Article

Communication Research
1–28

DOI: 10.1177/00936502211073398
journals.sagepub.com/home/crx
SAGE

Reading, Commenting and Sharing of Fake News: How Online Bandwagons and Bots Dictate User Engagement

Maria D. Molina[1], Jinping Wang[2], S. Shyam Sundar[3], Thai Le[3], and Carlina DiRusso[3]

Abstract

Do social media users read, comment, and share false news more than real news? Does it matter if the story is written by a bot and whether it is endorsed by many others? We conducted a selective-exposure experiment ($N = 171$) to answer these questions. Results showed that real articles were more likely to receive "likes" whereas false articles were more likely to receive comments. Users commented more on a bot-written article when it received fewer likes. We explored the psychological mechanisms underlying these findings in Study 2 ($N = 284$). Data indicate that users' engagement with online news is largely driven by emotions elicited by news content and heuristics triggered by interface cues, such that curiosity increases consumption of real news, whereas uneasiness triggered by a high number of "likes" encourages comments on fake news.

Keywords

engagement, social media, misinformation, automated journalism, heuristics

Dissemination of false information has received major attention since the 2016 presidential elections when a Buzzfeed article revealed that false articles generated more engagement on Facebook (shares, reactions, and comments) than articles reported by

[1]Michigan State University, East Lansing, MI, USA
[2]The Chinese University of Hong Kong, Hong Kong SAR, China
[3]Penn State University, University Park, PA, USA

Corresponding Author:
Maria D. Molina, Department of Advertising and Public Relations, Michigan State University, 404 Wilson Rd, East Lansing, MI 48824-1212, USA.
Email: molinad2@msu.edu

professional journalists (Silverman, 2016). In an effort to understand user engagement with false information on social media, scholars from a variety of fields have investigated the reasons behind its dissemination, identifying pre-existing attitudes and motivated reasoning as important predictors (Del Vicario et al., 2016; Kahan, 2017). Motivated reasoning implies that there is deliberation and cognitive reflection of the content, but Pennycook and Rand (2019) found that susceptibility to false information is attributable more to a lack of analytical thinking than to motivated reasoning. That is, people believe in misinformation "because they fail to think; not because they think in a motivated or identity-protective way" (Pennycook & Rand, 2019, p. 10).

This finding is consistent with years of investigation in users' online decision-making, which is known to be shallow and seldom effortful or analytical. But that does not mean their behavior is random; rather, it could be systematically biased by contextual factors. Humans tend to be cognitive misers (Fiske & Taylor, 1991), often relying on easily interpretable cues that trigger heuristics or "rules of thumb" without effortfully perusing the information. Such cues can be related to the content of the story. False news, for instance, is known to capitalize on users' emotions to generate engagement (Bakir & McStay, 2018). Other cues are external to the content of the story, such as cues displaying the authority or expertise of the source and cues conveying the popularity of a story (Metzger et al., 2010; Sundar, 2008). Thus, when assessing why false content becomes viral online, a missing piece of the puzzle could be the role of interface cues encountered on social media, specifically the source of the article and the number of likes a post receives (referred to as bandwagon cues). In the studies reported here, we assess the influence of these cues on engagement with false news online.

Literature Review

Theories suggest that the overloaded online information environment leads news consumers to rely on heuristics, or cognitive rules of thumb, in the context of news reception (Metzger et al., 2010; Sundar, 2008). For example, false news, defined as intentionally and verifiably false information (Allcott & Gentzkow, 2017), uses emotional appeals to attract clicks (Bakir & McStay, 2018). Users' reliance on emotions for decision-making is known as the "affect heuristic" (Slovic et al., 2007). Users' decisions are also based on non-content cues, such as the source of the content and the apparent support it has received from others (Metzger et al., 2010). According to the Modality-Agency-Interactivity-Navigability (MAIN) model (Sundar, 2008), such cues on the media interface affect perceptions by invoking cognitive heuristics, or mental shortcuts, about the nature of the underlying content. Making decisions by relying solely on such simple decision rules is called "heuristic processing" because it requires fewer cognitive resources compared to "systematic processing," which involves more cognitively effortful or analytical scrutiny of content (Eagly & Chaiken, 1993). This study investigates the bandwagon cue (the number of likes a story has received) and the source cue, specifically whether the article is written by a human or a bot, and their differential effects on false versus real content.

Bandwagon Cues and News Consumption on Social Media

A unique aspect of news consumption via social media is the salience of various popularity metrics (e.g., the number of likes, comments, and shares), which usually represent other users' reactions to the news. This is one of the strongest signaling devices utilized by platforms because social networks remind users about news that have received attention from others. When users are exposed to these metrics, the mental shortcut that "if others think this is good, then it must be good for me, too" can be triggered (Sundar, 2008). This rule of thumb is known as the "bandwagon heuristic." Extant research shows strong support for this proposition in the context of news consumption. News associated with a large number of "likes" or "diggs" is often perceived as more credible (Xu, 2013). As a contextual factor, bandwagon cues may also be one of the main driving forces for users' engagement with news on social media. News with either an implicit bandwagon cue (e.g., a larger number of views) or an explicit bandwagon cue (e.g., high ratings) attract more attention from users (Knobloch-Westerwick et al., 2005). Similarly, if the news is labeled as "the most viewed" in the recommendation system, its chances of being selected and read by users increase significantly (Yang, 2016).

All this suggests that it is quite likely for a piece of news with a higher number of likes to trigger the bandwagon heuristic, therefore be perceived as more credible, and lead to more active engagement with the news. In this way, the trustworthiness of the false news could be boosted by the strong social endorsement conveyed by bandwagon cues. Luo et al. (2020) found initial evidence for this—headlines associated with a high number of likes were perceived as more credible. However, it is also possible that bandwagon cues play a differential effect on real compared to false news. On the one hand, it is known that false news leverages users' emotions to increase attention (Bakir & McStay, 2018). This effect may be exacerbated by a large number of users who endorse the content. On the other hand, the large number of likes might serve as an indicator of the fakeness of an article, especially when coupled with more affective and provocative content, as is often the case with false news. Recent campaigns informing users about false news disseminated by bots (e.g., Center for Information Technology & Society, 2020) could have resulted in a more aware audience. We explore these possibilities further.

Automated Journalism and News Consumption

Another important cue is the source cue. Extensive research suggests that characteristics of the source, such as gender and ethnicity, influence how people evaluate and engage with content (e.g., Winter & Krämer, 2014). Less is known about users' evaluation of content when the writer is not human. Recent advances in machine learning and natural language generating techniques have given rise to "automated journalism." While automation is increasingly being adopted by mainstream media such as *Bloomberg* and *The Washington Post*, it has also been used for generating and spreading false news (Shao et al., 2018). This raises the question: If users learn that the news is written by a bot, how will it affect their judgments and reactions to false news vs. real ones? From one perspective, knowing that the author of the news is a bot may evoke the "machine heuristic"—the belief that machines are objective and free from ideological bias, leading to the attribution of a higher level of credibility to the news (Sundar, 2008). Previous studies lend support to this idea by showing that bot journalists decreased perceived extremity and bias of the news, due to invocation of machine heuristic (Liu & Wei, 2018). Moreover, when several articles were presented to readers concurrently, human-written news was expected to be higher in readability whereas bot-written news was expected to be more credible (Haim & Graefe, 2017). The awareness of the automated nature of the news may increase the trustworthiness of a piece of false news, triggering more interactions with it.

Alternatively, the heightened expectation originating from the machine heuristic may serve to negatively affect audiences' evaluation of bot-written news. Waddell's (2018) study revealed that the news generated by a bot was perceived as less trustworthy than news written by a human because it failed to meet readers' expectations of quality. This quality expectation might be further violated when the content is false because false news is often written using sensationalistic language and other characteristics that compromise quality (Molina et al., 2021a). In addition, human readers do not always hold a positive impression of bots because of the mental association between bots and artificiality, leading to less positive reactions to bot-written news. This negative mental association about machines might be exacerbated by literacy campaigns informing users about the use of bots for dissemination of misinformation.

Online Engagement

In order to assess how online users respond to the aforementioned cues on the interface, it is important to conceptualize the different types of engagement that can occur in social media. For example, users can opt to read, like, comment, or share content. Each of these actions has a different meaning for users and can be placed on a continuum from low to high effort. First, these actions differ in their required proactivity. Users might opt to simply read a piece of content out of curiosity (Tenenboim & Cohen, 2015) or they can perform one-click actions, such as "like," which are typically more reactive (than proactive) compared to composed comments (Burke & Kraut, 2016; Zell & Moeller, 2018). Users could also be much more actively engaged by sharing the content to their networks, thereby becoming *de facto* sources of information. Additionally, different engagement affordances are associated with different levels of publicness (Aldous et al., 2019). While reading content is a private action, liking or commenting can be seen by networked users, and sharing represents the most public engagement. More public reactions have greater implications for users' own identity and credibility compared to simply consuming news, which usually requires higher commitment and deliberation (Oeldorf-Hirsch & Sundar, 2015). In fact, Oh et al. (2018) identify a continuum in user engagement with interactive media, from clicking at the initial stage, followed by assessment and absorption, and culminating in sharing or "digital outreach." Therefore, we place these different types of engagement—reading, liking, commenting, and sharing—on a spectrum from low to high effort, representing the increasing level of commitment and deliberation involved in engaging in these actions. When users suspect the news is false, they may be less likely to perform public sharing, considering the consequences to their credibility. Alternatively, they may be more likely to share because of its potential to trigger discussion and debate, especially if the source of the news is a contentious one. If a piece of false news has already received bandwagon support, it might evoke sufficient curiosity to engage with it as a reader but not necessarily as a commenter or sharer. In this study, we investigate these possibilities and explore whether interface cues have a differential effect on user engagement with real versus false content, and if the effects vary across engagement actions. Conflicting evidence in the literature does not allow us to propose directional hypotheses. Therefore, we pose the following research question for Study 1:

> **RQ1:** What are the effects of type of content (real vs. false), bandwagon cues (high vs. low), and source cues (human vs. bot) on users' engagement (reading, liking, commenting, or sharing) with content on social networking sites?

