Original Research Article

Communication Research
1–25

DOI: 10.1177/0093650220921321
journals.sagepub.com/home/crx
SAGE

Credibility Perceptions and Detection Accuracy of Fake News Headlines on Social Media: Effects of Truth-Bias and Endorsement Cues

Mufan Luo[1], Jeffrey T. Hancock[1], and David M. Markowitz[2]

Abstract

This article focuses on message credibility and detection accuracy of fake and real news as represented on social media. We developed a deception detection paradigm for news headlines and conducted two online experiments to examine the extent to which people (1) perceive news headlines as credible, and (2) accurately distinguish fake and real news across three general topics (i.e., politics, science, and health). Both studies revealed that people often judged news headlines as fake, suggesting a deception-bias for news in social media. Across studies, we observed an average detection accuracy of approximately 51%, a level consistent with most research using this deception detection paradigm with equal lie-truth base-rates. Study 2 evaluated the effects of endorsement cues in social media (e.g., Facebook likes) on message credibility and detection accuracy. Results showed that headlines associated with a high number of Facebook likes increased credibility, thereby enhancing detection accuracy for real news but undermining accuracy for fake news. These studies introduce truth-default theory to the context of news credibility and advance our understanding of how biased processing of news information can impact detection accuracy with social media endorsement cues.

Keywords

fake news, deception detection, online credibility, social media, truth-default theory

[1]Stanford University, CA, USA
[2]University of Oregon, Eugene, USA

Corresponding Author:
Mufan Luo, Department of Communication, Stanford University, Stanford, CA 94305, USA.
Email: mufanl@stanford.edu

Misinformation is not a novel phenomenon in American journalism but has risen to prominence since the 2016 U.S. presidential election. Blatantly false claims about each presidential candidate (e.g., Donald Trump was endorsed for president by Pope Francis; Hillary Clinton abused children in a pizza restaurant) have raised substantial public concerns about the election results and, more broadly, the pervasiveness of deception in public discourse. Prior work estimates that one in four American voters visited a fake news website during the 2016 election (Guess et al., 2018), and the average U.S. adult saw and remembered more than one fake story before the election (Allcott & Gentzkow, 2017). The effects of fake news were exacerbated on social media, with fake news reaching a wider audience, more rapidly, than real news on Twitter (Vosoughi et al., 2018).

Fake news is concerning to the extent that it leads to misperceptions and facilitates decision-making based on false beliefs, posing threats to individual well-being and to society (Southwell et al., 2018). From the perspective of information processing, a burgeoning literature has examined how individuals process fake news, such as how motivated reasoning can drive false beliefs about controversial political and health issues (e.g., the existence of Weapons of Mass Destruction in Iraq, Nyhan & Reifler, 2010; risks of Genetically Modified Organisms, Bode & Vraga, 2015); or how cues in the information environment affect perceptions of fake information (e.g., source and intermediary; Shen et al., 2019).

We build on message credibility research by investigating how *accurately* people distinguish between fake and real news. Detection accuracy requires examining not only the perception of message credibility, but also whether a credibility judgment is correct relative to the "ground truth" of the message (i.e., the actual message veracity based on the best available evidence; Vrij, 2008). Thus, we define fake news as blatantly false content as judged by professional fact-checkers (Pennycook & Rand, 2020). Recent work suggests that deception detection for news is a difficult task. For example, McGrew and colleagues (2018) found that students ranging from middle school to college had trouble evaluating claims and sources of social and political information online. Pennycook and Rand (2019, 2020) demonstrated that discerning real from fake news may be challenging to the extent that it correlates with individuals' analytical thinking tendency. However, several important questions remain in the context of encountering news content on social media: How accurately can people detect both real and fake news, and is their performance affected by a truth-bias?

To address these questions directly, we adapted a deception detection paradigm from interpersonal communication research and asked people to judge the accuracy of multiple genuine news headlines for which ground truth was established. This allowed us to examine both *message credibility*[1] (Appelman & Sundar, 2016), the extent to which a participant judged a news headline as real or fake, and *detection accuracy*, the extent to which a participant's judgment is accurate.

We use truth-default theory (TDT; Levine, 2014) as our theoretical framework to ground predictions about how participants will detect fake and real news, and because fake news spans multiple topics (Vosoughi et al., 2018), we evaluate message credibility and detection accuracy across three topics: politics, health, and science. Finally, because people do not make online credibility judgments in a vacuum (Metzger et al., 2010), we focus here on how cues from social media may affect detection accuracy of news. Drawing on a proposition of TDT related to *sender demeanor cues* (discussed in detail below), the second study aims to investigate the effects of two endorsement cues—the number (high vs. low) and the source (friend vs. user) of likes—on message credibility and detection accuracy. We evaluate how these cues might change credibility ratings and accuracy for real versus fake news on Facebook, the most prevalent social media platform for news (Matsa & Shearer, 2018).

Study 1

Message Credibility and Detection Accuracy for News Shared on Social Media

To assess message credibility and detection accuracy of news, we followed the tradition of lab-based interpersonal deception research paradigms. In these paradigms, participants view multiple messages, with an equal number of lies and truths, and are tasked with evaluating whether each message is deceptive or truthful (Bond & DePaulo, 2006). The ground truth for each message is known and detection accuracy is calculated by comparing people's credibility evaluation with the ground truth. A meta-analysis of interpersonal deception using this paradigm revealed that people achieve an average of 54% accuracy in detecting deception with equal lie-truth base-rates, more accurately judging truths as honest (61%) than lies as deceptive (47%; Bond & DePaulo, 2006). The finding reflects the fact that people are truth-biased and tend to infer that messages are honest, independent of message veracity.

TDT conceptualizes the truth-bias as a core feature of veracity assessments. Given its primary assumption about the default nature of the truth-bias, TDT posits that judgments for truths should be more accurate than lies (Levine et al., 1999)—a robust observation called the *veracity effect*. The truth-bias has been extensively studied in interpersonal communication and has been foundational for assessing how people process social information (Gilbert, 1991). While recent research has discussed the implication of TDT for fake news detection in social media (Clare & Levine, 2019), little empirical evidence exists. Clementson (2018), for example, had people judge whether politicians answered or deceptively evaded questions during TV interviews. Participants perceived politicians to answer the questions most of the time and more accurately detected questions answered by politicians than those they evaded. Here, we use TDT to predict how people will judge actual news headlines on social media, regardless of the headline's veracity. Research on information credibility suggests that this is a judgment of message-level credibility, defined as "an individual's judgment of the veracity of the content of communication" that is separate from judgments related to source or media credibility (Appelman & Sundar, 2016, p. 63). Given the power of the truth-bias observed in prior deception detection research, we hypothesized:

Hypothesis 1 (H1): Participants will rate news headlines in social media as credible more often, independent of the headline's actual veracity.

According to the TDT, if H1 holds, then we should also observe a veracity effect:

Hypothesis 2 (H2): Participants will judge real news headlines more accurately than fake news headlines.

Fake News Across Topics

Health, politics, and science are three news topics that draw the most public interest among American internet users (Kennedy & Funk, 2015). They are also the topics most susceptible to fake news and therefore pose serious threats to individuals and society if the news is believed (Southwell et al., 2018). Misperception about health issues may adversely affect individuals' health behaviors (Dixon & Clarke, 2013); exposure to fake science messages may prevent environmental protection and constrain policy making (see Maki et al., 2018, for a review). Prior research in media effects often uses a single message or issue as stimuli (e.g., evaluating the persuasiveness of a message; O'Keefe, 2004); this limitation suggests that it is important to take a comparative approach to examine how credibility and detection of news vary across domains or topics.

How might credibility perceptions and detection accuracy vary as a function of topic? As news media are the main source of political information (Carpini & Keeter, 1996), scant media attention to public issues may limit people's exposure and constrain their knowledge of these issue domains. Between 2007 and 2012, while the majority of the annual news coverage pertained to government and politics, less than 5% focused on health and medicine and only 1.2% focused on science and technology (Kennedy & Funk, 2015). Furthermore, perceptions of the benefits and risks of public issues are tied to social trust, especially among less knowledgeable lay people (Siegrist & Cvetkovich, 2000), and people exhibit different levels of trust in information sources across topics. People often anticipate deception from politicians and political messages (Ekman, 2009), which may be related to the media depictions about political misbehavior and deceptions (Holan, 2015). In addition, a recent national survey suggested that more people trust science information from scientists than from elected officials (Funk, 2017). We pose a research question to investigate the understudied relationship between the different topics, message credibility and detection accuracy:

Research Question 1 (RQ1): How do (a) message credibility and/or (b) detection accuracy vary across topics?

