International Journal of Communication 19(2025), 3667–3690 1932–8036/20250005

Parental Perceptions of Dynamic Exchanges of Human Papillomavirus Vaccine Misinformation and Corrections From AI Checkers on Reddit

RITA TANG
University of Minnesota Twin Cities, USA

BENEDETTA BURSTON
Georgetown University, USA

JIKAI SUN
EMILY K. VRAGA
University of Minnesota Twin Cities, USA

LETICIA BODE[1]
Georgetown University, USA

As vaccination misinformation proliferates online and hesitancy increases in the United States, more research is needed to study how to best correct such myths. This study addresses two issues: exploring corrections from bot versus human actors and testing complicated social media interactions with two misinformation claims about human papillomavirus (HPV) vaccines. In our preregistered survey experiment (N = 1576) of parents, we found that when considering simple corrections to a single misinformation claim, a bot correction increased belief accuracy compared with a control (absent misinformation), whereas a user correction did not. However, as soon as a second, related false claim about the HPV vaccine was raised, audience beliefs about both false claims, as well as attitudes toward the vaccine, were stable, no matter the content and source of the corrections—from a repetitive bot, a responsive bot, or a social media user. We highlight the potential challenges of correction efforts in online environments.

Rita Tang: tang0768@umn.edu
Benedetta Burston: bjb142@georgetown.edu
Jikai Sun: sun00948@umn.edu
Emily K. Vraga: ekvraga@umn.edu
Leticia Bode: lb871@georgetown.edu
Date submitted: 2025-02-20

[1] This research was supported by a grant from the NSF, Award #2230692, and benefited from feedback from Dr. Michael Wagner, Sijia Yang, and Porismita Borah among others. We also appreciate the efforts of Yibing Sun, Liwei Shen, and Ji Soo Choi for their organization and support while fielding the study.

Keywords: artificial intelligence (AI), misinformation, correction, social media, conversation

There is growing scholarly concern about misinformation on social media (Farooq, Adlam, & Rutland, 2024; Farooq, Ketzitzidou Argyri, Adlam, & Rutland, 2022; Tang, Vraga, Bode, & Boulianne, 2024). Scholars often distinguish between misinformation—false information that is inconsistent with the best available scientific evidence—and disinformation, which is false information shared with deceptive intentions (Hameleers, Brosius, & de Vreese, 2022; Vraga & Bode, 2020). While this is an important theoretical distinction, in practice, at the content level, it is often difficult to distinguish the intent of sharing a single piece of content, making the determination between misinformation and disinformation challenging. For that reason, throughout this project, we refer to both types of false information collectively as misinformation.

Misinformation threatens society because it not only increases misbeliefs but also can discourage healthy behaviors—including vaccination intentions (Li & Yang, 2024; Schmid, Altay, & Scherer, 2023). One commonly offered solution to misinformation is providing corrections that introduce relevant facts that compete with inaccurate claims and, in doing so, can reduce misperceptions (Walter, Brooks, Saucier, & Suresh, 2021). However, much of the research on the benefits of corrections on social media in particular focuses on a simplified model, based on the immediate impact of a single correction (Bode & Vraga, 2025).

Two features of the modern social media environment complicate these recommendations. First, a single correction to misinformation may not end the conversation online. Instead, discussions are likely to evolve, often introducing new inaccurate claims surrounding the correction itself (Bond & Garrett, 2023). These responses may either endorse or challenge the correction, potentially affecting the effectiveness of the initial debunking efforts and the resulting beliefs of those witnessing the interactions (Mourali & Drake, 2022). This study focuses on the human papillomavirus (HPV) vaccine, an effective method for preventing the HPV infections most frequently associated with cancer (Centers for Disease Control and Prevention [CDC], 2024).

Second, technology is changing dramatically, particularly with the growing capabilities and popularity of large language model–powered chatbots like ChatGPT (Møller, Skovsgaard, & de Vreese, 2024; Tang, Fang, Sun, Bode, & Vraga, 2025). Bots are playing an increasingly important role across many social media platforms such as Reddit (Assenmacher et al., 2020), and interest in automated corrections is growing (Tang et al., 2025), using artificial intelligence (AI) for content moderation and decision making (Araujo, Brosius, Goldberg, Möller, & de Vreese, 2023; Araujo, Helberger, Kruikemeier, & de Vreese, 2020) and the potential for bots to offer corrections (Costello, Pennycook, & Rand, 2024). However, this emerging research has not yet answered whether bots are likely to be more effective correctors than human actors, especially given mixed evidence regarding their credibility (Edwards, Edwards, Spence, & Shelton, 2014; Liu & Wei, 2019).

Our study contributes to both of these open questions by simultaneously testing the effectiveness of bot versus human agents in correcting misinformation, both as a single correction to misinformation and as part of an extended conversation with multiple false and corrective claims. We study these two questions

in the context of misinformation and correction regarding HPV vaccines on Reddit, as these vaccines play a key role in reducing cervical cancer globally (World Health Organization [WHO], 2024b), and uptake of them remains below recommended levels in the United States (National Cancer Institute, 2024).