Sequence of Actions

The aforementioned actions underlying user engagement may imply a hierarchy when viewed in the context of traditional media, that is, one would share a news story only after reading it oneself. But, this is not necessarily true with social media, where the reverse appears to be more common. An important driver of the spread of false news on social media is that users engage with articles without even reading them. An analysis with a large-scale Twitter dataset found that nearly 60% of all shared URLs online do not receive a single click (Gabielkov et al., 2016). Guided by these findings, scholars have explored the different drivers of sharing behavior. Xu et al. (2020) found that the most shared articles typically reveal the author's name, although not necessarily other information about the authors (e.g., email address), and articles that emphasize authority as a moral frame are shared and liked the most. Likewise, Janét et al. (2020) found that the framing of a headline influences users' evaluation of the headlines, but not user engagement, and Molina et al. (2021b) found that non-clickbait headlines generate about the same (if not more) engagement than clickbait headlines. But it is unclear if cues in the interaction context, such as the bandwagon cue or the bot source, will promote more perusal of the news or more immediate actions such as sharing. The presence of a large bandwagon may elicit curiosity and persuade users to read the content further or it may boost the credibility of the content (Sundar, 2008) making users succumb to endorsing it without effortfully engaging with story details. Similarly, cueing the bot source of the story may either promote greater involvement or superficial endorsement of its sentiment without fully reading the story. Such possibilities may be moderated by the nature of the story, that is, whether it is true or false. To address these issues, we pose the following research question:

RQ2: What are the effects of type of content (real vs. false), bandwagon cues (high vs. low), and source cues (human vs. bot) on users' tendency to read the content before performing (a) effortful and committed actions (comment/ share) or (b) less effortful and committed actions (likes)?

Study 1

Method

To explore our research questions, we conducted a 2 (Type of News: real vs. false) × 2 (Bandwagon Cues: high vs. low) × 2 (Source: human vs. bot) within-participants online experiment utilizing an interface created especially for this experiment. The interface mimicked Facebook in its structure (See Figure B1 of Supplemental Appendix) and contained eight articles, randomized in its order of presentation. For each article participants were randomly assigned to one level of each independent variable, meaning that it was either the real or false version of the story, written by a human (staff writer) or a bot "Automated Insights," and had either high or low bandwagon cue. Since each participant saw eight articles in their feed, we had a total of 1,368 (171 participants × 8 articles) instances wherein participants had the opportunity to act. As such, the number of instances assigned to each cell of the 2 × 2 × 2 design ranged from 156 to 190.[1]

Participants. Participants were recruited from Amazon Mechanical Turk and paid $1 for completing the study. The sample consisted of 171 participants (72 female, 74 male, 25 did not report) of ages ranging from 23 to 69 years old ($M = 37.13$, $SD = 10.39$), and 70% of them identified themselves as Caucasian. Among the rest, 5% were Hispanic, 5% African American, 5% Asian, 2% other, and 13% did not report race. All participants were from the United States.

Procedure. After acknowledging consent, participants completed a questionnaire asking about their media use pattern and other individual-difference variables.[2] Following this, participants were directed to the interface and instructed to browse the site for 5 minutes as they normally would, clicking, reading, liking, and/or sharing as many or as few articles as they wanted. The 5-minute browsing time was selected because previous studies in the selective-exposure literature utilize this threshold (e.g., Knobloch-Westerwick & Meng, 2011). The interface unobtrusively recorded users' activities, including articles that participants read, article reading time (measured in terms of the time that the user is active on the article page), liking, commenting, sharing, as well as the messages that users typed when commenting and sharing. The interface additionally tracked the sequence of those actions. Importantly, the interface did not display user responses to other users. Once the 5 minutes elapsed, participants continued to a questionnaire eliciting their demographics and political affiliation.

Stimulus and experimental conditions. The user interface was written in Python programming language, with Flask and MongoDB database as backend. Upon starting a session, participants could see eight news posts, each ostensibly shared by a media organization (to control for source effects) (See Figure 1). The specific news articles used for this study and media organizations associated with each story were selected based on a pretest, as described in the next section. Multiple stories were used to achieve stimulus sampling and control for content-specific effects. Each of the eight articles was randomly selected to be the false or real version of the article, had either a high or low initial number of likes (high/low bandwagon), and was written by "automated insights" or "staff writer" (See Figure 1). Participants could click on the headline of the post and read the complete story (See Figure 2) and could return to the main page upon clicking the back button.

Story selection. We define false news as intentionally and verifiably false information (Allcott & Gentzkow, 2017; Molina et al., 2021a), and real news as information that can be verified as truthful based on collective consensus (Southwell et al., 2017). Following these definitions, we searched for false stories on Snopes.com. Once a false article was identified, we searched for its real counterpart, which was often provided by Snopes itself. If the real version was not on Snopes, we located it elsewhere on the internet. We followed this strategy because false news often originates from a real event, which is taken out of context or distorted in some way. A total of 12 false stories (with their respective real counterparts) were identified. Then, we conducted a pretest with a different group of participants to arrive at the final eight stories. The selection criteria were as follows: First, we selected eight different stories varying in their topic. Second, we selected stories where participants reported at least a moderate level of interest. Finally, we chose false stories that roughly corresponded to the same topic domain as the real ones but were perceived as significantly more false than their real counterparts (without being seen as patently false), confirming that our manipulation was successful (See Supplemental Appendix A1).

Media organization selection. Based on another pretest, four media organizations (Axios, NY Observer, Daily Progress, and Daily Cardinal) were chosen as they did not differ in credibility and familiarity, and their perceived objectivity was not overly low or high.

Independent variables. The following sections describe the manipulations of the independent variable of this study.

Type (real vs. false) manipulation. To manipulate the type of article (real vs. false), we searched for false stories on Snopes.com. Once a false article was identified, we searched for its real counterpart. To further strengthen our manipulation of false news, we altered certain elements of the false articles. We opted for this strategy to ensure that the false version looks as naturally false as possible in the interest of achieving ecological validity and to ensure that all the false articles possessed the same characteristics. For example, the image associated with false stories is typically of low quality or is taken out of context. Not all articles originally had an image, thus we selected one based on the standards identified by Molina et al. (2021a). Likewise, real news

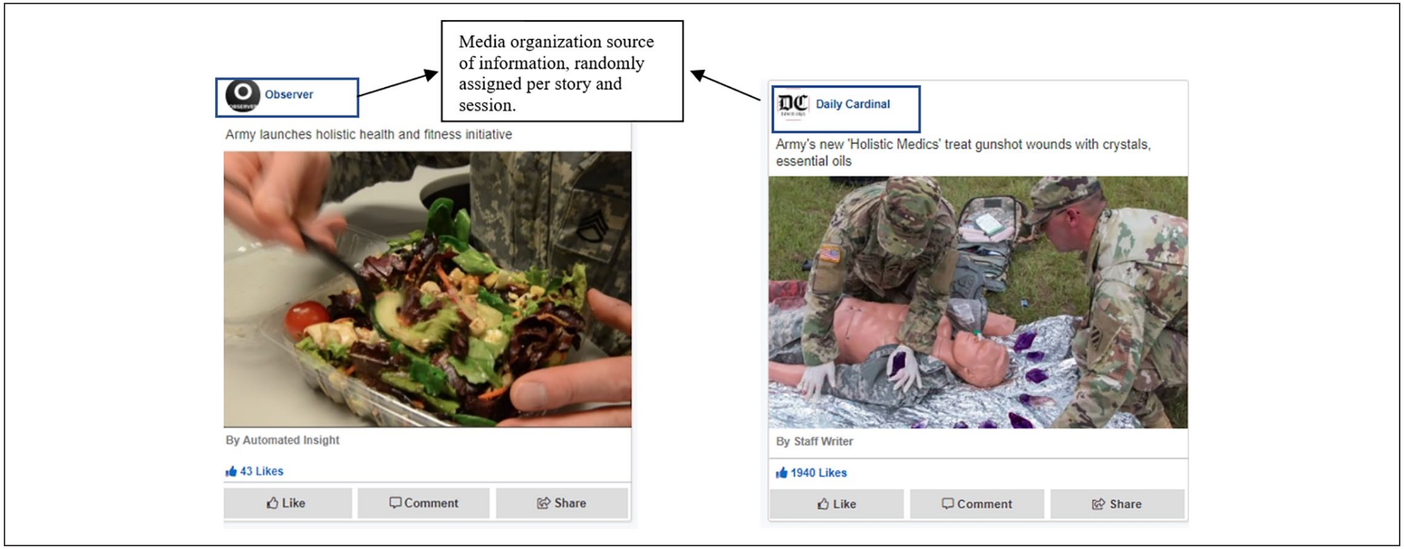

Figure 1. Sample news story posts.
Note. Each story was associated with one randomly assigned media organization for Study 1 and was kept constant in Study 2. The image on the left illustrates the real version of this particular story, written by a bot (Automated Insight), and featured a low bandwagon cue. Conversely, the image on the right illustrates the false version of the story, written by a human (staff writer), and featured a high bandwagon cue.