Method

This study obtained institutional review board (IRB) approval from the first author's university. We preregistered Study 1 on Open Science Framework (OSF; https://osf.io/98mz3/?view_only=ce5be533cd9149ed88692b9fbef1c4c4). We reported results of power analysis and randomization checks, and details about data exclusion in the Supplemental Online Appendix.

Participants. Participants (N = 379) were recruited from Amazon Mechanical Turk (AMT) and each was paid US$0.75 as compensation. The final sample size after data exclusion was 337. The majority of participants were White (n = 230, 68.24%), followed by 18.6% Asian (n = 62), and 7% Black (n = 25). The average age of the study sample was 35.4 years (SD = 10.31), wherein male (n = 170, 50.4%) and female (n = 167, 49.6%) were equally represented.

Experimental design and procedure. A 3 (Topic: politics, health, and science; between-subjects) × 2 (Veracity: real vs. fake; within-subjects) mixed-experimental design was employed. Participants were randomly assigned to one of three conditions of topics: politics (n = 115), health (n = 116), and science (n = 106). In each condition, participants saw five fake and five real headlines related to their topic condition in a randomized order and then rated message credibility of each news headline. Upon consenting to take part in the study, participants were told that "*all* headlines [they see] have been widely circulated on Facebook and *some* of them involve blatant fake content."

Materials. We compiled a database of real and fake headlines that were actually shared on social media to create an ecologically valid set of stimuli (see Supplemental Online Appendix). The final database involved 30 news items—five real and five fake for each of three topics (i.e., politics, health, and science). In our sample, political news items included statements about the U.S. government and politicians (e.g., "Trump's personal lawyer costing taxpayers $10,000 per hour"). Health news stories included content regarding food, medical treatment, and health behaviors (e.g., "Daily dose of diet soda tied to triple risk of deadly stroke"), whereas science news stories involved scientific discoveries and research findings (e.g., "Get ready! The brightest meteor shower in the recorded human history is happening"). News headlines were presented to participants in the format of a Facebook post, featuring a headline and a byline. Original news sources and images were omitted to focus on the effect of news veracity and content domain, with the goal of preventing other heuristics from confounding message credibility (Appelman & Sundar, 2016).

Dependent variables. We used an item of the message credibility scale to assess individuals' perceived accuracy of each news headline (Appelman & Sundar, 2016). Participants reported the extent to which they believed that the news was fake or real on a 7-point Likert-type scale from 1 = *definitely fake* to 7 = *definitely real*.

We followed prior work for calculating detection accuracy from Likert-type scale responses (Levine et al., 2010): scores of 1 to 3 were rescored as fake and were coded as accurate if the news was actually fake and as inaccurate if the news was actually real, while scores of 5 to 7 were rescored as real and were coded for accuracy in the same manner. The midpoint 4 was always coded as half accurate. Deception detection accuracy was then measured as the percentage of accurate judgments across experimental conditions.

Results

Table 1 includes all means and standard errors for message credibility and detection accuracy across conditions.

Message credibility. A one-sample t test was conducted to test H1, showing that the average message credibility ($M = 3.85$, $SD = 2.08$) was significantly below the midpoint of the 7-point scale, $t(336) = -4.01$, $p < .001$, Cohen's $d = .22$. This result suggests an overall bias toward deception rather than truth, failing to support H1. We also used the dichotomous measure of message credibility to examine the truth-bias as it is a typical approach in interpersonal deception research. Responses above 4 were coded as credible judgments. Participants were overall *deception-biased*, judging only 44.6% ($SD = 17.5\%$) of headlines as real. The deception-bias was observed for political, $t(114) = -3.34$, $p < .01$, and health news, $t(115) = -3.25$, $p < .01$, but not for science news, $t(105) = -.38$, $p = .71$. Finally, a one-way analysis of variance (ANOVA) examined the effect of topic on message credibility, which was not significant, $F(2, 334) = 2.62$, $p = .07$ (RQ1a).

Detection accuracy. A one-sample t test revealed that the average accuracy across all the news headlines was 53.5% ($SD = 17.5\%$), which was significantly above the chance rate of 50%, $t(336) = 3.65$, $p < .01$, Cohen's $d = .20$. A mixed-effects ANOVA, with veracity (fake vs. real) as a repeated measure and topic (politics, science, and health) as a between-subjects factor revealed a main effect of topic, $F(2, 334) = 31.90$, $p < .001$, $\eta^2 = .08$. A post hoc analysis revealed that people detected health news more accurately than science news, $t(334) = 2.92$, $p = .004$, but less accurately than political news, $t(334) = -5.07$, $p < .001$. Neither the main effect of veracity, $F(1, 334) = .05$, $p = .82$, $\eta^2 < .001$ (H2), nor the interaction between topic and veracity was significant, $F(2, 334) = 1.63$, $p = .20$, $\eta^2 = .01$ (RQ1b).

Discussion

Study 1 empirically examined predictions from TDT for veracity assessments of news in social media, including the truth-bias and veracity effects in detection accuracy of fake and real news. Inconsistent with expectations of a truth-bias, which is a strong effect in most interpersonal deception studies, our findings revealed a deception-bias for health and political news. People were inclined to suspect fake news when prompted to make judgments for both topics. Although these data run counter to our initial prediction, TDT also specifies several trigger events that may facilitate suspicion, such as a lack of coherence or correspondence between people's knowledge of the reality and the news content (Levine, 2014). People who have a general suspicion about messages in a given context may also abandon the truth-default state (McCornack & Levine,

Table 1. Mean Message Credibility and Detection Accuracy Across Experimental Conditions.

		Message credibility			Detection accuracy (%)		
Topics	*N*	Overall	Real	Fake	Overall	Real	Fake
(a) Study 1 (*N* = 337)							
Health	116	3.78 (0.06)	4.08 (0.08)	3.47 (0.10)	51.8 (1.6)	50.6 (2.0)	52.9 (2.4)
Politics	115	3.80 (0.09)	4.72 (0.08)	2.89 (0.10)	62.6 (1.6)	61.0 (1.9)	64.2 (2.4)
Science	106	3.97 (0.06)	4.00 (0.08)	3.95 (0.10)	45.5 (1.6)	47.7 (2.3)	43.2 (2.2)
Overall	337	3.85 (0.04)	4.27 (0.05)	3.42 (0.06)	53.5 (0.09)	53.3 (1.4)	53.7 (1.2)

			Message credibility				Detection accuracy (%)			
			Real		Fake		Real		Fake	
Topic	Source	*n*	High	Low	High	Low	High	Low	High	Low
(b) Study 2 (*N* = 622)										
Politics	Friend	94	4.82 (0.14)	4.62 (0.14)	2.88 (0.14)	2.64 (0.14)	61.2 (3.6)	59.6 (3.6)	63.3 (3.6)	70.7 (3.4)
	User	112	4.77 (0.12)	4.49 (0.14)	2.63 (0.12)	2.71 (0.12)	61.2 (3.5)	58.5 (3.3)	74.1 (3.2)	69.6 (3.3)
Health	Friend	100	3.89 (0.14)	3.19 (0.12)	3.57 (0.13)	3.48 (0.13)	45.5 (3.3)	29.0 (3.2)	51.0 (3.5)	49.5 (3.5)
	User	110	3.88 (0.15)	3.04 (0.12)	4.08 (0.15)	3.31 (0.13)	49.1 (3.4)	29.1 (2.9)	53.6 (3.3)	37.3 (3.4)
Science	Friend	96	3.60 (0.13)	3.28 (0.16)	4.38 (0.16)	3.85 (0.15)	38.0 (3.3)	35.4 (3.5)	31.8 (3.8)	46.4 (3.7)
	User	110	3.79 (0.13)	3.48 (0.16)	3.98 (0.15)	3.94 (0.15)	42.3 (3.2)	40.0 (3.5)	41.4 (3.4)	42.2 (3.2)

Note. The mean scores and standard errors of the variables are presented. Message credibility was measured by a continuous scale (1 = *definitely real* to 7 = *definitely fake*), with midpoint referring to "neither fake or real." Detection accuracy was calculated by the percentages of correct judgments.