We used a preregistered survey experiment with parents of children not fully vaccinated for HPV (N = 1576) to answer these two fundamental research questions. Results showed that for a single misinformation claim, a correction labeled as coming from an "AI_checker" bot significantly improved belief accuracy compared with the control condition (absent any misinformation on the topic), while the user correction showed no significant effects. However, as soon as a second, related false claim was raised, audience beliefs about both specific false claims regarding the HPV vaccine—as well as attitudes toward the vaccine overall—were resistant to change, no matter what the source of the second correction was and whether the second correction was directly targeted at the second false claim. This research highlights the potential challenges of debunking efforts in online environments.

Literature Review

Observed Correction of Misinformation Reduces Misperceptions

Misinformation can directly lead to misperceptions (or belief in misinformation), along with related attitudes and behavioral choices. Although researchers argue that misinformation issues may be overestimated (e.g., Allen, Howland, Mobius, Rothschild, & Watts, 2020), its harm has been studied and validated in the context of vaccine beliefs and intentions (Carrieri, Madio, & Principe, 2019; Lee, Sun, Jang, & Connelly, 2022). Currently, the rise of AI-generated content can exacerbate existing concerns that misinformation is easy to generate and spread, but hard to detect (Aïmeur, Amri, & Brassard, 2023; Kreps, McCain, & Brundage, 2022), which could further worsen misinformation issues.

Given the harms associated with misinformation, multiple methods to tackle it have been proposed, including correction to specific misinformation claims, prebunking (inoculation), and media literacy interventions (Tang et al., 2025; Tang, Tully, Bode, & Vraga, 2025). While many factors affect the effectiveness of correction, including how the correction is labeled (Huang & Wang, 2022; Zhang, Featherstone, Calabrese, & Wojcieszak, 2021), correction through simple rebuttal or factual elaboration (Jin, van der Meer, Lee, & Lu, 2020; van der Meer & Jin, 2020), correction in a narrative or non-narrative way (Dahlstrom, 2021; Huang & Wang, 2022), and others, in general, correction has been consistently shown to effectively reduce belief in misinformation (Walter et al., 2021). Therefore, we generally expect correction to be effective in this study as well. However, existing research does not sufficiently account for two factors: whether the correction comes from a bot versus a human, and how the effectiveness of correction differs when multiple misinformation and correction claims intersect in a single conversation.

Bot Versus Human Correctors on Social Media

Bots, automated programs or accounts controlled by software that attempt to mimic human behavior, have emerged as significant actors capable of influencing public opinion (Chang, Chen, Zhang, Muric, & Ferrara, 2021). On one hand, malicious bots can play an important role in spreading misinformation

and low-credibility content (Shao et al., 2018). However, bots are not always malicious and can, in fact, be harnessed for good. For example, social media bots were used to challenge corporate hypocrisy (Armstrong, Neal, Tang, Rim, & Vraga, 2024).

Given these different potential uses for bots online, it is perhaps not surprising that research on the general evaluation of bot versus human sources (not in the context of corrections) produces mixed findings. For example, Edwards et al. (2014) compared the source credibility of two simulated Twitter accounts for the CDC, the national public health agency of the United States, one labeled as a CDC scientist and the other as an automated CDC Twitter bot. Their results suggest there is no difference in source credibility ratings between a CDC human scientist and the CDC Twitter bot, though they found that people rate the CDC human scientist as more attractive than the bot (Edwards et al., 2014). However, other research indicates that human journalists are perceived as more trustworthy and credible than robot journalists (Hong, Chang, & Tewksbury, 2024; Waddell, 2018), though people may also rate robot journalists as less biased (Hong, Chang, & Tewksbury, 2024). Furthermore, according to a meta-analysis, experimental evidence suggests that human-written news is perceived as more credible than automated news, though descriptive evidence suggests the opposite (Graefe & Bohlken, 2020). Given the mixed results in how people evaluate the two sources (bot versus human), it is unclear whether a correction from one source may outperform a correction from the other source.

In the specific context of correction, research is also mixed. In a study that tested socially recommended versus algorithmically recommended corrections, Huang and Wang (2022) found that the effects of the correction mechanism (social versus algorithmic), when accompanied by explicit endorsement, on issue attitudes were contingent on the type of correction shared (story versus non-story). However, a similar study where corrections came either from other social media users or from algorithmically produced "related stories" found that the two corrective mechanisms were equally effective (Bode & Vraga, 2018), as did a subsequent study comparing a bot and a human responding to COVID-19 misinformation on Facebook (Vraga & Bode, 2021). Therefore, while we expect both a bot and a human to be able to correct a single piece of misinformation on social media, their relative effectiveness remains an open question.

H1: Correction from either a bot or a user will (a) increase belief accuracy (in the first false claim)[2] and (b) increase positive attitudes toward the HPV vaccines, as compared with the control.

RQ1: Do bots or users produce (a) higher belief accuracy (in the first false claim) and (b) increased positive attitudes toward the HPV vaccines?

Social Media Conversation Does Not End With a Correction

Social media platforms advance a dialogic communication model, where users can and often do actively contribute to the flow of information through the affordances of the platform (Kent & Taylor, 2002). Although this leads to social media conversations that are often long and complicated, correction research has generally employed a simplified framework, in which a false claim is followed by a single correction

[2] Note that "belief accuracy" refers to decreased beliefs in the false claim.