Army's new 'Holistic Medics' treat gunshot wounds with crystals, essential oils

By Automated Insights

In an effort to match the broad range of medical treatments available to the civilian population, the U.S. Army has introduced a new breed of battlefield medic, the service announced Friday.

The new MOS, Complementary Medicine Specialist (69W), or "holistic medic," will be trained in a variety of alternative medical treatments, ranging from aromatherapy to interventional prayer, and will be authorized to prescribe medications like megavitamins and homeopathic dilutions.

"There's great demand for treating warfighter injuries with more natural, holistic remedies," said Col. Thomas Sorrenson, commander of the new program. "So where a traditional combat medic treats an amputated limb with a tourniquet and hemostatic gauze, our protocol opts for a healing-hands reiki session followed by some cleansing herbs."

"It's a more integrative approach," he said. "After all, what's really more important: stopping massive blood loss and preventing shock, or curing a devastating disruption of a soldier's chi?"

In lieu of conventional trauma supplies like bandages, tourniquets, and chest seals, holistic medics' aid bags will be stocked with colon cleanse kits, prayer beads, and ginkgo biloba auto-injectors.

Sgt. 1st Class Nick Bye, an instructor in the holistic medic course, told Duffel Blog that the training to become a 69W is intense.

"The course includes both lecture and hands-on instruction," said Bye. "Holistic medics have to not only know their pharmacology, such as which essential oils to use for local anesthesia, but also practical skills, like stopping a sucking chest wound in under 30 seconds through the power of positive thinking."

"Sure, some have dismissed our methods as 'pseudoscience,' but we've demonstrated a 100% success rate," he added with a confident smile. "We haven't lost a training mannequin yet."

Army launches holistic health and fitness initiative

By Automated Insight

WASHINGTON -- The Army has more obese Soldiers than ever before, and to address the problem, senior Army leaders discussed a holistic approach that includes vegan options at dining facilities and box breakfasts at fitness centers.

Currently, one in 20 Soldiers fails the Army Physical Fitness Test annually and 13 percent of Soldiers are clinically obese, according to statistics from the Army Surgeon General's Office that were presented Monday at the 2017 Medical Symposium. The conference was hosted by the Association of the U.S. Army in San Antonio, Texas.

"I do think that we've been recognizing that what we are doing has not been working," said Michael McGurk, director of the Research and Analysis Directorate at the Army Center for Initial Military Training. "Holistic health and fitness is a radical change -- radical departure. It is going to cost the Army money, time and people. And we've got to be willing to give that up to make the changes that we need."

The Army's new campaign plan, called "Holistic Health and Fitness" focuses on improving the Soldier selection process, physical performance, performance education, and transforming and improving Soldier fitness/training centers.

Tim Higdon, Healthy Army Communities program manager, said one suggestion is to have fitness centers provide box breakfasts so that Soldiers in a time crunch will not need to choose between having breakfast and working out. Higdon said HAC is also planning to attempt innovations such as requiring 15 to 50 percent of vending machines to have healthier options, while others will be 100-percent comprised of healthy items. Vegetarian and vegan options may also be added to dining hall menus.

Figure 2. Complete story after clicking on post headline.
Note. Structure of the interface after participants clicked on the headline to read more. The false story includes a less professional image and no dateline (left). The real story (right) includes a more professional image and a dateline, among other indicators of real and false news. The source manipulation appears in the form of a byline at the beginning of the story. We used the same structure and format to display stories to participants in Study 1 and Study 2.

typically has a dateline whereas false news often does not. Again, some original false articles had a dateline and not others. We also modified the articles to keep all the articles (real and false) relatively consistent in terms of structure, length, and complexity.

Bandwagon manipulation. Each story was randomly assigned to show either a low or high bandwagon cue, operationalized as the number of likes received by the story. To randomly generate the number of likes, we used the following formula: #likes = a^b + c, in which a is an integer from 40 to 50, b is either 1 or 2, and c is a natural number from 0 to 9. This results in low versus high bandwagon conditions, with the number of likes ranging from 40 to 57 and 1600 to 2409, respectively.

Source manipulation. Each article was randomly assigned to a bot "Automated Insights" or a human "Staff Writer." Upon accessing the interface, participants could see the source in the left bottom corner of each post (See Figure 1). If the participant clicked the story to read it, s/he viewed the full story with the source in the byline (See Figure 2).

Dependent variables. The various actions performed on each article by participants, including liking, commenting, sharing, and reading, served as dependent variables of this study. Two subtypes of variables were created for analyses: *performance of action* and *first action.* The *performance of action* variable refers to whether a given action (liking, commenting, sharing, reading) was performed or not, and was coded as a binary no/yes response, coded as 0 and 1 respectively. The *first action* variable refers to the action that participants performed first for each article, coded as a nominal variable: Read, Like, Comment, Share. Like was classified as a low commitment or effort whereas Comment and Share were combined into one variable considered high commitment or effort.

Results

Predictors of action. Before conducting the analyses to answer our specific research questions, we assessed the general pattern of users' actions by calculating the rate of occurrence of each action. To achieve this, we divided each action distribution by the number of participants (171) and stories (8). Results revealed that users tended to read articles more often than performing any other action, with a 0.44 rate of occurrence, meaning that users read the article 44% of the time (See Table A2 of Supplemental Appendix). Liking behavior was the second most performed action occurring 23% of the time, with liking of real news occurring more often. Conversely, commenting behavior occurred 10% of the time, with false articles being commented on more often. Sharing behavior was the action that occurred least often, with it occurring only 3% of the time.

To answer RQ1, a series of generalized linear mixed models with binary logistic distribution was conducted to assess the relationship between the three independent variables (content type, bandwagon, source) and the performance of each action (like, read more, share, and comment). We opted for this data analysis strategy due to the unbalanced nature of our design and the dichotomous dependent variables. For all models, story was treated as a repeated-measures variable.

When analyzing users' "liking" of an article as the dependent variable, the type of article (real or false) was a significant predictor (F (1, 1360) = 29.92, $p < .001$). Analysis of the main effect indicates that the odds of false articles receiving a "like" were 52% less than real news, OR: 0.48 (95% CI: 0.37, 0.63). Bandwagon cues and source were not significant predictors, and there were no significant interactions.

For users' "commenting" behavior as the dependent variable, the type of article (false vs. real) was again a significant predictor of commenting F (1, 1360) = 21.48, $p < .001$. Analysis of the main effect reveals that the odds of false articles receiving a comment were more than twice that of real articles, OR: 2.74 (95% CI: 1.84, 4.08). There was also an interaction effect between bandwagon cues and source (F (1, 1360) = 13.08, $p < .001$). Analysis of the two-way interaction effect reveals that when the writer is a bot (vs. staff writer), users were less likely to comment on articles with high bandwagon (vs. low bandwagon) (OR: 0.22, 95% CI: 0.10, 0.49), but when the writer is a staff writer, users were more likely to comment on articles with high bandwagon (OR: 4.48, 95% CI: 2.05, 9.80). In other words, when the bot is the writer, users are more likely to comment on articles with *few* likes, and conversely, when a human (staff) is the writer, users are more likely to comment on articles with *many* likes. Analyses with users' sharing and reading behaviors did not yield any significant findings.

We conducted a follow-up analysis to assess if the interaction effect between bandwagon cue and source on users' commenting behavior was the same for users who commented before reading the article or after reading the article. We ran two different models, one for users who read the article first and one for users who did not read the article first. The model remained significant only for users who did not read the article first (See Supplemental Appendix C) showing that cues in the interaction serve to elicit immediate action or heuristic processing of information.

First action. To answer RQ2a, investigating the conditions under which the user is more likely to read the article before performing a more effortful and committed engagement (commenting/sharing), we conducted a series of generalized linear mixed models with a binary logistic distribution using the *first action* variable as the dependent variable. For this analysis, we combined participants who either commented or shared as a first action into one "effortful or committed" action. This helped us test if our independent variables influenced users' decision to comment or share even before reading the article—commenting and sharing representing a more committed type of engagement (Aldous et al., 2019; Burke & Kraut, 2016; Zell & Moeller, 2018). Only participants who either read first or commented/shared first were included in the analyses (N = 606). Two separate logistic regressions were employed, one examining the main effects of type of content (real vs. false) and bandwagon cue (low vs. high), and their interaction; and the other examining the main effects of type of content and source, and their interaction.[3] The story was treated as a repeated-measures variable.

When entering type of content and bandwagon as independent variables, we found significant main effects for type of content (F (1, 602) = 6.59, p = .01). Analysis of the main effect revealed that the odds of a user commenting on the article first (vs. reading) are almost double for false articles (compared to real) (OR: 1.98, 95% CI: 1.16, 3.40). These effects should be interpreted in light of a potential interaction between type of content and bandwagon, suggesting that for false news, users were more likely to comment first (than read) when articles had high bandwagon (compared to low bandwagon) whereas for real articles, they were more likely to comment first (than read) when the articles had low bandwagon (compared to high bandwagon) (OR = 2.46, 95% CI: 0.81, 7.46). This data pattern implies that when paired with high bandwagon cues, false articles might elicit more immediate action. However, given that the interaction effect fell short of the conventional cut-off for statistical significance (F (1, 602) = 2.54, p = .11), we decided to further investigate it in Study 2. The second regression entering type and source as independent variables did not yield any significant results.