1990; Millar & Millar, 1997). The deception-bias we observed in Study 1 may be due to general suspiciousness of news stories presented on social media (Jones, 2018). Indeed, a recent Pew report suggests that a majority (57%) of Americans believe that news on social media is largely inaccurate, compared with only 42% who believe that it is largely accurate (Matsa & Shearer, 2018). This general suspicion of news on social media may be driving the deception-bias we observed.

Furthermore, TDT predicts that a deception-bias will lead to higher fake news detection than real news detection although the effect should be small, given that the observed deception-bias was small. Indeed, Study 1 showed a null result for the effects of veracity on detection accuracy ($\eta^2 < .001$). In addition, participants achieved an average of 53.5% detection accuracy and performed best when judging political news. The results suggest poor news detection accuracy across topics, which is consistent with the 54% accuracy reported in the meta-analysis of interpersonal deception detection studies with equal truth-lie base-rates (Bond & DePaulo, 2006).

In Study 1, we did not consider variables that may undermine an individual's truth-bias and affect detection accuracy across news topics. For example, one's dispositional suspicion about news in social media might influence the truth-bias (Appelman & Sundar, 2016). Prior exposure to news headlines may also affect message credibility due to dynamics associated with the illusory truth effect, a type of processing heuristic that suggests that false information is often believed when repeated (Pennycook et al., 2018). Study 1 also did not provide cues that are typically included with news presented on social media, such as the number of likes or shares for a news headline. Fake news stories are often accompanied by these implicit endorsement cues on social media. To resolve these concerns and provide a replication, Study 2 included measures that may explain the deception-bias and its effect on detection accuracy, and examined how endorsement heuristics may affect credibility and detection accuracy.

Study 2

Credibility research suggests that endorsements can lead to online social influence because cues, such as the number and source of likes on social media, affect how people assess credibility (Metzger et al., 2010). A similar idea is proposed by TDT (Levine, 2014): *sender demeanor cues* during interpersonal communication signal believability and can affect accuracy rates depending on message veracity. Sincere demeanor cues (e.g., direct eye contact and calm appearance) can facilitate people's ability to detect honest messages accurately but undermine that ability for deceptive messages.

Drawing on these two perspectives, Study 2 aims to (1) examine whether a high number and friend-generated likes can increase message credibility compared with a low number and user-generated likes as predicted by credibility research (Metzger et al., 2010), and to (2) extend prior credibility work by examining how biased perceptions of cues translate to detection accuracy for both real and fake news, as suggested by the *sender demeanor cue* proposition of TDT.

Social Endorsement Cues and Message Credibility

When people fail to elaborate on a message, they tend to rely on simple cognitive shortcuts to process information (e.g., Petty & Cacioppo, 1986). Similarly, when making credibility judgments online, people often rely on cues that are available in the environment to conserve cognitive resources and make simpler, less effortful judgments (Metzger et al., 2010; Sundar, 2008). Prior work suggests that traces of social media (e.g., number of views and likes) are cues that people use to judge the validity or value of online information (Messing & Westwood, 2014; Spartz et al., 2017), suggesting that they will likely play a crucial role in veracity judgments as well.

Number of endorsements. Social media provide aggregated user representations (AURs; Walther & Jang, 2012), which are algorithm-generated statistics representing the number of users' evaluations (e.g., likes on Facebook, up or down votes on Reddit) and behaviors, such as watching and sharing (e.g., viewership on YouTube, shares on Facebook). "Liking" metrics are a form of AURs that generally represent the frequency of endorsement. They can affect message credibility by triggering a *bandwagon heuristic* where "people tend to believe things if others believe them" (Metzger et al., 2010, p. 429; Sundar, 2008). The effect of bandwagon heuristics on credibility perceptions has been observed in news (e.g., more votes for a news headline led to higher perceived credibility; Xu, 2013) and online health forums (e.g., the presence of a five-star rating increased the credibility of a post; Jucks & Thon, 2017). Thus, we predicted:

> **Hypothesis 3 (H3):** A high number of likes will lead to higher message credibility of news headlines than a low number of likes, regardless of actual news veracity.

Source of endorsement. Aggregated social endorsement can be derived from either *known* or *unknown* others (Metzger et al., 2010). Overall, messages endorsed by known others are considered more credible due to the *liking/agreement* heuristic, whereby people tend to trust those they feel close to and like. Similarly, the *reputation* heuristic argues that people tend to trust information from a source they recognize rather than an unfamiliar source (Metzger et al., 2010). For example, emails with political rumors or news stories recommended by friends and family were perceived as more credible than those from individuals outside of one's immediate social network (Garrett, 2011) or generated from news organizations (Turcotte et al., 2015). Alternatively, warranting theory suggests that online credibility evaluations depend on the degree to which people perceive that the information is immune to source manipulation (see DeAndrea, 2014). While aggregated Facebook likes may have warranting value because they indicate third-party evaluations, friend-generated Facebook likes may be perceived to be less susceptible to manipulation by social bots than user-generated likes, thereby making the news content more credible. Thus, we hypothesized:

Hypothesis 4 (H4): Friend-generated likes will lead to higher message credibility of news headlines than user-generated likes, regardless of the actual news veracity.

Social Endorsement Cues and Detection Accuracy of News

While research on online credibility is less concerned with message veracity and detection accuracy, TDT connects message credibility with accuracy rates. Specifically, people perceive news stories with endorsement cues as credible (e.g., many likes), especially from in-group contacts (e.g., likes from friends; Garrett, 2011; Xu, 2013). TDT's *sender demeanor cue* proposition argues that sender demeanor—either honest (i.e., being confident, friendly, and engaged) or dishonest (i.e., avoiding eye contact and appearing nervous), can affect credibility judgments but may not directly predict detection accuracy (Levine et al., 2011). Instead, the effects of demeanor cues on detection accuracy may depend on whether demeanor matches message veracity. People achieve higher accuracy rates when demeanor cues match veracity (i.e., honest demeanor matched with a truth, or dishonest demeanor matched with a lie) than when they are mismatched (i.e., honest demeanor matched with a lie, or dishonest demeanor matched with a truth).

As high-frequency endorsement cues are perceived as more trustworthy than low frequency, real news with many likes or fake news with few likes (matched veracity and endorsement) should be judged more accurately than real news with few likes or fake news with many likes (mismatched veracity and endorsement). In sum, there should be an interaction between endorsement cues and news veracity on accuracy such that if endorsement cues bias processing toward perceiving news as more real, then this will lead to reduced accuracy rates for fake news (because readers tend to rate it credible and be wrong) but higher accuracy rates for real news (because readers tend to rate it credible and be correct):

Hypothesis 5 (H5): A high number of likes will (a) increase detection accuracy for real news and (b) decrease detection accuracy for fake news.
Hypothesis 6 (H6): Friend-generated likes will (a) increase detection accuracy for real news and (b) decrease detection accuracy for fake news.

Method

We conducted the preregistration, power analysis, and randomization checks following procedures from Study 1 (see Supplemental Online Appendix and OSF).

Participants. A total of 736 participants were recruited from AMT and we ensured that these people did not complete Study 1. The final sample after data exclusion (see OSF) included 392 males and 230 females (M = 32 years old). The sample was 80% White, 7.6% Asian, and 5.8% Black. Fewer than 10% had a high school education or less, 28% had some college education, 13% had an associates degree, and 52% had a bachelor's degree and more.

Design and procedure. Study 2 employed a 3 (Topic: politics, science vs. health; between-subjects) × 2 (Veracity: real vs. fake; within-subjects) × 2 (Source of Likes: friend vs. user; between-subjects) × 2 (Number of Likes: low vs. high; within-subjects) mixed design. There were six between-subjects conditions: participants randomly assigned to each condition saw eight headlines with four different combinations of veracity and number. For example, a participant saw eight political headlines whose source of likes came from Facebook friends and then saw both real and fake news accompanied by a high number and a low number of likes. All the news headlines were drawn from the same database and were presented in the same format as Study 1 with no other contextual information (such as news source or visuals). This way, participants focused on the news content and two forms of endorsement cues.