(Bode & Vraga, 2018; Smith & Seitz, 2019; Vraga & Bode, 2017, 2018). This design, while effectively isolating the interaction of interest, does not allow for the possibility of a back-and-forth interaction between the two sides in the conversation. This simple format of misinformation claim rebutted by a single correction only imitates a portion of actual social media exchanges, excluding those that involve users engaging with, refuting, or affirming the correction (see Figure 1 as an example).

Because different reactions to a correction are possible on social media, this study examines what happens when the conversation does not end in a correction. Specifically, we focus on how people react when, in response to the first correction of the initial misinformation, another related piece of misinformation emerges in the conversation—a phenomenon we refer to as *a pivot to another false claim*. Each layer of misinformation and correction presents a new signal according to which information has to be recalibrated in mental frameworks (Botvinick, Braver, Barch, Carter, & Cohen, 2001). When conversations continue beyond a correction and pivot to another false claim, the additional information could overwhelm users with limited cognitive resources (Sweller, 1988). If users are overwhelmed, processing corrections to multiple misinformation claims—especially as compared with a baseline of no information on the topic—may be much more difficult. But if each instance of misinformation is still corrected (i.e., has a direct rebuttal with correct information), then existing evidence suggests these corrections can mitigate misperceptions (Walter et al., 2021). This is especially true when we compare it with when a second misinformation claim is offered, absent any rebuttal. Therefore, we offer both the following hypothesis and research question to capture these different relevant comparisons of whether correction is "successful."

H2: *A second correction by either a human or a bot will (a) increase belief accuracy (in both false claims) and (b) increase positive attitudes toward the HPV vaccines compared with when a pivot to another false claim goes uncorrected.*

RQ2: *Will any condition that includes a second correction after a pivot to another false claim (a) increase belief accuracy (in both false claims) and (b) increase positive attitudes toward the HPV vaccines compared with the control?*

Not All Bots Are Created Equal

Notable in the context of considering the effectiveness of bots as correctors of misinformation is their evolution over time. Particularly in recent years, with advances in large language models (LLMs) and generative AI, bots have become more sophisticated and better able to tailor their responses in conversations with users. Earlier social bots were equipped with automation modules, but many still fell short of being responsive (Assenmacher et al., 2020). For example, early social media bots often used scripts and automation tools to simulate user behavior, but did so in highly predictable patterns: They often posted at regular intervals or responded with templated phrases, which likely made them identifiable by their repetitive behaviors (Martini, Samula, Keller, & Klinger, 2021; Murthy et al., 2016). Moreover, these early bots could not engage in meaningful conversations and had limited capabilities in responding to social media comments (Assenmacher et al., 2020; Grimme, Preuss, Adam, & Trautmann, 2017). Social media bots have evolved with recent developments in AI, leveraging machine learning algorithms and more complex models that enable them to simulate humanlike interactions and adapt responses based on user

inputs (Hajli, Saeed, Tajvidi, & Shirazi, 2022). This adaptability allows these bots to engage in interactive, context-aware conversations, making them more effective tools for influencing public discourse and moderating content.

In the context of misinformation correction, smart bots with responsive correction strategies should be more effective than "dumb" bots that rely on repetitive corrections when combating misinformation, given their ability to generate more reasonable, meaningful, and context-aware interactions. Specifically, smart bots can adapt their responses based on the specifics of the misinformation being addressed and tailor their content in response to it, which may enhance the credibility and effectiveness of the correction. Indeed, early research suggests that extended conversations with LLM-enabled chatbots can result in decreased conspiracy beliefs (Costello et al., 2024). This adaptability should give them an advantage compared with the static, formulaic approach of "dumb" bots, which can appear mechanical and may be disregarded as irrelevant or spamlike. Thus, we propose the following hypothesis:

H3: A smart bot responding to a pivot to another false claim will produce (a) higher belief accuracy (in both false claims) and (b) increased positive attitudes toward the HPV vaccines, more than a repetitive bot.

Method

A preregistered[3] online survey experiment (N = 1576) was conducted to answer the research questions and test the hypotheses. This study was approved by the Institutional Review Board at the University of Wisconsin–Madison.

The topic of the experiment is the HPV vaccine, which provides safe, effective, and lasting protection against the HPV infections that most commonly cause cancer (CDC, 2024). However, according to the World Health Organization's (2024a) data, by 2023, only around 20% of 15-year-old girls had received the recommended doses of HPV vaccine globally. Misinformation on the HPV vaccine is prevalent on social media (Suarez-Lledo & Alvarez-Galvez, 2021), which can result in HPV vaccine hesitancy (Calo et al., 2021). For these reasons, the experiment focuses on the topic of the HPV vaccine.

[3] Find the preregistration (https://osf.io/tyjra) for Team I for details. In the preregistration, there are nine research questions and four hypotheses. We reported H1(a), H1 (c), RQ1 (a), RQ1 (d), H3 (a), H3 (d), H4 (a), H4 (d), RQ6 (a), and RQ6 (c), which correspond to H1 (a), H1 (b), RQ1 (a), RQ1(b), H2(a), H2 (b), RQ2 (a), RQ2 (b), H3 (a), and H3 (b), in the study because of the Special Section's focus on AI. Meanwhile, we only focused on two outcomes in this study because of the space limit of the article. Additionally, we also adjusted the wording of the hypotheses and questions to increase clarity, such as making it clear which claim we refer to for the outcome of belief accuracy and changing the wording of the outcome from "increase attitudes" to "increase positive attitudes."

