi, S., Roy, D., and Aral, S. The spread of true and false news online. *Science* 359, 6380 (2018), 1146-1151.

Dorit Nevo (nevod@rpi.edu) is an associate professor at the Lally School of Management, Rensselaer Polytechnic Institute in Troy, NY, USA.

Benjamin D. Horne is an assistant professor in the School of Information Sciences, University of Tennessee, Knoxville in Knoxville, TN, USA.