In the "friend-likes" conditions, participants were asked to report their total number of Facebook friends and then to write three names of their close Facebook friends prior to the experiment (Figure 1). This procedure aimed to create a more personalized and realistic experimental setting, where participants would have believed that their Facebook friends actually "liked" the news headlines they saw during the experiment. The number of likes indicated in the "high number" conditions was approximately 20% of the participants' reported network size. The number of likes in the "low number" condition was either 1 or 0, determined randomly. In the "user-likes" conditions, participants were led to believe that the number of likes reflected the total times that each headline had been liked by general Facebook users (Figure 2). The high number involved a random number ranging from 50K to 100K, whereas the low number was randomly generated between five and 10. Such large differences between the high and low frequencies were used to ensure a reasonable chance of detecting effects based on subtle manipulations.

Measures

Manipulation check. We measured how high or low participants perceived the number of likes of each news headline on a Likert-type scale from 1 = *very low* to 7 = *very high*. We adapted the perceived homophily scale from McCroskey et al. (1975) to assess the extent to which participants felt that people who "liked" the news were "unlike/like me," "different from/similar to me," "do not think like me/think like me," and "concern unlike/like me" on a bipolar scale ranging from 1 to 7. Manipulation checks for number and source of likes were successful, as reported in the Online Appendix.

Dependent variables. We measured message credibility by asking participants to report the extent to which they thought the news headline was fake or real on a 7-point Likert-type scale from 1 = *definitely fake* to 7 *definitely real* (M = 3.68, SD = .71). Detection accuracy was calculated as the average percent of correct judgments (M = 49.2%, SD = 21%).

Covariates. The trust in social media scale was adapted from Tsfati and Cappella's (2003) scale of trust in media from 1 = *very little* to 7 = *very much*. Participants reported their

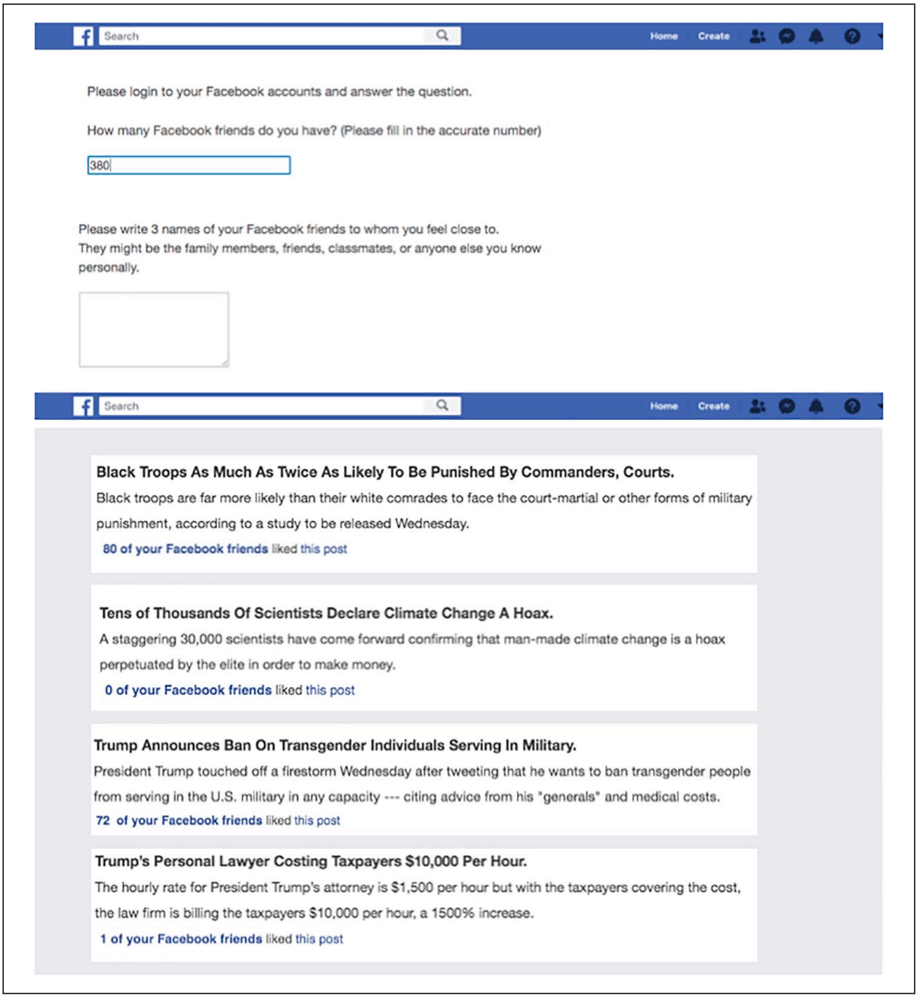

Figure 1. Political news (fake and real) accompanied with a high and a low number of likes from Facebook friends.

perceived fairness, completeness, trustworthiness, and accuracy of news posts on Facebook (Cronbach's α = .92, M = 2.80, SD = 1.31). Prior exposure to the news headlines was assessed by asking participants whether or not they recognized each headline before the study in a dichotomous scale (0 = *did not recognize*, 1 = *recognized*).

Results and Discussion

Consistent with Study 1, all means and standard errors were reported in Table 1.

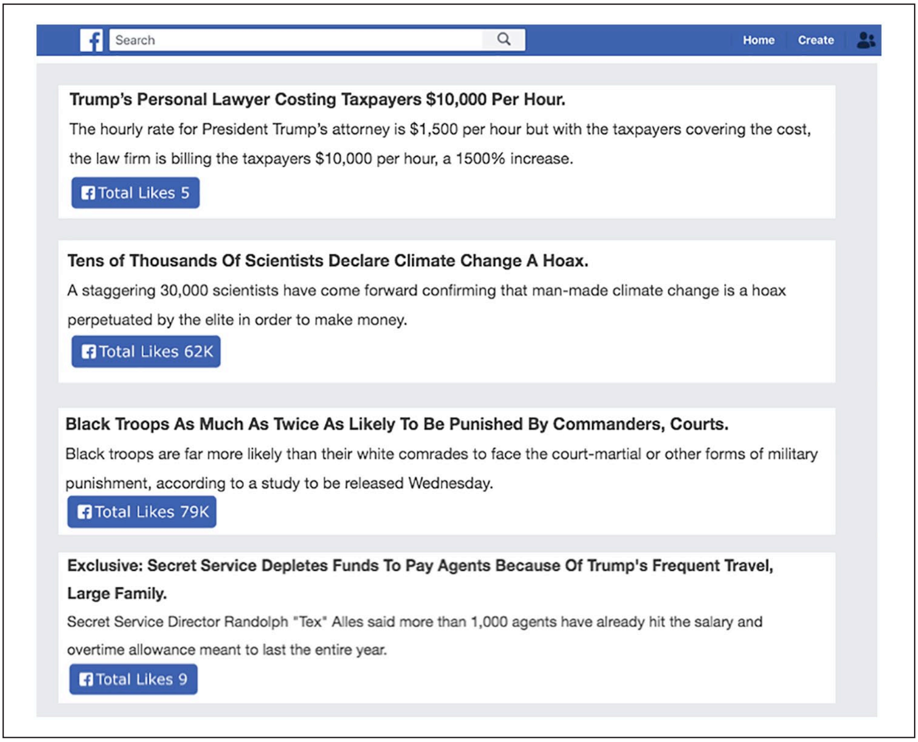

Figure 2. Political news (fake and real) accompanied with a high and a low number of likes from the internet users.

Message credibility. To test H1, a one-sample *t* test showed that the average message credibility score was significantly lower than the midpoint of the scale ($M = 3.68$, $SD = .71$), $t(621) = -11.20$, $p < .001$, Cohen's $d = .45$. Using the dichotomous measure, participants perceived 40.0% of the headlines as real ($SD = 17.9\%$), replicating the deception-bias for news headlines from Study 1. Furthermore, participants were deception-biased toward news for all three topics: politics, $t(205) = -7.47$, $p < .001$, health, $t(209) = -8.48$, $p < .001$, and science, $t(205) = -3.96$, $p < .001$.