The experiment involved seven conditions,[4] See Figures 1–2 as example stimuli and Figure 3 for an illustration of the experimental conditions: (1) a pure control group (who saw nothing), (2) a correction labeled as from an "AI_checker" bot in response to a false claim ("HPV vaccines increase the risk of MS and other autoimmune diseases. My kids will not be getting the shot—and everyone else should avoid it too!") posted on a Reddit-like social media platform (see Figure 1), (3) a user correction (identical to condition 2, but posted by a human user), (4) a pivot to a second false claim ("Well, even if it doesn't cause MS, my kids aren't even having sex, so why would they get a vaccine to protect them against a disease you get from sex?") posted by the a new (human) user, (5) the same bot repeating the initial correction in response to the second false claim (repetitive bot), (6) the same bot offering a different response directly targeted at the second false claim (responsive bot; see Figure 2), and (7) a new user correction to the second false claim (parallel to condition 6, but from a human user).

[4] There were eight conditions in total in the initial design (N = 1800), but given the focus of this article, we analyze only the seven relevant conditions (N = 1576). Condition labels were also adjusted from the preregistration because of this focus. Find Appendices A–D at https://osf.io/edgva/?view_only=c47493d70c6041ae9f4b224b9a1311fb. Specifically, see Appendix A for full descriptions of seven experimental conditions. Randomization checks were performed to ensure there were no significant differences across experimental conditions in terms of age, education, political affiliation, and gender (see Appendix B). Manipulation checks were performed to test the effectiveness of the source of the two corrections. To evaluate whether participants correctly perceived the source of the first correction (bot vs. user), we conducted a chi-squared test of independence comparing participants' responses across experimental conditions, modeled after prior research (van der Meer & Jin, 2020). The results revealed a significant association between condition and perceived correction source, $\chi^2(10, N = 1329) = 85.55$, $p < .001$. However, we must also note that despite the significant differences, overall, the recall accuracy of the correction source is modest: On average, between 46% and 60% of participants correctly identify the source of the first correction (see Appendix C for details). Meanwhile, to assess whether participants recognized the source of the second correction, a chi-squared test of independence was conducted. In this case, the association between experimental condition and perceived source of the second correction was not statistically significant, $\chi^2(4, N = 665) = 6.09$, $p = .19$. Additionally, as with the first correction, recall of the second correction source is modest, with between 37% and 49% correctly recalling it.

Figure 1. Example stimuli (bot correction to claim 1).