Similar analyses were run for RQ2b assessing under what conditions the user was more likely to read the article before liking it—liking representing a less committed action, and less public than commenting or sharing. In this case, the dependent variable only included participants who liked the article first and those who read the article first. We found a main effect for type of content (F (1, 736) = 8.76, p = .003), such that the odds of reading the article first, relative to liking it first, was higher when the article was false (vs. real), (OR: 1.65, 95% CI: 1.19, 2.31. Results revealed no effect of bandwagon cue and no interactions. The second regression entering type and source as independent variables did not yield significant results.

Discussion

Data reveal that engagement with online news is contingent upon specific content and non-content characteristics, and it differs based on the action being performed (i.e., reading, liking, commenting, sharing). First, "liking" behavior was largely driven by content characteristics of the story. Specifically, false articles were liked less than real articles. Although we cannot be sure that participants in the experiment were able to detect the fakeness of the article, this finding suggests that, at least stylistically, they were able to perceive a difference between these two types of content. We also found that users are more likely to comment on false articles (versus real articles). There are two possible reasons for this finding. On the one hand, it could be that the emotionality of false news triggered fast responses from users without careful analysis of the content. On the other hand, it could be that the style of false news raised concerns about the veracity of the content, leading users to engage in more careful scrutiny. To test both possibilities, we entered the actual comments left by participants into LIWC software to analyze the degree to which they were analytical. We also coded the comments to assess if users could identify them as fake (See Supplemental Appendix D for details). We found that comments on false articles tended to be more analytical, revealing that when engaging in more effortful and committed action, users do tend to interrogate falsehoods. Metzger et al. (2021) found similar results. In a dataset of over 2.5 million comments, 15% of user comments about false content expressed disbelief.

It is likely that users' motivations for engaging with false news are not always based on the believability of the content, but entertainment, education, or debate over conflicting information (Buttliere & Buder, 2017; Metzger et al., 2021).

Furthermore, commenting behavior is predicted by interface cues. For example, we found an interaction effect between bandwagon cue and source, revealing that when content is written by staff and has a high bandwagon cue, users are more likely to comment. However, bandwagon cues have the opposite effect when articles are written by a bot. Possibly, a bot-written article without endorsement from other users induces curiosity among users; thus, they decide to engage with the content and express their opinions. Conversely, it might be that an article written by a bot that has received many likes raises a red flag about its potential fakeness, thereby inhibiting commenting action. The negative associations of a bot-written article (Waddell, 2018) could have been exacerbated by a high number of likes. We also found initial evidence of a potential interaction between type of content and bandwagon cue. When content is false and it has received many likes, users are more likely to comment before reading the article first, compared to when the article is real or has received only a few likes. Nonetheless, this finding should be investigated further as it fell short of significance.[4]

Given the behavioral nature of our measures in Study 1, we are unable to uncover the mechanisms underlying our findings, thus necessitating a follow-up study.

Study 2

Results of Study 1 reveal the differential effects of type of content and interface cues on the various forms of social media engagement. In Study 2, we explore potential reasons for these effects.

Effects of Type of Content on User Engagement

Results of Study 1 reveal that users' engagement with content online differs based on whether it is true or false. It is possible that false stories, being inherently more emotional and arousing than real news, alert users to their potential falsehood and diminish their credibility. To test this possibility, we propose the following hypothesis:

> **H1:** False news will be (a) perceived as more false and (b) less credible than real news.

It is likely that this perceived falsehood could have driven the results of Study 1—users liked real news more but commented on false articles more. Preliminary analyses of the comments left by users also reveal greater analytical thinking when engaging with false news. This means that the nature of false news could have prompted users to process information more systematically, influencing engagement.

However, engagement with content is not only driven by cognitive factors but also affective ones (Kormelink & Meijer, 2018), that is, the emotions that each type of news invokes in readers. This is what Slovic et al. (2007) call the "affect heuristic," wherein individuals consult "positive and negative tags consciously or unconsciously associated with the representations" (p. 1335) during decision-making. For example, social media posts driven by positive emotions are known to receive more engagement than those driven by negative emotions (Gerbaudo et al., 2019). Likewise, Kormelink and Meijer (2018) found that users engage with content that makes them feel good or is moderately disheartening. It appears that real stories are liked more because they elicit positive emotions whereas false stories tend to evoke negative emotions, leading users to deliberate on them.

Research also suggests that cognitive and affective factors might not be sufficient to elicit engagement. To elicit engagement, the headline should provide just enough information to peak one's interest, but also leave something wanting—an associative gap whereby users are not provided all the expected information to satiate their curiosity about the topic, thus persuading them into action (Kormelink & Meijer, 2018; Loewenstein, 1994). This associative gap could be manifested in the form of likes for real news (compared to false) given the higher interest that users have toward positive content (Gerbaudo et al., 2019; Kormelink & Meijer, 2018). After all, liking behavior serves as a signal to indicate importance or interest (Zell & Moeller, 2018). On the other hand, for false news, the associative gap could be manifested in the form of comments. We know that stories that are more controversial tend to receive more comments from users (Tenenboim & Cohen, 2015). To test all the aforementioned mechanisms, we propose:

H2: Perceived fakeness (H2a), cognitive elaboration (H2b), invoked emotions (H2c), and aroused curiosity (H2d) will each mediate the relationship between the type of content (real vs. false) and user engagement, such that users will like real news more, but comment more on false news.

Role of Interface Cues on Commenting Behavior

Study 1 revealed that aside from the news story being false or real, cues on the interface can also predict commenting behavior. Specifically, the data pattern (although not statistically significant) suggests that when false content has high bandwagon cues, users comment before reading the article. It is possible that they feel compelled to comment in an effort to stem the tide of opinion on the topic. Perhaps they are driven by current media literacy campaigns informing users about the inflation of likes on social media by using bots. Thus, users are probably being more cautious about endorsement cues, particularly when the content of the article appears to be negative or false. In these instances, it might be that they comment to warn others about false content. Given the lack of prior empirical evidence, we pose the following research question instead of a hypothesis:

RQ3: What is the role of (a) bandwagon perception, (b) fakeness perception, and (c) emotions induced by the post in explaining the effect of bandwagon cues and type of content upon commenting and reading behaviors?

Study 1 also revealed that when content is written by a human, users are influenced by the high number of likes. However, bandwagon cues have the opposite effect when articles are written by a bot. Again, these findings may be explained by increased media literacy, in that users may be perceiving bots as less objective, especially when receiving a high number of likes. That is, users might question the accuracy of bot-written stories, and have doubts about the real number of people who have endorsed the content, thus perceiving the article as more fake and therefore dismissing them without being curious about its content. However, they might avoid commenting on bot-written stories with a high number of likes for the opposite reason. It may be that they do not comment on it because of the operation of the machine heuristic—"if the story is written by a bot then it is accurate and objective." Thus, the story might be perceived as a fairly solid story that does not require a comment. The machine heuristic could also make the story be perceived as unequivocal and boring, inhibiting curiosity. We propose the following research question to explore these possibilities:

> **RQ4:** What is the role of (a) bandwagon perception, (b) fakeness perception, (c) machine heuristic, and (d) perceived curiosity in explaining the effect of bandwagon and source cues upon commenting behavior?

Study 2 Method

To assess the different perceptual, affective, and cognitive mechanisms that could drive users' engagement with content in social media, we conducted a 2 (Type of News: real vs. false) × 2 (Bandwagon Cues: high vs. low) × 2 (Source: human vs. bot) between-participants online experiment. Unlike Study 1, each participant was exposed to only one post and the associated news story, as explained below, given our interest in psychological mediators.

Participants. Participants ($N=284$) for this study were recruited using Amazon Mechanical Turk (46.2% female, 53.8% male). Their ages ranged from 18 to 70 years ($M=39.16$, $Md=36$, $SD=12.16$). Most participants reported being White/ Caucasian (68.9%), 12.2% African American, 6.6% Asian, 4.9% Hispanic, and 0.7% Native American. The remaining 6.6% reported having another ethnicity or did not report.

Procedure. After acknowledging informed consent, participants were randomly assigned to one of the eight conditions varying in the type of content they received (real vs. false), level of bandwagon cue (high vs. low), and source (human vs. bot). Participants were also randomly assigned to one of three story topics (derived from Study 1) for stimulus sampling purposes. The topics chosen included a military story, a story about a TV show, and a story about human rights.

As a first step, participants received the *post* of their assigned condition as presented in social media, without the associated story (see Figure 1). The source of the post for Study 2 was kept constant across conditions. After 10 seconds, participants could hit "continue" and were directed to answer a questionnaire assessing their intentions to engage with the post (like, read, share, comment), as well as the mediators of interest. Then, participants were shown the *story* associated with the post (see Figure 2). After reading the article, they were redirected to answer a questionnaire assessing their intention to engage with the story, as well as the mediators of interest. The questions were the same as the questions about the post, but asked participants to assess the article instead. At the end of the questionnaire, participants were asked demographic questions as well as questions pertaining to control and moderating variables. Note that for Study 2, we assessed users' perceptions and engagement intentions, both after viewing the post (before viewing the story associated with the post) and after viewing the story. This is because both events are theoretically and practically different. Users' decision to act (like, comment, or share) without reading the associated article is based on the headline of the story, the image, source, and bandwagon cues only. Their decision to act *after* reading the associated article is based on their perusal of the actual content of the story. It is likely that liking, commenting, or sharing based on reading only the headline of the post is guided more strongly by heuristic cues, compared to those same actions taken after reading the content of the story. Study 1 data indicate that this is the case. We test this possibility formally in Study 2.