Message credibility was then entered as a dependent variable in a 3 (Topic: politics, health, or science) × 2 (Source: friend or user) × 2 (Number: low or high) mixed ANOVA with the first two variables as between-subjects factors. There was a significant main effect of topic, $F(2, 616) = 5.11$, $p < .05$, $\eta^2 = .01$. Post hoc contrasts showed that health news was perceived as significantly less real than science news, $t(616) = -3.20$, $p < .05$, suggesting a stronger deception-bias for health news than science (RQ1a). The other pairs of topics were not significant.

In support of H3 (the bandwagon effect), the main effect of number of likes was significant, $F(1, 616) = 36.39$, $p < .001$, $\eta^2 = .03$, suggesting that headlines with

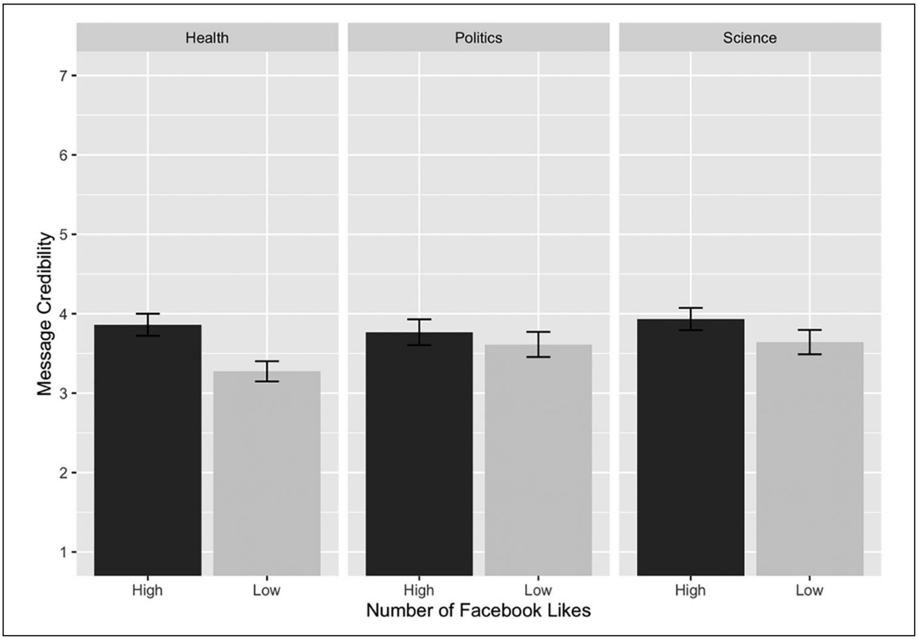

Figure 3. Message credibility by number of Facebook likes and news topics in Study 2.
Note. Error bars denote 95% confidence intervals.

many likes ($M = 3.85$, $SE = .04$) were perceived as more credible than those with few likes ($M = 3.51$, $SE = .04$), $t(616) = 6.02$, $p < .001$. The interaction effect between number and topic was significant, $F(2, 616) = 4.99$, $p = .01$, $\eta^2 = .01$. Post hoc contrasts revealed that the bandwagon effect was significant for health and science topics, ts > 3.00, ps $< .001$, but not for political news, $t(616) = 1.60$, $p = .11$ (see Figure 3).

H4 predicted that friend-generated likes can facilitate message credibility, which was not supported, $F(1, 616) < 1$, $p = .99$, suggesting that people rated friend-liked headlines ($M = 3.68$ $SE = .04$) as equally credible as user-liked headlines ($M = 3.68$, $SE = .04$). The other two-way interactions regarding source were not significant, Fs < 1, ps $> .05$.

Detection accuracy. A one-sample t test revealed that participants had an overall detection accuracy of 49.2%, which was not significantly different from the chance rate (50%), $t(621) = -.99$, $p = .32$, Cohen's $d = .04$. A mixed ANOVA tested detection accuracy with veracity (fake or real) and number (high or low), both entered as repeated measures, and topic (politics, health, or science) and source (friend or user) as between-subjects factors. There was a significant main effect of veracity, $F(1, 616) = 37.76$, $p < .001$, $\eta^2 = .01$, suggesting that people were more accurate at judging fake ($M = 52.6\%$, $SE = .01$) than real news ($M = 45.8\%$, $SE = .01$), consistent with

TDT's prediction that a deception-bias will lead to higher accuracy for fake over real news.

The main effect of topic was also significant, $F(2, 616) = 124.02, p < .001, \eta^2 = .09$ (see Table 1). Participants more accurately detected political news than health, $t(616) = 12.57, p < .01$, or science, $t(616) = 14.42, p < .001$. A significant interaction effect between veracity and topic, $F(2, 616) = 4.13, p = .02, \eta^2 = .01$, revealed that fake news was more accurately detected than real news in both political and health contexts (ts > 4.04, ps $< .001$) but not for science, $t(616) = .66, p = .51$ (RQ1b).

H5 predicted that the number of likes would be independent from detection accuracy but affect accuracy depending on news veracity. Aligning with TDT's proposition of *sender demeanor cue*, the main effect of number of likes was not significant, $F(1, 616) = .67, p = .41$, but the interaction between number and veracity was significant, $F(1, 616) = 21.07, p < .001, \eta^2 = .01$. Within-subject contrasts revealed that people more accurately detected real news when headlines had many likes ($M = 49.6\%$, $SE = .01$) than when they had few likes ($M = 41.9\%$, $SE = .01$), $t(1232) = 3.77, p < .001$. In contrast, people were less accurate when the fake headlines had a high number of likes ($M = 49.8\%$, $SE = .01$) than when they had a low number of likes ($M = 55.4\%$, $SE = .01$), $t(1232) = -2.75, p = .006$. This two-way interaction was qualified with a significant three-way interaction between number of likes, veracity, and topic, $F(2, 616) = 5.79, p < .05, \eta^2 = .004$ (Figure 4). Decomposing this interaction by topic suggested that the interaction between veracity and number of likes was only significant for health and science topics, $t(616) > 2.05, p < .05$, but not for politics, $t(616) = -.73, p = .46$. Finally, the main effect and interactions of source were not significant, Fs < 1, ps $>.05$. H6a and H6b were not supported.

Additional analyses. We first examined whether general trust in social media played a role in the deception-bias observed in Studies 1 and 2. Trust in social media was positively skewed and the average rating was lower than the midpoint ($M = 2.80$, $SD = 1.31$, median $= 2.75$, skewness $= .72$), suggesting that participants generally considered social media an untrustworthy medium for news. A logistic regression predicting truth or deception-bias, with trust in social media and demographic variables as predictors, was significant, $\chi^2(5) = 13.35, p = .02$, and revealed that the less a participant generally trusted social media, the more likely they were to be deception-biased in their judgments ($B = .13$, $SE = .07$, $p = .05$), while education decreased the likelihood ($B = -.14$, $SE = .07$, $p = .04$; see Online Appendix Table A2, Figure A1). These data suggest that low trust in social media may explain the deception-bias.