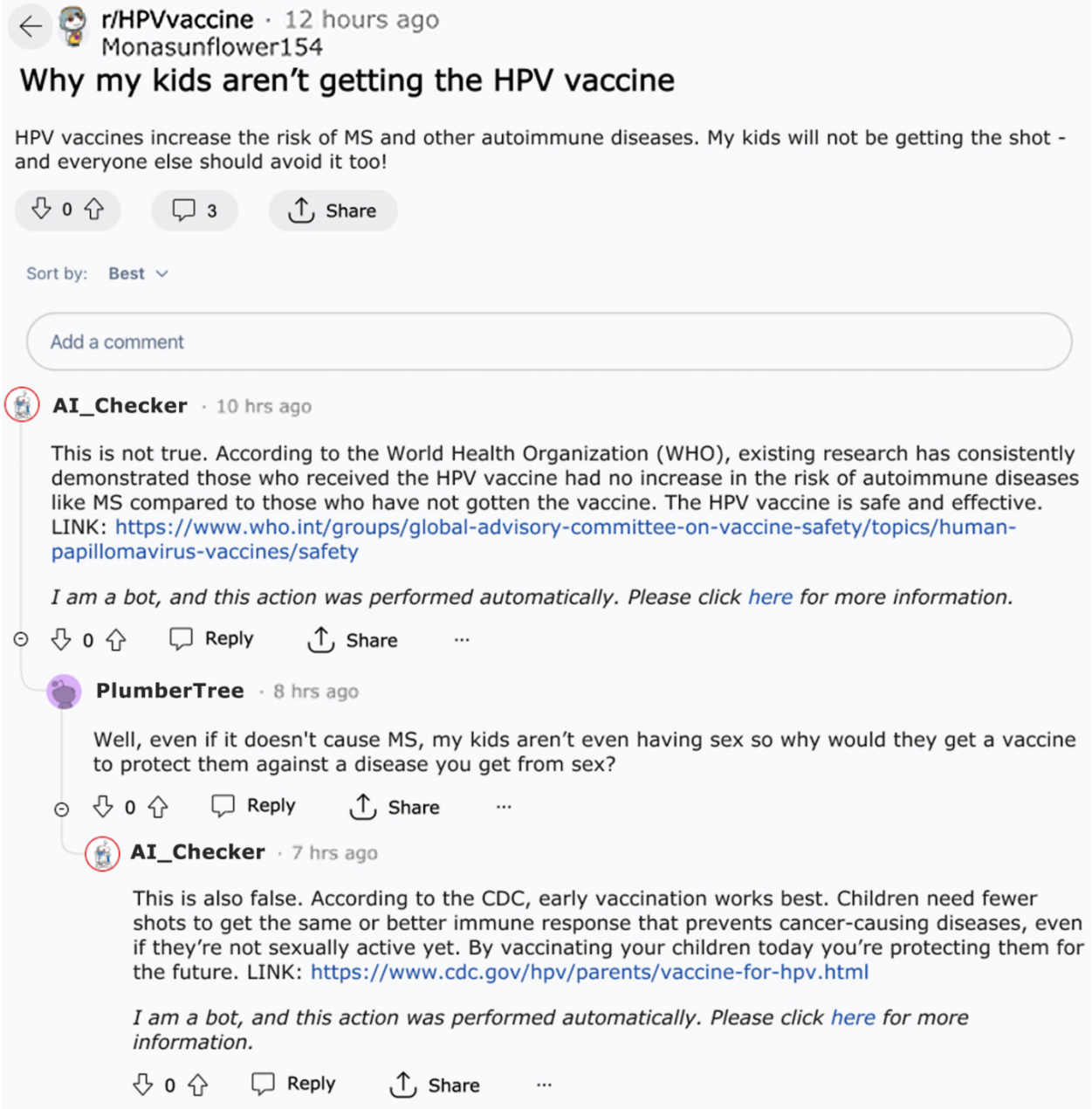

Figure 2. Example stimuli (bot responsive correction to claim 2).

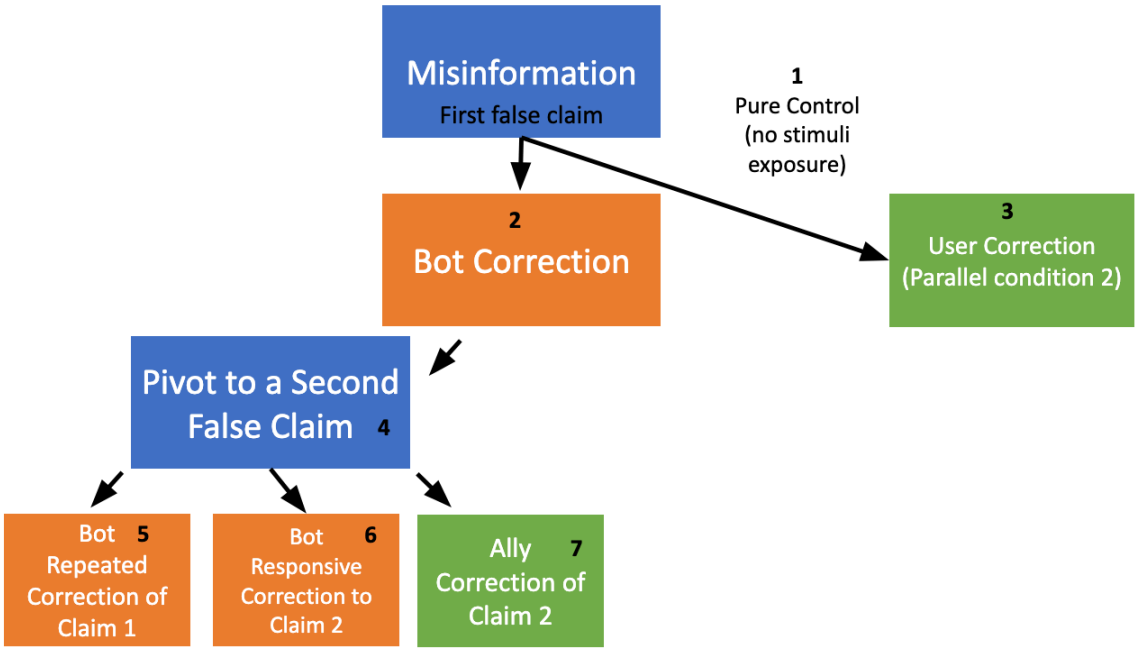

Figure 3. Illustration of the experimental conditions.

On entering the survey, participants answered a short pretest questionnaire, including their demographic characteristics, and then were randomly assigned to one of the seven conditions. After reading the stimuli, participants were asked to answer questions about key outcomes, including belief accuracy in the first false claim (HPV vaccines increase the risk of autoimmune diseases), belief accuracy in the second false claim (children are not sexually active, so they do not need HPV vaccines),[5] and attitudes toward HPV vaccines. Participants were debriefed at the end of the survey about the veracity of the information they had just seen.

Participants were recruited from the national online panel maintained by the survey company Lucid. Previous research suggests that parents declining vaccination for their children is a key reason for low HPV vaccination coverage in the United States (Kornides, McRee, & Gilkey, 2018); thus, our study specifically focuses on how parents perceive HPV vaccine-related misinformation and corrections. Participants were limited to parents currently living in the United States and able to read and write in English, including parents with younger children (< 9 years old) and those with older children (9–17 years old). We excluded those whose HPV vaccine-eligible children were already fully vaccinated. In this study, 1,576 participants are included in the analysis. 63.3% of participants were women; 47.4% were White (with 34.1% Black or African American, 10.3% Hispanic or Latin, and 8.2% other races); median education

[5] Note that these two false claims were adapted from common HPV vaccination myths (Taumberger et al., 2022).

was some college, no degree; median age was 35; median party was Independent; and median income was between $50,000 and $74,999 per year.