Independent variables. The independent variables for Study 2 were the same as in Study 1. The only difference was in the bandwagon conditions. In Study 2, all participants in the high bandwagon condition saw a post with 1,600 likes and those in the low bandwagon condition saw one with 47 likes.

Dependent variables. The dependent variables were all measured on a 1 to 7 scale and were worded as described below.

Engagement. To assess engagement with the post and the story, participants were asked their likelihood (in a 1–7 scale) of reading the article associated with the post ($M = 3.63$, $SD = 2.21$), liking the post or story ($M_{post} = 2.50$, $SD_{post} = 1.97$; $M_{story} = 2.82$, $SD_{story} = 2.00$), commenting on the post or story ($M_{post} = 2.56$, $SD_{post} = 1.98$; $M_{story} = 2.61$, $SD_{story} = 1.99$) or sharing the post or story ($M_{post} = 2.49$, $SD_{post} = 1.98$; $M_{story} = 2.77$, $SD_{story} = 2.13$). We did not combine actions into committed versus not committed actions due to its continuous nature.

Emotional response. Participants' *emotional responses* to the post and the story were measured through three dimensions proposed by Schimmack and Grob (2000)—*pleasure (enjoyment), alert*, and *calm.* Participants were asked how they felt after engaging with the post (or after reading the story) through a 9-item semantic differential scale. Items measuring enjoyment ($M_{post} = 3.79$, $SD_{post} = 1.62$, $\alpha_{post} = .89$; $M_{story} = 3.69$, $SD_{story} = 1.69$, $\alpha_{story} = .90$) were unpleasant/pleasant, good/bad, and positive/negative. Items measuring alertness ($M_{post} = 4.62$, $SD_{post} = 1.14$, $\alpha_{post} = .69$; $M_{story} = 4.63$, $SD_{story} = 1.19$, $\alpha_{story} = .73$) included awake/sleepy, tired/energetic, and drowsy/alert. Items measuring calmness ($M_{post} = 4.30$, $SD_{post} = 1.25$, $\alpha_{post} = .71$; $M_{story} = 4.20$, $SD_{story} = 1.32$, $\alpha_{story} = .82$) were tense/relaxed, clutched up/calm, and at rest/jittery.

Content perception. Participants' perception of content was assessed via three measures: fakeness perception, curiosity arousing, and credibility. Participants were presented with a battery of adjectives and were asked to rate how well each of the terms describe the post/story. *Credibility* (M_{post}=3.67, SD_{post}=1.56, α_{post}=.74; M_{story}=4.11, SD_{story}=1.56, α_{story}=.75) was assessed using Sundar's (1999) credibility scale and included the items objective, fair, and biased. Items assessing *curiosity* (M_{post}=3.55, SD_{post}=1.63, α_{post}=.91; M_{story}=4.06, SD_{story}=1.62, α_{story}=.92) were created by the authors and included want to know more, intriguing, aroused curiosity. Items assessing *fakeness perception* (M_{post}=4.42, SD_{post}=1.58, α_{post}=.86; M_{story}=3.90, SD_{story}=1.56, α_{story}=.84) were also created by the authors and included fake, deceptive, dishonest, tricky, disturbing, and sensationalistic.

Machine heuristic. The *invocation of the machine heuristic* was measured by a 6-item scale based on Sundar (2008) (e.g., machine-like precision, does not have human touch, formulaic) asking participants to rate how well each term described the post (M=3.93, SD=1.33, α=.83) and the story (M=3.52, SD=1.42, α=.87).

Bandwagon perception. Bandwagon perception was assessed after participants viewed the post only since this was where the manipulation was located. Participants were asked to rate, on a 1 to 7 scale, the extent to which they agree with four questions adapted from Sundar (2008), including how likely are other people to think that this is a credible post and how likely are other people to recommend this post (M=3.52, SD=1.42, α=.87).

Elaboration. Elaboration was assessed through Reynolds' (1997) 12-item measure asking participants their agreement with a 7-point scale. Items included: while reading the news story I was "attempting to analyze the issues in the message" and "not very attentive to the ideas (rc)" (M=4.89, SD=1.14, α=.88).

Control variables. Political ideology and issue involvement were entered as control variables. *Political orientation* was measured via 4 items from Janoff-Bulman et al. (2008) (M=3.68, SD=1.47, α=.83). We also assessed users' *involvement* with the topics covered in the stories shown to participants—human rights issues (M=5.67, SD=1.35, α=.89), military issues (M=4.45, SD=1.65, α=.86), and TV show news (M=4.17, SD=1.71, α=.90)—through three subscales asking participants to what extent they consider each topic personally relevant, important, and interesting.

Study 2 Results

In Study 2, we tested user perception and engagement intention at two time points—after viewing the post only and after reading the story associated with the post. This allowed us to test our hypotheses at both time points and assess if heuristic cues operate differently before and after users have read the complete story. Thus, separate analyses were conducted for users' responses pertaining to the post and those pertaining to the story.

To test H1, two separate 2 (Type: real vs. false) × 2 (Bandwagon Cue: high vs. low) × 2 (Source: human vs. bot) multivariate analyses of covariance (MANCOVAs) were conducted for fakeness and credibility as dependent variables, one for participants' perceptions of the post and one for their perceptions of the story. Issue involvement, political orientation, and story were entered as covariates in all MANCOVAs.

Results for users' perceptions of the post[5] revealed a significant multivariate main effect of type of content, Pillai's Trace = .32, F (2, 268) = 61.54, $p < .001$, partial $\eta^2 = .32$. Specifically, the false version of the post ($M = 5.21$, $SE = .11$) was perceived as significantly more false than the real version ($M = 3.67$, $SE = .11$), F (1, 269) = 104.16, $p < .001$, partial $\eta^2 = .28$. In addition, the false version of the post ($M = 2.91$, $SE = .11$) was perceived as less credible than the real version ($M = 4.41$, $SE = .11$), F (1, 269) = 88.08, $p < .001$, partial $\eta^2 = .25$.

Results for assessment of the story[6] revealed the same patterns, with a significant multivariate effect of type of content (fake vs. real), Wilks' $\Lambda = .69$, F (2, 268) = 61.61, $p < .001$ partial $\eta^2 = .32$, and a near-significant effect for bandwagon cue, Wilks' $\Lambda = .98$, F (2, 268) = 2.74, $p < .07$ partial $\eta^2 = .02$. The univariate analyses for fakeness of the story revealed that the false version of the story ($M = 4.64$, $SE = .11$) was perceived as more false compared to the real version ($M = 3.18$, $SE = .11$), F (1, 269) = 87.59, $p < .001$, partial $\eta^2 = .25$, and a significant main effect for bandwagon such that articles in the high bandwagon condition were perceived as more false ($M = 4.09$, $SE = .11$) compared to articles with low bandwagon ($M = 3.72$, $SE = .11$), F (1, 269) = 5.37, $p = .021$, partial $\eta^2 = .02$. The univariate analysis for credibility, on the other hand, only showed a significant main effect for type of content, such that the false stories ($M = 3.29$, $SE = .11$) were perceived as less credible than their real versions ($M = 4.94$, $SE = .11$), F (1, 269) = 111.74, $p < .001$, partial $\eta^2 = .29$.

To test H2 and assess the mediating role of fakeness perception, aroused curiosity, and emotions on the relationship between the type of content and engagement with the *post*, a series of mediation analyses with PROCESS macro Model 4 (Hayes, 2018) was conducted, one for each dependent variable (like, comment, read, share). All models were run with a 95% bias-corrected confidence interval based on 5,000 bootstrapped samples. Fakeness perception, aroused curiosity, and emotions (enjoyment, calm, alert) were entered as mediators. Issue involvement, political orientation, and story (two dummy coded variables) were entered as covariates. When entering "like," comment, or share as dependent variables, results reveal significant indirect effects via curiosity ($a_2b_{2liking} = -0.79$, CI [−1.08, −0.54]; $a_2b_{2commenting} = -0.79$, CI [−1.06, −0.54]; $a_2b_{2sharing} = -0.83$, CI [−1.11, −0.57]) and enjoyment ($a_3b_{3liking} = -0.30$, CI [−0.50, −0.14]; $a_3b_{3commenting} = -0.27$; CI [−0.48, −0.11], $a_3b_{3sharing} = -0.29$, CI [−0.49, −0.13]), such that users reported higher intentions to "like," comment, or share real (0) (vs. false news (1)) because it arouses curiosity and enjoyment (See Figures B2–B4 of Supplemental Appendix). The indirect effects through fakeness perception ($a_1b_{1liking} = 0.11$, CI [−0.11, 0.34]; $a_1b_{1commenting} = 0.14$, CI [−0.13, 0.39]; $a_1b_{1sharing} = 0.17$, CI [−0.09, 0.41]) and alert ($a_4b_{4liking} = 0.04$, CI [−0.01, 0.10]; $a_4b_{4commenting} = 0.03$, CI [−0.01, 0.08]; $a_4b_{4sharing} = 0.02$, CI [−0.01, 0.08]) were not significant for liking, commenting or sharing. However, there was an indirect effect through calm, but only for commenting behavior ($a_5b_{5liking}$=0.01, CI [−0.07, 0.09]; $a_5b_{5commenting}$ =0.09, CI [0.01, 0.22]; $a_5b_{5sharing}$=0.07, CI [−0.01, 0.17]), indicating that false news makes users feel less calm or more uneasy, which is associated with greater likelihood of commenting about it (compared to real news), but not liking or sharing. When entering intentions to read the story, only curiosity (a_2b_2=−1.11, CI [−1.47, −0.76]) was a significant mediator, again in the same direction (See Figure B5 of Supplemental Appendix). There was no significant difference for fakeness perception (a_1b_1=−0.13, CI [−0.36, 0.09]), enjoyment (a_3b_3=0.09, CI [−0.05, 0.23]), alert (a_4b_4=−0.01, CI [−0.06, 0.03]), or calm (a_5b_5=−0.03, CI [−0.14, 0.06]).