Next, we examined whether prior exposure to a headline played a role in how participants judged a headline's credibility. We used a Poisson regression to predict the number of headlines people recognized using news topic and veracity. For real news, the number of recognized news headlines was not significantly different across topics. However, for fake news, political news headlines were significantly less frequently recognized than health ($B = -.63$, $SE = .09$, $p < .001$) or science ($B = -.63$, $SE = .09$, $p < .001$). This observation that participants were less familiar with our sample of fake political headlines than fake headlines of the other topics may explain the higher

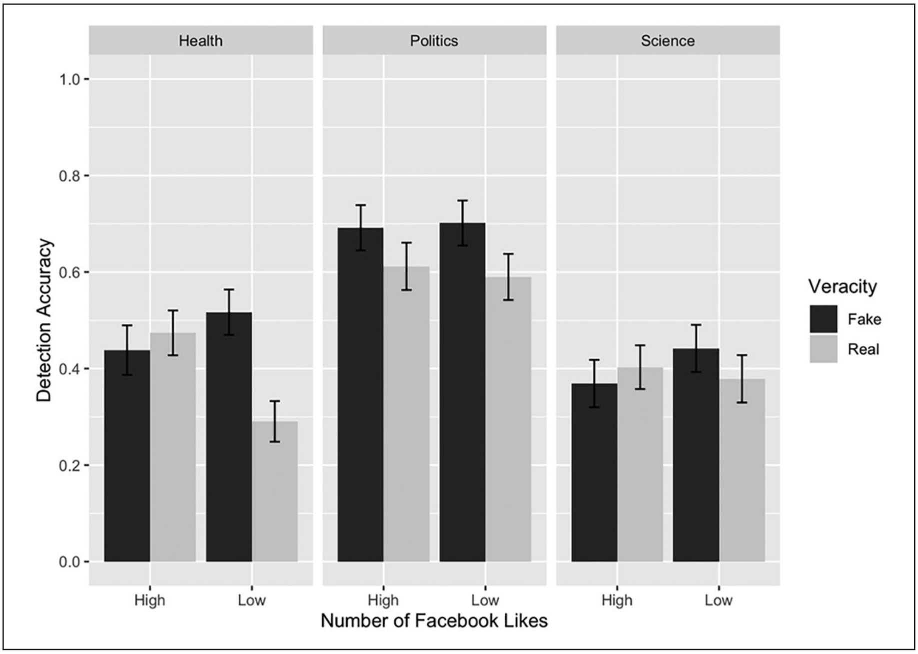

Figure 4. Detection accuracy by number of Facebook likes, news veracity, and topics in Study 2.
Note. Detection accuracy is calculated as percent of judgments that are correct. Error bars denote the 95% confidence intervals.

detection accuracy observed for political news. Indeed, a linear mixed model revealed a higher detection accuracy rate for real than fake news when headlines have been recognized before, $t(4714) = -10.68$, $p < .001$, and a higher rate for fake than real news when headlines have not been recognized, $t(4572) = 16.21$, $p < .001$ (see Supplemental Online Appendix for model information), suggesting that participants used prior exposure as a heuristic for credibility.

Finally, we examined whether the plausibility of our headlines affected credibility or detection accuracy. We followed Pennycook and Rand's (2019) approach to obtain the mean plausibility score for each news headline. We recruited out-of-sample participants from AMT ($N = 615$) to rate our sample of headlines, both real and fake, by indicating how likely a news headline was true. The mean plausibility rate was obtained by averaging participants' responses to each headline (min = 1.85, max = 5.71, median = 3.93, $SD = 1.01$). The main effect of news veracity in a two-way ANOVA was not significant, $F(1, 18) = 1.18$, $p = .29$, although a significant interaction between topic and veracity emerged, $F(2, 18) = 3.63$, $p = .05$. Pairwise contrasts showed that real news ($M = 4.74$, $SD = .72$) was significantly more plausible than fake news ($M = 2.83$, $SD = .37$) only for political news, $t(18) = 2.83$, $p = .01$, but not for the other two topics.

Did this difference in plausibility play a role in the higher detection rates for political news relative to science and health? A linear mixed model using the out-of-sample plausibility ratings for each headline to predict detection accuracy in Study 2 revealed that plausibility rating was negatively associated with fake news detection ($B = -.19$, $SE = .01$, $p < .001$) but positively correlated with real news detection ($B = .19$, $SE = .01$, $p < .001$; see Supplemental Online Appendix for model information). This set of analyses using the out-of-sample data suggests that, along with prior exposure, plausibility cues may also help explain higher detection accuracy of political news.

General Discussion

In two studies, which included nearly 1,000 participants and involved over 8,000 message credibility evaluations and detection accuracy ratings for genuine news headlines, we examined the truth-bias and veracity effects for news headlines on social media. We observed that in processing news headlines a deception-bias prevailed (H1), and this bias led fake news judgments to be more accurate than real news judgments overall (H2). Study 2 demonstrated that a high number of Facebook likes not only enhanced message credibility (and reduced deception-bias; H3), but also increased real news detection and decreased fake news detection accuracy (H5). These findings extend theoretical frameworks of both TDT (Levine, 2014) and online credibility assessment (Metzger et al., 2010) to the processing of news headlines in social media.

Deception-Bias and Veracity Effect in News on Social Media

Contrary to the robust truth-bias observation in interpersonal deception research (Levine, 2014), participants in our studies were inclined to perceive news headlines as fake more often than real. Why did we see a deception-bias, given the strength of the truth-bias in interpersonal contexts? The *truth-bias* in interpersonal communication is grounded in a fundamental assumption that communicators are cooperative partners who share accurate and relevant information (Levine, 2014). The deception-bias, in contrast, often occurs in settings that can trigger a default assessment that communicators are likely to lie (Bond et al., 2005). The *deception-bias* we observed therefore reflects a presumption of news on social media being fake. This is consistent with recent national survey findings that American adults estimate 65% of the information on social media to be misinformation (Jones, 2018), and that 57% social media news consumers perceive news on social media as largely inaccurate (Matsa & Shearer, 2018). Consistent with this inference, we found a positive association between trust in social media and message credibility in Study 2, with higher trusting individuals less likely to be deception-biased than less trusting people.

We also found that the deception-bias led to a higher detection accuracy for fake news than real news (only significant in Study 2), which is logically compatible with

TDT's proposition that the truth-bias can induce a veracity effect. This finding not only extends TDT's *veracity effect* to a news setting, but also extends prior fake news detection studies that aggregated the detection accuracy of real and fake news (e.g., media truth discernment; Pennycook & Rand, 2019, 2020) by revealing distinct detection accuracy rates for each type. This latter point is an important methodological requirement noted by Levine (2014). Without isolating accuracy by veracity, we would have made an overly simple conclusion that people's ability to judge news is approximately 51%, which obscures the fact that fake (52.6%) and real news (45.8%) detection accuracies were different. The reverse of the veracity effect, therefore, highlights a presumption of deception in news for the data in this article.

The results also have implications for news detection in the media ecosystem where the prevalence of real news is usually much higher than fake news. The deception-bias may lead to more accurate judgments for fake news but lower accuracy for real news. This observation is consistent with the Park–Levine Probability Model (Park & Levine, 2001), which shows that the truth-bias and the ratio of truth to deception can predict accuracy rates, but here we extend it to the detection of news. The finding is concerning as it indicates that the average detection accuracy in the real world may be worse than 51%—the rate observed in our study where the ratio of fake to real news was 50:50—given people's greater exposure to real than fake news. As Allcott et al. (2019) have shown, fake news exposure appears to be declining on Facebook, suggesting that people are less likely to be exposed to fake news than in the past several years. As the Park-Levine Probability Model predicts, if individuals continue to demonstrate a deception-bias on social media, their overall accuracy will decline as fake news exposure declines. Future research needs to examine the role of base-rates of fake and real news in detection accuracy. It is important to note, however, that the prior implication is based on the deception-bias observed in the research lab, unlike everyday media consumption where no prompt would be made through task instructions or measurement (Levine, 2018). Therefore, the truth-bias may still be present for news headlines outside the lab, which would lead to higher overall detection accuracy (see the section "Limitations").

Linking Credibility Heuristic Cues to Detection Accuracy

Study 2 integrates ideas from TDT and the heuristic approach to argue that message credibility and detection accuracy may be influenced by endorsement cues in social media. First, our data confirmed the bandwagon effect on credibility perceptions (e.g., Jucks & Thon, 2017; Xu, 2013). Aggregate endorsement ratings in a statistical form can influence perceptions of news credibility, with a high number of likes leading to higher credibility and a weaker deception-bias. Second, our data support the prediction of TDT's *sender demeanor cue* proposition in which cues may directly affect credibility perceptions but affect accuracy, depending on whether they are matched with news veracity (Levine et al., 2011).

This prediction from TDT maps onto key insights from credibility research: Credibility heuristics can induce believability in mediated communication that affects detection accuracy of news. One practical implication of this finding is for educational practitioners who should consider integrating the strategies about how to evaluate social endorsement cues into the curriculum of online credibility assessment, given its effectiveness on online reasoning among college students (McGrew et al., 2018).