Measures

Belief Accuracy in the First False Claim

This was measured by asking participants to rate the accuracy of the first false claim, "HPV vaccines increase the risk of autoimmune diseases (e.g., MS, lupus, IBD)," on a 5-point scale, ranging from "definitely false" to "definitely true." Responses were reverse-coded so that higher scores indicated greater belief accuracy (M = 3.42, SD = 1.18).

Belief Accuracy in the Second False Claim

This was measured by asking participants to rate the accuracy of the false claim "Children and adolescents are not sexually active, so there is no need for them to get HPV vaccines at a young age," on a 5-point scale, ranging from "definitely false" to "definitely true." Similarly, responses were reverse-coded so that higher scores indicated greater belief accuracy (M = 3.57, SD = 1.28).

Positive Attitudes Toward HPV Vaccines

This was measured by asking participants to rate five sets of adjectives for the HPV vaccines on a bipolar 11-point scale, including "negative/positive," "bad/good," "useless/valuable," "foolish/wise," and "harmful/beneficial" (α = .94, M = 7.56, SD = 2.82).

Results

To address H1 and RQ1, we used a one-way ANOVA comparing those who saw a user correction, those who saw a bot correction, and those in the control condition, using Bonferroni comparisons as specified in the preregistration. The one-way ANOVA revealed a statistically significant effect of group on belief accuracy in the first false claim, $F(2, 685) = 3.21$, $p = .04$, $\eta^2 = .009$; for positive attitudes toward HPV vaccines, there was no significant group difference, $F(2, 685) = .02$, $p = .98$, $\eta^2 = .000$. Furthermore, the post hoc Bonferroni comparison results suggest H1a gained partial support, as those who saw a bot correction have more accurate beliefs for the first myth compared with those in the control condition (p = .04), but there is no significant difference between the control and user correction condition in terms of belief accuracy (p = 1.00). Moreover, the post hoc Bonferroni comparison results suggest H1b is not supported: There are no significant differences between the control and either the bot correction (p = 1.00) or user correction (p = 1.00) in terms of attitudes toward the HPV vaccine. Finally, per RQ1, there was no difference between bot correction and user correction for both outcomes (see Table 1).

Table 1. Comparing Conditions 2– 3 Against the Control Condition on Dependent Variables.

Condition	Belief accuracy in the first false claim Mean (SD)	Positive attitudes toward HPV vaccines Mean (SD)
Pure control (n = 241)	3.26 (1.18)	7.56 (2.92)
Bot correction (n = 221)	**3.53** (1.17)	7.59 (2.75)
User correction (n = 226)	3.34 (1.19)	7.61 (2.76)
ANOVA results	$F(2, 685) = 3.21$, $p = .04$, $\eta^2 = .009$	$F(2, 685) = .02$, $p = .98$, $\eta^2 = .000$

Note. Bolded numbers signal a significant difference from the control condition. Belief accuracy was measured on a 5-point Likert scale, while positive attitudes toward the HPV vaccines were measured on a bipolar 11-point scale, with a higher number indicating higher belief accuracy and more favorable attitudes toward the HPV vaccine.

H2 hypothesized that a second correction by either a human or a bot will (a) increase belief accuracy (in both false claims) and (b) increase positive attitudes toward the HPV vaccines compared with a second misinformation claim presented, but left uncorrected. Per the preregistration, we tested this hypothesis using a one-way ANOVA with Bonferroni comparisons among the four relevant conditions. For belief accuracy in the first false claim, the ANOVA was not significant, $F(3, 884) = .20$, $p = .89$, $\eta^2 = .001$. For belief accuracy in the second false claim, no significant group differences were found, $F(3, 884) = .74$, $p = .53$, $\eta^2 = .003$. Similarly, for positive attitudes toward HPV vaccines, the ANOVA was not significant, $F(3, 884) = .55$, $p = .65$, $\eta^2 = .002$. Therefore, H2 was rejected because results suggest there are no significant differences in individuals' belief accuracy in both claims and attitudes toward the HPV vaccines among these conditions (see Table 2).

Table 2. Comparing Conditions 4, 5, 6, and 7 on Dependent Variables.

Condition	Belief accuracy in the first false claim Mean (SD)	Belief accuracy in the second false claim Mean (SD)	Positive attitudes toward HPV vaccines Mean (SD)
A pivot to the second false claim (n = 218)	3.46 (1.13)	3.60 (1.27)	7.37 (2.85)
Bot repeated correction to the second false claim (n = 223)	3.42 (1.20)	3.45 (1.31)	7.58 (2.88)
Bot responsive correction to the second false claim (n = 220)	3.47 (1.16)	3.58 (1.31)	7.47 (2.72)
User correction to the second false claim (n = 227)	3.50 (1.20)	3.61 (1.28)	7.70 (2.84)
ANOVA results	$F(3, 884) = .20$, $p = .89$, $\eta^2 = .001$	$F(3, 884) = .74$, $p = .53$, $\eta^2 = .003$	$F(3, 884) = .55$, $p = .65$, $\eta^2 = .002$

Note. Belief accuracy was measured on a 5-point Likert scale, while positive attitudes toward the HPV vaccines were measured on a bipolar 11-point scale.