The same analyses were run to assess engagement with the *story*, by adding elaboration as another mediator. All models were run with a 95% bias-corrected confidence interval based on 5,000 bootstrapped samples. Data revealed significant indirect effects through fakeness perception ($a_1b_{1liking}$=0.21, CI [0.02, 0.40]; $a_1b_{1commenting}$=0.35, CI [0.14, 0.456]; $a_1b_{1sharing}$=0.14, CI [0.04, 0.25]), curiosity ($a_2b_{2liking}$=−1.01, CI [−1.31, −0.72]; $a_2b_{2commenting}$=−1.02, CI [−1.35, −0.72]; $a_2b_{2sharing}$=−0.44, CI [−0.61, −0.30]), enjoyment ($a_3b_{3liking}$=−0.45, CI [−0.70, −0.25]; $a_3b_{3commenting}$=−0.27, CI [−0.47, 0.11]; $a_3b_{3sharing}$=−0.16, CI [−0.27, −0.07]), and calm ($a_5b_{5liking}$=0.09, CI [0.0005, 0.20]; $a_5b_{5commenting}$=0.12, CI [0.02, 0.26]; $a_5b_{5sharing}$=0.07, CI [0.02, 0.15]) for all dependent variables. The indirect effect via alertness ($a_4b_{4liking}$=−0.0001, CI [−0.05, 0.05]; $a_4b_{4commenting}$=−0.001 CI [−0.03, 0.03]; $a_4b_{4sharing}$=−0.01, CI [−0.04, 0.02]) and elaboration ($a_6b_{6liking}$=0.04, CI [−0.03, 0.13]; $a_6b_{6commenting}$=0.02, CI [−0.02, 0.08]; $a_6b_{6sharing}$=0.02, CI [−0.02, 0.08]) were not significant (See Figures B6–B8 of Supplemental Appendix).

To answer RQ3 about the role of bandwagon perception, fakeness perception, and emotions, in the relationship between bandwagon cues, type of content and users' intentions to comment and read *the post*, two moderated mediation models using PROCESS macros Model 7 (Hayes, 2018) were run, one testing users' intention to read the article associated with the post and the other for users' intention to comment on the post. Type of content (real (0) vs. false (1)) was entered as the independent variable and bandwagon as the moderator (low (0) vs. high (1)). Bandwagon perception, fakeness perception, and emotions (enjoyment, alert, calm) were entered as mediators. Issue involvement, political orientation, and story (two dummy coded variables) were entered as covariates.

When entering "reading" intention as the dependent variable, a 95% bias-corrected confidence interval based on 5,000 bootstrap samples showed significant moderated mediation via alertness (Index=−0.20, CI [−0.49, −0.02]), such that after being presented with the real (vs. false) post, users felt more alert and attentive, in turn motivating reading intention, but only when the real news has a high number of likes (Indirect Effects=−0.15, CI [−0.35, −0.02)—there is no such mediation effect when the number of likes is low (Indirect Effects=0.04, CI [−0.06, 0.19]) (See Figure B9 of Supplemental Appendix). The indexes of moderated mediation were not significant for bandwagon perception (Index=−0.03, CI [−0.33, 0.25]), fakeness perception (Index=−0.01, CI [−0.18, 0.15]), enjoyment (Index=0.01, CI [−0.15, 0.15]), and calm (Index=0.04, CI

[−0.14, 0.28]). When entering intention to comment on the post as the dependent variable, on the other hand, analyses revealed a significant moderated mediation via calmness (Index = 0.21, CI [0.02, 0.50]). Participants felt less calm (more uneasy) when the story was false (vs. real) and had a high number of likes (Indirect Effects = 0.23, CI [0.04, 0.48]), leading them to express greater commenting intentions (See Figure B10 of Supplemental Appendix). There is no such mediation when the number of likes is low (Indirect Effects = 0.02, CI [−0.11, 0.16]). The indexes of moderated mediation were not significant for bandwagon perception (Index = −0.03, CI [−0.33, 0.25]), fakeness perception (Index = 0.001, CI [−0.07, 0.05]), enjoyment (Index = −0.21, CI [−0.54, 0.04]), and alert (Index = −0.04, CI [−0.21, 0.09]). These findings indicate that when real news has received many likes, people feel more alert, motivating them to read the story associated with the post. Conversely, when the article is false and has many likes, people feel more uneasy, leading to greater commenting intention.

To test RQ4 about the possible mechanisms for the effects of bandwagon and source cues on users' commenting behavior, two moderated mediation models using PROCESS macro Model 7 (Hayes, 2018) were run, one for participants' intention to comment before reading the story and one after. The indices of moderated mediation were not significant for either model (See Table A3 of the Supplemental Appendix for statistics).

Study 2 Discussion

In Study 2, we tested plausible explanatory mechanisms for the differential effects of interface cues on user engagement with real vs. false news articles. We found that these explanatory mechanisms vary based on the type of user action (like, comment, sharing) and whether it was performed before or after reading the complete article. When assessing engagement with the post (before reading the article), users reported higher intentions to read real posts further because they arouse more curiosity than false ones. "Liking," commenting, and sharing intentions were also higher for real (vs. false) posts because real posts elicit higher curiosity, and because users enjoy real posts more. These findings are consistent with recent research suggesting that, contrary to popular perception, non-clickbait arouses more curiosity than clickbait headlines (Molina et al., 2021b). Our data also showed that when a false post makes users feel uneasy, users have higher intentions to comment on them (compared to real posts), possibly because they want to set the record straight about its fakeness (Metzger et al., 2021). Furthermore, our findings suggest that the effects of content type (real vs. false) on user engagement with the post are also contingent on the number of likes the story has received. When a real post has received many likes, it makes people feel more alert and attentive, motivating them to read the article further. However, when a false post has received many likes, it makes users feel uneasy, motivating commenting behavior, even before reading the story.

Importantly, while users' motivation for engagement with the *post* varies based on the specific action, once users have read the story associated with the post, the reasons for engagement are the same for all actions (liking, commenting, and sharing). We found that after reading the article, users were able to identify a false article as false, as indicated by the higher fakeness perception of false articles (vs. real). Results of the mediation analyses further reveal that the perception of fakeness, in turn, motivates user action.

General Discussion

Overall, findings of this study suggest that engagement with false news online is not always driven by the believability of content. Consistent with the premises of Loewenstein (1994) and Slovic et al. (2007), our study demonstrates that engagement with different types of content is largely driven by curiosity elicited by characteristics of the content, as well as emotions elicited by both content characteristics and cues on the interface of social media platforms.

First, our findings suggest that users are more likely to read real rather than false news. This is because real news evokes more curiosity, as shown by Study 2. Higher curiosity and user enjoyment of real posts also result in "liking behavior." Molina et al. (2021b) found similar results—users were more likely to click "read more" for non-clickbait compared to clickbait headlines. Clickbait headlines were also perceived as less credible and less curiosity-arousing. These findings run counter to previous arguments that false news is more engaging because it evokes more curiosity (Dempsey, 2017). According to Loewenstein (1994), curiosity will be greater when information is perceived as better at accomplishing a particular task. It is possible that the sensationalistic nature of false news does not elicit much curiosity toward it when users do not perceive any information value in it. It is also possible that users' curiosity is fully satiated by false news headlines such that they may not feel the need to go beyond the post because false headlines often convey the crux of the story.

Furthermore, although participants generally engaged with real news more compared to false news, users commented more on false content than real content in Study 1. Study 2 replicated this effect and revealed that this occurs because false news makes users feel uneasy, motivating them to comment on the post even before reading the associated article. This uneasiness was also a significant factor after users read the story, mediating the relationship between the type of content and all engagement actions (like, comment, and share). Consistently, past research suggests that emotions play a key role in users' responses to misinformation. For example, Weeks (2015) found that when faced with false information, anger motivates partisan processing of information, while anxiety reduces the reliance on partisanship, suggesting that "the experience of anxiety can diminish the effects of motivated reasoning" (p. 702). Our study extends this research by revealing that the feeling of uneasiness derived from a false article is not only manifested in information processing, but also in user action.

Furthermore, in Study 2, we found that when individuals recognize a story as clearly false after they have read the associated story, they are more likely to like, comment, and share it. This finding suggests that sometimes users engage with false news even when they recognize it as false. In fact, comments left by users under false articles in Study 1 revealed more fakeness perception and were more analytical than comments on real stories. For example, one person said, "good example of an article just screaming to be 'fact checked'. . . which I did, and these quotes are taken WAY out of context!" A recent analysis of over 2.5 million social media comments revealed similar results—15% of user comments about false content expressed disbelief (Metzger et al., 2021). These findings suggest that there are other good reasons why users may be motivated to engage with false news. One possibility is that the ability to perceive fakeness in the story might make users feel agentic, which manifests not only affectively in terms of liking, but also behaviorally in terms of commenting and sharing. When users perceive an article as false, they might comment to warn others about its fakeness. Buttliere and Buder (2017) found that participants were more likely to respond to others when they disagreed with a position. The higher commenting on false articles is nonetheless problematic because it would contribute to the virality of the article and potentially reach other users who might not be able to interrogate the veracity of the content. As Vaccari et al. (2016) note, politically active citizens on social media are likely to reach users who are less engaged, exercising influence over them.