The endorsement finding also adds some urgency for social media platforms to prevent fake news content from being shared and endorsed. Our findings suggest that Facebook likes can undermine accuracy for detecting fake news by as much as 6% when the number of likes increase from a few to many. If a total of 156 fake news stories were circulated on Facebook right before the election, as estimated by Allcott and Gentzkow (2017), then people would have detected approximately 9 fewer fake news stories (5.6% fewer, or 77 rather than 86 stories) because these news stories received a high number of Facebook likes, though the effect is likely moderated by topic. These data also suggest that detecting and removing follower bots, which manipulate endorsement cues to create the illusion of popularity (Lazer et al., 2018), should be a priority for improving detection accuracy. Given limits for news validation from third-party fact-checkers, our findings highlight the necessity to fact-check news that receive a high number of social endorsement as these posts may have outsized impacts on people's credibility perceptions and ultimately detection accuracy if they are fake.

Our data suggested that friend-generated likes did not stimulate higher credibility perceptions than user-generated likes (with a very small effect size, $\eta^2 < .001$). The nonsignificant finding may be due to the power of the bandwagon heuristic. Research has shown that bandwagon effects appear to be so powerful that they can overshadow other heuristics such as the expertise heuristic (Sundar et al., 2007). Furthermore, due to privacy concerns and methodological constraints, we did not display the names or profile pictures of participants' real friends alongside the news headlines, which may make the meaning of aggregate endorsement rating from "Facebook friends" ambiguous and unreal to receivers (Turcotte et al., 2015). Participants may have perceived that friend-generated likes were controlled by the experimenters and perceived the associated news headlines as low in warranting value (DeAndrea, 2014).

Topical Differences in Message Credibility and Detection Accuracy

Our studies revealed several distinct dynamics that suggest fake news is not perceived the same across topics. For example, the bandwagon effect occurred among health and science topics but not for politics. That may have occurred because perceptions of political news are closely tied to existing political beliefs (Giner-Sorolila & Chaiken, 1997) and may be less influenced by social heuristic cues than motivated reasoning (Nyhan & Reifler, 2010), compared with the other topics.

Both studies indicated that political news was more accurately detected than health or science news. We identified two possible mechanisms for this effect. The first is information repetition as a processing fluency heuristic (e.g., Pennycook et al., 2018) in which people perceive familiar information as more credible, which in turn can lead to increased detection accuracy for real news. Our data revealed that participants previously encountered fewer of the political fake headlines than those about health or science. This finding was consistent with a recent observational study on Twitter

(Grinberg et al., 2019) where only 1% of individuals accounted for 80% of exposure of fake news sources in the 2016 U.S. presidential election. Future work will need to establish whether this finding holds with a broader sample of fake news headlines on social media. Other types of fluency heuristics in news headlines may be considered in future research, such as communication frames and the use of jargon (Bullock et al., 2019; Shulman & Sweitzer, 2018). A second possible mechanism is the plausibility of the headlines—an important message feature of fake news (Pennycook & Rand, 2019). Using out-of-sample data, we found that our political fake news headlines were perceived as less plausible overall than fake news from science and health topics. Whether plausibility differences across news topics sustain in the natural news ecosystem is an important question that warrants future examination.

Limitations and Future Research

Our emphasis on assessing how people process genuine fake and real news from a deception detection paradigm prioritized experimental control over external validity. We followed the tradition of interpersonal deception detection research to prompt participants to make explicit fake-real judgments about news veracity (Levine, 2018), with minimal contextual information that is often available in actual social media platforms (e.g., source, visuals, and comments). This approach may limit the generalizability of our conclusions in two aspects: (1) people may be less suspicious about text-based headlines in actual social media platforms because prompted judgments tend to be more biased toward deception (Clare & Levine, 2019), and (2) the real-world news detection may also be more accurate because people can rely on a wide range of context-level heuristics (e.g., source, visuals) to make credibility judgments.

Other limitations concerning external validity include (1) the message stimuli, which are only a small sample of all possible fake and real news that people might encounter, and (2) the manipulation of "a high number of likes" from "Facebook friends," which may have lacked experimental realism. Future work should evaluate a broader set of news headlines and consider the believability of the number of likes manipulation.

Finally, our measures were limited by (1) a single-item scale to measure message credibility, which, while consistent with the deception detection paradigm, prevents assessments of perceived authenticity and believability (Appelman & Sundar, 2016), and (2) we did not collect data on several individual differences in Study 2, such as political orientation and social media engagement (Shen et al., 2019), which can affect information processing. As partisanship can affect individuals' beliefs about public issues across topic (Matsa & Shearer, 2018; Nyhan & Reifler, 2010) and exposure to fake news on social media (Grinberg et al., 2019), political orientation should be considered in future research.

Conclusion

We adopted a deception detection paradigm from interpersonal deception research to examine both the message credibility and detection accuracy of fake and real news presented in social media. The findings reveal a deception-bias (i.e., individuals perceive news as fake most of the time) and a near-chance accuracy rate (approximately, 51% across the two studies) for detecting fake and real news across political, health, and science headlines. On social media, a high number of likes not only facilitated beliefs that news headlines were real or credible, as suggested by online credibility research, but also increased accuracy rates for real news and undermined accuracy for fake news, which is consistent with TDT. Our findings highlight the urgency of empowering individuals to assess both news veracity and endorsement cues appropriately on social media.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: The authors received a grant from the Stanford Cyber Policy Center.

ORCID iDs

Mufan Luo https://orcid.org/0000-0003-0762-9058
David M. Markowitz https://orcid.org/0000-0002-7159-7014

Supplemental Material

Supplemental material for this article is available online.

Note

1. Message credibility is defined as a multidimensional construct involving perceived accuracy, authenticity, and believability (Appelman & Sundar, 2016). The current study focuses on perceived accuracy.

References

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*, 211–236. https://doi.org/10.1257/jep.31.2.211

Allcott, H., Gentzkow, M., & Yu, C. (2019). Trends in the diffusion of misinformation on social media. *Research & Politics*, *6*, 1–8. https://doi.org/10.1177/2053168019848554

Appelman, A., & Sundar, S. S. (2016). Measuring message credibility: Construction and validation of an exclusive scale. *Journalism & Mass Communication Quarterly*, *93*, 59–79. https://doi.org/10.1177/1077699015606057

Bode, L., & Vraga, E. K. (2015). In related news, that was wrong: The correction of misinformation through related stories functionality in social media. *Journal of Communication*, *65*, 619–638. https://doi.org/10.1111/jcom.12166

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, *10*, 214–234. https://doi.org/10.1207/s15327957pspr1003_2

Bond, G. D., Malloy, D. M., Arias, E. A., Nunn, S. N., & Thompson, L. A. (2005). Lie-biased decision making in prison. *Communication Reports*, *18*, 9–19. https://doi.org/10.1080/08934210500084180

Bullock, O. M., Colón Amill, D., Shulman, H. C., & Dixon, G. N. (2019). Jargon as a barrier to effective science communication: Evidence from metacognition. *Public Understanding of Science*, *28*, 845–853. https://doi.org/10.1177/0963662519865687

Carpini, M. X. D., & Keeter, S. (1996). *What Americans know about politics and why it matters*. Yale University Press.

Clare, D. D., & Levine, T. R. (2019). Documenting the truth-default: The low frequency of spontaneous unprompted veracity assessments in deception detection. *Human Communication Research*, *45*, 286–308. https://doi.org/10.1093/hcr/hqz001

Clementson, D. E. (2018). Truth bias and partisan bias in political deception detection. *Journal of Language and Social Psychology*, *37*, 407–430. https://doi.org/10.1177/0261927X17744004

DeAndrea, D. C. (2014). Advancing warranting theory. *Communication Theory*, *24*, 186–204. https://doi.org/10.1111/comt.12033

Dixon, G., & Clarke, C. (2013). The effect of falsely balanced reporting of the autism–vaccine controversy on vaccine safety perceptions and behavioral intentions. *Health Education Research*, *28*, 352–359. https://doi.org/10.1093/her/cys110

Ekman, P. (2009). *Telling lies: Clues to deceit in the marketplace, politics, and marriage* (Revised ed.). W.W. Norton.

Funk, C. (2017). *Mixed messages about public trust in science*. Pew Research Center.