RQ2 is similar to the previous analysis, but compares the second correction from any source (repetitive bot, responsive bot, or human) with the pure control condition, where no misinformation regarding HPV vaccination was presented. Results from the preregistered one-way ANOVA (with four conditions, Bonferroni pairwise comparisons) produced no significant differences in terms of belief accuracy for either false HPV vaccine claims or attitudes toward HPV vaccines across these four conditions (see Table 3). Specifically, for belief accuracy in the first false claim, the ANOVA was not statistically significant, $F(3, 907) = 2.01$, $p = .11$, $\eta^2 = .007$. For belief accuracy in the second false claim, the ANOVA was also not significant, $F(3, 907) = .69$, $p = .56$, $\eta^2 = .002$. Similarly, there was no significant group effect on positive attitudes toward HPV vaccines, $F(3, 907) = .25$, $p = .86$, $\eta^2 = .001$.

This same ANOVA was used to test H3, which predicted that a responsive bot engaging in a second correction would (a) produce higher belief accuracy (in both false claims) and (b) increase positive attitudes toward the HPV vaccines more than a repetitive bot. H3 was also rejected because results show there are no differences in participants' belief accuracy (in both false claims) and attitudes toward the HPV vaccines when seeing a second correction from a repetitive versus responsive bot[6] (see Table 3).

[6] All the above results were consistent when eliminating those who failed the attention check, keeping only those who chose "strongly disagree" or "disagree" for the item "I swim the Atlantic Ocean to work every morning" (N = 1338).

Table 3. Comparing Conditions 1, 5, 6, and 7 on Dependent Variables.

Condition	Belief accuracy in the first false claim Mean (SD)	Belief accuracy in the second false claim Mean (SD)	Positive attitudes toward HPV vaccines Mean (SD)
Control condition (n = 241)	3.26 (1.18)	3.58 (1.29)	7.56 (2.92)
Bot repeated correction to the second false claim (n = 223)	3.42 (1.20)	3.45 (1.31)	7.58 (2.88)
Bot responsive correction to the second false claim (n = 220)	3.47 (1.16)	3.58 (1.31)	7.47 (2.72)
User correction to the second false claim (n = 227)	3.50 (1.20)	3.61 (1.28)	7.70 (2.84)
ANOVA results	$F(3, 907) = 2.01$, $p = .11$, $\eta^2 = .007$	$F(3, 907) = .69$, $p = .56$, $\eta^2 = .002$	$F(3, 907) = .25$, $p = .86$, $\eta^2 = .001$

Note. Belief accuracy was measured on a 5-point Likert scale, while positive attitudes toward the HPV vaccines were measured on a bipolar 11-point scale.

Discussion

This study set out to expand knowledge of how observed correction functions in complicated social media environments, wherein a single correction often does not end the conversation. However, social media is not just complicated by the number of responses but also by the actors who participate in the conversation. Given interest in the potential for AI bots to be sources of both misinformation (Shao et al., 2018) and correction (Costello et al., 2024; Vraga & Bode, 2021), we explored bots as a source of initial and second correction, as well as whether the second correction from a bot agent is responsive to new (misinformed) concerns or whether it simply repeats the same initial correction. The study suggests that while bots may have a slight advantage in correcting a single misinformation post, as soon as a second, related false claim is raised, audience beliefs about both specific false claims regarding the HPV vaccine and their attitudes toward the vaccine overall are resistant to change.

First, we find that a correction labeled as coming from an "AI_checker" bot, in response to a single misinformation claim on Reddit, leads to higher belief accuracy among the parent participants who witnessed the correction than those who saw nothing at all (i.e., neither misinformation nor correction). This finding echoes earlier research on observed correction arising from bots (Vraga & Bode, 2021), but also offers a useful takeaway for organizations hoping to scale their debunking activities. Bot corrections also have the added advantage of insulating the correctors from hostile attacks that often plague other experts who

attempt to respond to misinformation online (Kim & Shin, 2022; Royster et al., 2024), heightening their potential value to the correction landscape.

Nonetheless, we did not see these same benefits in terms of belief accuracy when the first correction came from a user. This null finding stands in contrast to a long line of existing work demonstrating successful observed correction from users (Bode & Vraga, 2025; Walter et al., 2021). Two explanations seem particularly promising for these null results. First, it may be that the misinformation regarding HPV vaccination is unusually difficult to correct, as has been shown for other vaccine misinformation (Bode & Vraga, 2015; Nyhan, Reifler, Richey, & Freed, 2014). Second, it may be that the affordances of the Reddit environment reduced the potential for user corrections to function as documented on other platforms. Indeed, the very features of Reddit—notably the presence of upvotes and downvotes—that make the platform democratic (Forestal, 2021) may be more important in explaining whether or not corrections are effective. In our study, we eliminated these cues (i.e., all posts had zero upvotes/downvotes); doing so may have thus signaled that the posts were not credible and therefore reduced their ability to correct misinformation. However, we also acknowledge that few of our participants were likely Reddit users, given that only 22% of the American public uses Reddit (Gottfried, 2024), so the absence of social cues may not have struck them as unusual. Unfortunately, our study cannot answer this question, but we hope future research can adopt other approaches to understand how social cues operate on Reddit, specifically in the context of misinformation, as well as whether social cues may be more important in reinforcing corrections from humans as compared with bot correctors.