Although content type was a strong predictor of engagement, findings of both studies highlight the importance of interface cues in motivating user engagement with posts. For example, in Study 2, we found that false content paired with high bandwagon cues makes individuals feel worried and uneasy, which in turn motivates them to comment before reading. On the other hand, when real news is paired with high bandwagon cues, users felt more alert and attentive, in turn motivating reading intention. This means that the affect heuristic is not only triggered by content characteristics but also interface cues. The many media literacy campaigns informing users about bots' ability to amplify likes (e.g., Center for Information Technology & Society, 2020) might be motivating users to set the record straight and comment on the article to alert others, especially when content is suspected of being false. On the other hand, and consistent with previous research (e.g., Xu, 2013), when content is real, the number of likes serves as a positive heuristic cue that motivates them to read further. Overall, results suggest that a high bandwagon cue plays a persuasive role in attracting readership of real stories and motivating user action in the case of false stories.

Implications and Limitations

Research on misinformation has attributed user engagement with false information to users' lack of analytical thinking (Pennycook & Rand, 2019). Consistently, years of investigation on decision-making online show that it is seldom effortful. However, user behavior online is not arbitrary. The MAIN Model (Sundar, 2008) suggests that news consumers rely on heuristics, or cognitive rules of thumb, for decision-making and credibility assessment. While some heuristics are elicited by characteristics of the content itself—for example, false news is known to engage users by triggering emotions and accentuating user identity (Bakir & McStay, 2018; Kahan, 2017)—other heuristics derive from the interface, such as the bandwagon heuristic (Sundar, 2008). In this study, we explore the role of content characteristics and interface cues in user assessment of real versus false content online, and find that (1) user engagement is driven by emotions elicited by both content characteristics and heuristic cues on the interface, (2) each engagement action (liking, commenting, sharing, reading) has its own affective (enjoyment, calmness, alertness) and perceptual (fakeness) explanatory mechanism, and (3) these mechanisms differ based on whether the actions occur before or after users read the entire story.

While intentions to read a story associated with a social media post are driven by curiosity elicited by real content, commenting behavior is driven by positive and negative emotions elicited by content characteristics and heuristic cues on the interface. When a real post is perceived as enjoyable and curiosity-arousing, users seem to be motivated to comment. However, when the post is false, users feel uneasy and are more likely to comment on it, even without reading it first. This effect is stronger when false news has received many likes. Research exploring the bandwagon cue has consistently found that bandwagon cues increase credibility and engagement (e.g., Xu, 2013), but our research reveals that social endorsement cues may not always promote resharing or signify believability. When the post is clearly false, cues seem to propel users to become more agentic and express their unease by commenting on it. This raises questions about how to properly display metrics to users and categorize the valence of their resulting comments. This is especially important given that when users actually read the story and perceive it as false, they tend to like, comment, and share it maybe because they think other users will also perceive it as false. This is problematic because other users who receive the post might not read it further to assess its veracity. These findings also have implications for social media data scientists because they (1) suggest that we should not equate engagement with credibility—sometimes users are able to recognize falsity of content and still engage with it; and (2) highlight the importance of analyzing engagement actions beyond its quantitative value.

Our study also sheds light on the importance of automated journalism cues on users' engagement with content. Findings of Study 1 reveal that when users are aware that the post is written by a bot, they tend to be wary of engaging with content, specifically commenting, when these have high bandwagon cues. However, in Study 2, we were unable to detect the theoretical mechanisms of this effect. Given the rapid pace with which false content disseminates online and the fact that endorsement of this content is often amplified by bots, it is likely that users find this combination overly synthetic, with both the story writing and the endorsements coming from non-human entities. Clearly, more research is needed to understand user perceptions of bandwagons surrounding AI or bot-generated news stories.

In assessing our findings and their implications, it is important to note certain limitations. First, in the current experimental design, it is possible that participants engaged in more systematic processing of content than they would in a naturalistic setting. This might explain why our results showed that users engaged with real news more than false news. In addition, the website we used in these studies was not a real social media platform, and participants were not connected with other users. Their engagement could be different from real-life settings where their actions are public to their networks, and users' self-presentation motive, network characteristics, and the norm of the platform may all affect their decisions. Future research can consider exploring how results vary on different platforms and using actual engagement datasets to corroborate our findings. Additionally, while we utilized real-world examples of real and false articles, we purposefully selected articles where the false version was perceived as more false than its real counterpart. This is not always the case in our information environment where false information can sometimes be indistinguishable from real information. Finally, due to the design of Study 1, we were unable to include manipulation checks of our variables because participants interacted through a platform mimicking Facebook and saw the eight stories simultaneously (See Figure B1 of Supplemental Appendix). Randomization of experimental conditions occurred at the story level and not at the participant level. Thus, participants were exposed to all levels of the manipulations simultaneously and at random, making it infeasible to administer manipulation checks without interrupting their browsing of the stimulus site.

Despite its limitations, this study expands our understanding of user engagement with online information (false and real) by adding heuristic cues of the interface as important components driving these behaviors and by exploring the psychological mechanisms that drive each engagement action. Our study also extends the MAIN model (2008) in two ways. First, by suggesting that interface cues elicit different heuristics based on the message with which they are associated. Secondly, our study introduces and highlights the importance of the affect heuristic as a mechanism through which interface cues affect content perception on social media platforms. Theoretically, our findings invite future scholars to include interface cues as well as interactions between them in the array of factors that shape news engagement of online users. Furthermore, our results reveal that each engagement action should be analyzed in isolation, and that engagement does not necessarily mean believability. Online users are quite purposive in their engagement with real vs. false news stories, driven by curiosity and emotion, and cognizant of the level of bandwagon support for these stories when they decide to read, like, and comment.

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

This research is supported by the U. S. National Science Foundation (NSF) via Standard Grant No. CNS-1742702.

ORCID iD

Maria D. Molina https://orcid.org/0000-0001-8115-6303

Supplemental Material

Supplemental material for this article is available online.

Notes

1. Because all variables of this study were within-subjects, but not fully crossed since the independent variables were randomized per article, some participants received more than one instantiation of a particular condition (as evidenced by cells with $N > 171$) or less than one instantiation of a condition (as evidenced by cells with $N < 171$).
2. Although we measured these variables as control variables, we did not use them in the final analysis. The within- subjects design allows us to account for all potential individual differences that could have influenced engagement because the eight stories were seen by all participants.
3. We ran two models rather than one full-factorial model utilizing the three independent variables of interest because we did not have enough data points to reliably predict all the main effects and interaction terms within one model. This is because analyses were run only with participants who read the article first or commented/shared, reducing the number of instantiations to 606 from the original 1,368. Note as well that, of the 606 instantiations, only 71 represent comments or shares. Likewise, for the comparison between participants who liked first versus read first, instead of 1,368 data points, we had only 740, out of which 205 were likes.
4. While this result fell short of significance in Study 1, this might have occurred due to a lack of power of these analyses (See Footnote 3). Thus, we test this effect further in Study 2.
5. Box's test of equality of variance: $p = .03$. Levine test of equality of error variance: $p_{\text{credibility}} = .01$, $p_{\text{fakeness}} = .59$. We continued with the analysis despite assumption violation, given equal cell sizes (ranging from 32 to 40) and F-max lower than 10. We report Pillai's Trace for this analysis because it is more robust against such violations.
6. Box's test of equality of variance: $p = .60$. Levine test of equality of error variance: $p_{\text{fakeness}} = .38$, $p_{\text{credibility}} = .09$.

References

Aldous, K. K., An, J., & Jansen, B. J. (2019). View, like, comment, post: Analyzing user engagement by topic at 4 levels across 5 social media platforms for 53 news organizations. *Proceedings of the International AAAI Conference on Web and Social Media*, *13*(1), 47–57. https://ojs.aaai.org/index.php/ICWSM/article/view/3208

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*(2), 211–236. https://doi.org/10.1257/jep.31.2.211

Bakir, V., & McStay, A. (2018). Fake news and the economy of emotions: Problems, causes, solutions. *Digital Journalism*, *6*(2), 154–175. https://doi.org/10.1080/21670811.2017.1345645

Burke, M., & Kraut, R. E. (2016). The relationship between Facebook use and well-being depends on communication type and tie strength. *Journal of Computer-Mediated Communication*, *21*(4), 265–281. https://doi.org/10.1111/jcc4.12162

Buttliere, B., & Buder, J. (2017). Reading more vs. writing back: Situation affordances drive reactions to conflicting information on the internet. *Computers in Human Behavior*, *74*, 330–336. https://doi.org/10.1016/j.chb.2017.04.041

Center for Information Technology & Society. (2020). *How is fake news spread? Bots, people like you, trolls, and microtargeting*. https://www.cits.ucsb.edu/fake-news/spread

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, *113*(3), 554–559. https://doi.org/10.1073/pnas.1517441113

Dempsey, K. (2017, May). What's behind fake news and what you can do about it. *Information Today*. http://www.infotoday.com/IT/may17/Dempsey–Whats-Behind-Fake-News-and-What-You-Can-Do-About-It.shtml

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt, Brace, & Janovich.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.) Mcgraw-Hill Book Company.