Garrett, R. K. (2011). Troubling consequences of online political rumoring. *Human Communication Research*, *37*, 255–274. https://doi.org/10.1111/j.1468-2958.2010.01401.x

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*, *46*, 107–119. https://doi.org/10.1037/0003-066X.46.2.107

Giner-Sorolila, R., & Chaiken, S. (1997). Selective use of heuristic and systematic processing under defense motivation. *Personality and Social Psychology Bulletin*, *23*, 84–97. https://doi.org/10.1177/0146167297231009

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, *363*, 374–378. https://doi.org/10.1126/science.aau2706

Guess, A., Nyhan, B., & Reifler, J. (2018). *Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign*. European Research Council.

Holan, A. D. (2015, December). All politicians lie: Some lie more than others. *The New York Times*. https://www.nytimes.com/2015/12/13/opinion/campaign-stops/all-politicians-lie-some-lie-more-than-others.html

Jones, M. J. (2018, June). Americans: Much misinformation, bias, inaccuracy in news. *Gallup*. https://news.gallup.com/opinion/gallup/235796/americans-misinformation-bias-inaccuracy-news.aspx

Jucks, R., & Thon, F. M. (2017). Better to have many opinions than one from an expert? Social validation by one trustworthy source versus the masses in online health forums. *Computers in Human Behavior*, *70*, 375–381. https://doi.org/10.1016/j.chb.2017.01.019

Kennedy, B., & Funk, C. (2015). *Science and health interest linked to Americans' gender, age and personality*. Pew Research Center Science & Society.

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., . . . Zittrain, J. L. (2018). The science of fake news. *Science*, *359*, 1094–1096. https://doi.org/10.1126/science.aao2998

Levine, T. R. (2014). Truth-default theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology*, *33*, 378–392. https://doi.org/10.1177/0261927X14535916

Levine, T. R. (2018). Ecological validity and deception detection research design. *Communication Methods and Measures*, *12*, 45–54. https://doi.org/10.1080/19312458.2017.1411471

Levine, T. R., Park, H. S., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect." *Communication Monographs*, *66*, 125–144. https://doi.org/10.1080/03637759909376468

Levine, T. R., Serota, K. B., Shulman, H., Clare, D. D., Park, H. S., Shaw, A. S., . . . Lee, J. H. (2011). Sender demeanor: Individual differences in sender believability have a powerful impact on deception detection judgments. *Human Communication Research*, *37*, 377–403. https://doi.org/10.1111/j.1468-2958.2011.01407.x

Levine, T. R., Shaw, A. S., & Shulman, H. (2010). Assessing deception detection accuracy with dichotomous truth–lie judgments and continuous scaling: Are people really more accurate when honesty is scaled? *Communication Research Reports*, *27*, 112–122. https://doi.org/10.1080/08824090903526638

Maki, A., Carrico, A. R., & Vandenbergh, M. P. (2018). Doing the wrong things for the right reasons: How environmental misinformation affects environmental behavior. In B. G. Southwell, E. A. Thorson & L. Sheble (Eds.), *Misinformation and mass audiences* (pp. 177–191). University of Texas Press.

Matsa, K. E., & Shearer, E. (2018). *News use across social media platforms 2018*. Pew Research Center.

McCornack, S. A., & Levine, T. R. (1990). When lovers become leery: The relationship between suspicion and accuracy in detecting deception. *Communication Monographs*, *57*, 219–230. https://doi.org/10.1080/03637759009376197

McCroskey, J. C., Richmond, V. P., & Daly, J. A. (1975). The development of a measure of perceived homophily in interpersonal communication. *Human Communication Research*, *1*, 323–332. https://doi.org/10.1111/j.1468-2958.1975.tb00281.x

McGrew, S., Breakstone, J., Ortega, T., Smith, M., & Wineburg, S. (2018). Can students evaluate online sources? Learning from assessments of civic online reasoning. *Theory & Research in Social Education*, *46*, 165–193. https://doi.org/10.1080/00933104.2017.1416320

Messing, S., & Westwood, S. J. (2014). Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news online. *Communication Research*, *41*, 1042–1063. https://doi.org/10.1177/0093650212466406

Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication*, *60*, 413–439. https://doi.org/10.1111/j.1460-2466.2010.01488.x

Millar, M. G., & Millar, K. U. (1997). The effects of cognitive capacity and suspicion on truth bias. *Communication Research*, *24*, 556–570. https://doi.org/10.1177/009365097024005005

Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, *32*, 303–330. https://doi.org/10.1007/s11109-010-9112-2

O'Keefe, D. J. (2004). Trends and prospects in persuasion theory and research. In J. S. Seiter & R. H. Gass (Eds.), *Readings in persuasion, social influence, and compliance gaining* (pp. 31–43). Pearson/Allyn and Bacon.

Park, H. S., & Levine, T. (2001). A probability model of accuracy in deception detection experiments. *Communication Monographs*, *68*, 201–210. https://doi.org/10.1080/03637750128059
Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, *147*, 1865–1880. https://doi.org/10.1037/xge0000465
Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, *188*, 39–50. https://doi.org/10.1016/j.cognition.2018.06.011
Pennycook, G., & Rand, D. G. (2020). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of Personality*, *88*, 185–200. https://doi.org/10.1111/jopy.12476
Petty, R. E., & Cacioppo, J. T. (1986). Message elaboration versus peripheral cues. In R. E. Petty & J. T. Cacioppo (Eds.), *Communication and persuasion* (pp. 141–172). Springer.
Shen, C., Kasra, M., Pan, W., Bassett, G. A., Malloch, Y., & O'Brien, J. F. (2019). Fake images: The effects of source, intermediary, and digital media literacy on contextual assessment of image credibility online. *New Media & Society*, *21*, 438–463. https://doi.org/10.1177/1461444818799526
Shulman, H. C., & Sweitzer, M. D. (2018). Advancing framing theory: Designing an equivalency frame to improve political information processing. *Human Communication Research*, *44*, 155–175. https://doi.org/10.1093/hcr/hqx006
Siegrist, M., & Cvetkovich, G. (2000). Perception of hazards: The role of social trust and knowledge. *Risk Analysis*, *20*, 713–720. https://doi.org/10.1111/0272-4332.205064
Southwell, B. G., Thorson, E. A., & Sheble, L. (2018). *Misinformation and mass audiences*. University of Texas Press.
Spartz, J. T., Su, Y.-F., Griffin, R., Brossard, D., & Dunwoody, S. (2017). YouTube, social norms and perceived salience of climate change in the American mind. *Environmental Communication*, *11*, 1–16. https://doi.org/10.1080/17524032.2015.1047887
Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. In M. J. Metzger & A. J. Flanagin (Eds.), *Digital media, youth, and credibility* (pp. 72–100). MIT Press.
Sundar, S. S., Knobloch-Westerwick, S., & Hastall, M. R. (2007). News cues: Information scent and cognitive heuristics. *Journal of the American Society for Information Science and Technology*, *58*, 366–378. https://doi.org/10.1002/asi.20511
Tsfati, Y., & Cappella, J. N. (2003). Do people watch what they do not trust?: Exploring the association between news media skepticism and exposure. *Communication Research*, *30*, 504–529. https://doi.org/10.1177/0093650203253371
Turcotte, J., York, C., Irving, J., Scholl, R. M., & Pingree, R. J. (2015). News recommendations from social media opinion leaders: Effects on media trust and information seeking. *Journal of Computer-Mediated Communication*, *20*, 520–535. https://doi.org/10.1111/jcc4.12127
Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*, 1146–1151. https://doi.org/10.1126/science.aap9559
Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). John Wiley.
Walther, J. B., & Jang, J. (2012). Communication processes in participatory websites. *Journal of Computer-Mediated Communication*, *18*, 2–15. https://doi.org/10.1111/j.1083-6101.2012.01592.x
Xu, Q. (2013). Social recommendation, source credibility, and recency: Effects of news cues in a social bookmarking website. *Journalism & Mass Communication Quarterly*, *90*, 757–775. https://doi.org/10.1177/1077699013503158

Author Biographies

Mufan Luo is a PhD candidate researches psychological processes in social media, including social media use and well-being, perceptions and detection of misinformation, and information-sharing processes.

Jeffrey T. Hancock focuses on understanding psychological and interpersonal processes in social media.

David M. Markowitz examines how language reflects social and psychological dynamics, such as deception, persuasion, and status.