Gabielkov, M., Ramachandran, A., Chaintreau, A., & Legout, A. (2016, June). *Social clicks: What and who gets read on Twitter?* [Conference session]. Proceedings of the 2016 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Science, Antibes Juan-les-Pins, France.

Gerbaudo, P., Marogna, F., & Alzetta, C. (2019). When "positive posting" attracts voters: User engagement and emotions in the 2017 UK election campaign on Facebook. *Social Media + Society*, *5*(4), 1–11. https://doi.org/10.1177/2056305119881695

Haim, M., & Graefe, A. (2017). Automated news: Better than expected? *Digital Journalism*, *5*(8), 1044–1059. https://doi.org/10.1080/21670811.2017.1345643

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). The Guilford Press.

Janét, K., Richards, O., & Landrum, A. R. (2020). Headline format influences evaluation of, but not engagement with, environmental news. *Journalism Practice*, *16*, 1–21. https://doi.org/10.1080/17512786.2020.1805794

Janoff-Bulman, R., Sheikh, S., & Baldacci, K. G. (2008). Mapping moral motives: Approach, avoidance, and political orientation. *Journal of Experimental Social Psychology*, *44*(4), 1091–1099. https://doi.org/10.1016/j.jesp.2007.11.003

Kahan, D. M. (2017). *Misconceptions, misinformation, and the logic of identity-protective cognition*. https://papers.ssrn.com/abstract=2973067

Knobloch-Westerwick, S., & Meng, J. (2011). Reinforcement of the political self through selective exposure to political messages. *Journal of Communication*, *61*(2), 349–368. https://doi.org/10.1111/j.1460-2466.2011.01543.x

Knobloch-Westerwick, S., Sharma, N., Hansen, D. L., & Alter, S. (2005). Impact of popularity indications on readers' selective exposure to online news. *Journal of Broadcasting & Electronic Media*, *49*, 296–313. https://doi.org/10.1207/s15506878jobem4903_3

Kormelink, T. G., & Meijer, I. C. (2018). What clicks actually mean: Exploring digital news user practices. *Journalism*, *19*(5), 668–683. https://doi.org/10.1177/1464884916688290

Liu, B., & Wei, L. (2018). *Reading machine-written news: Effect of machine heuristic and novelty on hostile media perception* [Conference session]. Proceedings of the 2018 International Conference on Human-Computer Interaction, Las Vegas, NV, USA.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*(1), 75–98. https://doi.org/10.1037/0033-2909.116.1.75

Luo, M., Hancock, J. T., & Markowitz, D. M. (2020). Credibility perceptions and detection accuracy of fake news headlines on social media: Effects of truth-bias and endorsement cues. *Communication Research*. 1–15. Advanced online publication. https://doi.org/10.1177/0093650220921321

Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication*, *60*(3), 413–439. https://doi.org/10.1111/j.1460-2466.2010.01488.x

Metzger, M. J., Flanagin, A. J., Mena, P., Jiang, S., & Wilson, C. (2021). From dark to light: The many shades of sharing misinformation online. *Media and Communication*, *9*, 134–143. https://doi.org/10.17645/mac.v9i1.3409

Molina, M. D., Sundar, S. S., Le, T., & Lee, D. (2021a). "Fake news" is not simply false information: A concept explication and taxonomy of online content. *American Behavioral Scientist*, *65*(2), 180–212. https://doi.org/10.1177/0002764219878224

Molina, M. D., Sundar, S. S., Rony, M. M. U., Hassan, N., Le, T., & Lee, D. (2021b, May). *Does clickbait actually attract more clicks? Three clickbait studies you must read* [Conference session]. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI'21), Yokohama, Japan.

Oeldorf-Hirsch, A., & Sundar, S. S. (2015). Posting, commenting, and tagging: Effects of sharing news stories on Facebook. *Computers in Human Behavior*, *44*, 240–249. https://doi.org/10.1016/j.chb.2014.11.024

Oh, J., Bellur, S., & Sundar, S. S. (2018). Clicking, assessing, immersing, and sharing: An empirical model of user engagement with interactive media. *Communication Research*, *45*(5), 737–763. https://doi.org/10.1177/0093650215600493

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

Reynolds, R. A. (1997). A validation test of a message elaboration measure. *Communication Research Reports*, *14*(3), 269–278. https://doi.org/10.1080/08824099709388670

Schimmack, U., & Grob, A. (2000). Dimensional models of core affect: A quantitative comparison by means of structural equation modeling. *European Journal of Personality*, *14*(4), 325–345. https://doi.org/10.1002/1099-0984(200007/08)14:4<325::aid-per380>3.0.CO;2-I

Shao, C., Ciampaglia, G. L., Varol, O., Yang, K.-C., Flammini, A., & Menczer, F. (2018). The spread of low-credibility content by social bots. *Nature Communications*, *9*(1), 4787. https://doi.org/10.1038/s41467-018-06930-7

Silverman, C. (2016, November 16). This analysis shows how viral fake election news stories outperformed real news on Facebook. *BuzzFeed*. https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2007). The affect heuristic. *European Journal of Operational Research*, *177*(3), 1333–1352. https://doi.org/10.1016/j.ejor.2005.04.006

Southwell, B. G., Thorson, E. A., & Sheble, L. (2017). The persistence and peril of misinformation. *American Scientist*, *105*(6), 372–375. https://doi.org/10.1511/2017.105.6.372

Sundar, S. S. (1999). Exploring receivers' criteria for perception of print and online news. *Journalism & Mass Communication Quarterly*, *76*(2), 373–386. https://doi.org/10.1177/107769909907600213

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). The MIT Press.

Tenenboim, O., & Cohen, A. A. (2015). What prompts users to click and comment: A longitudinal study of online news. *Journalism*, *16*(2), 198–217. https://doi.org/10.1177/1464884913513996

Vaccari, C., Valeriani, A., Barbera, P., Jost, J. T., Nagler, J., & Tucker, J. A. (2016). Of echo chambers and contrarian clubs: Exposure to political disagreement among German and Italian users of Twitter. *Social Media + Society*, *2*(3), 1–24. https://doi.org/10.1177/2056305116664221

Waddell, T. F. (2018). A robot wrote this? How perceived machine authorship affects news credibility. *Digital Journalism*, *6*(2), 236–255. https://doi.org/10.1080/21670811.2017.1384319

Weeks, B. E. (2015). Emotions, partisanship, and misperceptions: How anger and anxiety moderate the effect of partisan bias on susceptibility to political misinformation. *Journal of Communication*, *65*(4), 699–719. https://doi.org/10.1111/jcom.12164

Winter, S., & Krämer, N. C. (2014). A question of credibility – Effects of source cues and recommendations on information selection on news sites and blogs. *Communications*, *39*(4), 435–456. https://doi.org/10.1515/commun-2014-0020

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

Xu, W. W., Sang, Y., & Kim, C. (2020). What drives hyper-partisan news sharing: Exploring the role of source, style, and content. *Digital Journalism*, *8*(4), 486–505. https://doi.org/10.1080/21670811.2020.1761264

Yang, J. (2016). Effects of popularity-based news recommendations ("most-viewed") on users' exposure to online news. *Media Psychology*, *19*, 243–271. https://doi.org/10.1080/15213269.2015.1006333

Zell, A. L., & Moeller, L. (2018). Are you happy for me . . . on Facebook? The potential importance of "likes" and comments. *Computers in Human Behavior*, *78*, 26–33. https://doi.org/10.1016/j.chb.2017.08.050

Author Biographies

Maria D. Molina (PhD, Pennsylvania State University) is an Assistant Professor in the Advertising and PR Department at Michigan State University. Her research investigates the real-life and societal benefits of sharing online, and the role played by technological affordances in the process. Currently, three areas define her research agenda: the role played by technological affordances in promoting sharing of both positive content (e.g., reporting workouts in fitness apps for motivating healthy lifestyle) and negative content (e.g., spreading false news via social media), as well as user responses to automated tools using artificial intelligence (AI) for flagging problematic content such as misinformation and hate speech.

Jinping Wang (PhD, Pennsylvania State University) is an Assistant Professor in the School of Journalism and Communication at The Chinese University of Hong Kong. Her research lies at the intersection of the psychological effects of new media and strategic communication, focusing on how affordances of media technologies shape communication behaviors and communication effectiveness. She employs multiple methods in her research, including experiments, surveys, and computational analysis.

S. Shyam Sundar (PhD, Stanford University) is James P. Jimirro Professor of Media Effects and founding director of the Media Effects Research Laboratory at The Pennsylvania State University. His research investigates social and psychological effects of technological elements unique to online communication. His experiments investigate the role played by technological affordances in shaping user experience of mediated communications, in a variety of interfaces from websites and social media to mobile media and robotics. He edited the first-ever *Handbook of the Psychology of Communication Technology* (Blackwell Wiley, 2015) and served as editor-in-chief of the *Journal of Computer-Mediated Communication*, 2013 to 2017.

Thai Le is pursuing his PhD degree at the Pennsylvania State University. He is majoring in Information Science and Technology. Thai mainly contributes to the field of Machine Learning and AI. His research interest lies in trustworthy machine learning for NLP applications at large, with narrow topics in explainable AI, adversarial attack, robustness, and computational misinformation. He is an ACM member.

Carlina DiRusso (PhD, Pennsylvania State University) is an analytics consultant at Ketchum Inc. She's part of a team that helps corporations build their communications strategy through consumer research with a focus on healthcare and sustainability communications.