Our second contribution to the literature is an initial exploration of how people respond to corrections in environments where several misinformation claims are being offered and corrected. Our results suggest that as soon as participants saw two different misinformation claims about the HPV vaccine, their attitudes became relatively fixed (at least immediately after their exposure to the stimuli). In other words, their belief accuracy, as well as their attitudes toward the HPV vaccine, did not differ from their initial beliefs entering the study, no matter what combination of misinformation and correction claims they saw or the source of those messages. Surprisingly, this was true for both a more responsive bot, which responded to evolving misinformation claims, and a repetitive bot simply repeating the same accurate information as the first correction. We had expected the responsive bot to be more effective (H3), given that it more closely resembles effective debunking conversations.

The fact that attitudes became relatively fixed after multiple misinformation claims and corrective responses can be seen as both a positive and a negative outcome. While it means people are not left better off seeing the misinformation and correction than when they saw nothing on the topic, they also were not worse off, and this makes sense in many ways. As John Zaller (1992) famously posited, two-sided and competing information flows are unlikely to produce mass persuasion, leaving people tethered to their existing beliefs. Stable attitudes in the face of competing information may be especially likely in the case of hard-to-change vaccination attitudes, but we argue strenuously that future research needs to examine observed correction in these more competitive information environments. It also complicates the recommendations to users and experts about how best to respond to misinformation publicly: If doing so is not a net positive nor a net negative, it may mean that encouraging people to respond to misinformation

when they are likely to get pushback is not the best use of resources, but this is something that needs much more testing in ever more realistic social media environments.

We also want to return to the question of whether bots may be especially well-suited for these kinds of interactions. We can offer only mixed evidence on this point. While the bot was successful (and a user was not) in single correction, neither the repetitive nor the responsive bot was able to increase belief accuracy when confronting a second related misinformation claim. However, we think the natural advantages of bots as correctors—in terms of both scalability and safety from harassment (Bautista, Zhang, & Gwizdka, 2021; Kim & Shin, 2022)—merit special attention to the role they can play. As AI becomes ever more advanced, bots may be able to serve in even more powerful roles, for example, by engaging in long conversations with people inclined toward misinformation or conspiracy theories that both ordinary users and especially expert sources may not have the bandwidth for—and in doing so, successfully address their evolving concerns (Costello et al., 2024).

Limitations

Of course, we must acknowledge several limitations of this study. First, we chose to study the important topic of HPV vaccines, but given the stickiness of vaccine attitudes, doing so may have limited our ability to perceive the effects of corrections, and the use of a single-message design may limit the ability to draw causal inferences and make it difficult to draw robust conclusions (Thorson, Wicks, & Leshner, 2012). Therefore, future research needs to replicate this research with another topic and use a multiple-message design to strengthen conclusions. Second, we chose to explore the understudied platform of Reddit for correction, but we cannot know how its affordances shaped audience responses. This is especially true because we did not sample Reddit users or place them in a true Reddit environment, but instead showed a simulated interaction between Reddit users. Future research needs to continue to push the boundaries of building realistic social media platforms where people can interact with information, and researchers can carefully track the interactions.

Third, there might be credibility differences between the first and second false claims (i.e., the second false claim may be harder to refute, which could also explain why people resist changing their attitudes after being exposed to the second piece of misinformation). Therefore, future research can switch the order of the two false claims and examine people's responses to them. Meanwhile, people's responses to the second false claim could arise from the claim itself—that children who are not sexually active do not need the HPV vaccine—or from the source of the claim, an account named PlumberTree. While this type of pseudonym is common on Reddit, some people may not recognize this as a human account. Our manipulation checks suggest that overall recall for both the first and second correction sources (AI vs. human user) is modest (between 37% and 60%; see Appendix C). While this generally aligns with correction recall on social media (Bode & Vraga, 2025), it may reduce the power of source cues in these spaces, although our supplemental analyses suggest our results are largely consistent when limited to those with accurate correction source recall. However, future research should investigate how to increase the salience

of source cues, especially for AI sources, and whether this influences the effectiveness of misinformation corrections.[7]

In addition, we test the effects of chatbots in a particular sociotechnical context, and these results may differ as that context changes. When this study was conducted, various political signals, such as the Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence (DiResta & Willner, 2023), highlighted the importance of guardrails for AI, including the LLMs that power chatbots, such as the ones in the experiment. In the current political environment, however, leaders of multiple technology firms are perceived to have garnered favor with the current U.S. administration, which some suggest may weaken regulations and safeguards around AI (Chan, Liedtke, O'Brien, Ortutay, & Parvini, 2024). Indeed, congressional legislation under consideration freezes state-level efforts to regulate AI (Hendrix & Lima-Strong, 2025). As the sociotechnical context changes, we might therefore expect chatbots to behave in different ways. For example, in the health sector, weakened regulations could lead to fewer requirements for transparency, accuracy, or pre- and post-deployment testing of LLMs. This, in turn, could reduce the alignment of chatbot outputs with public health guidance, increase the bias they display, and subject users to a greater likelihood of chatbot "hallucinations," a specific type of LLM-generated misinformation. At the same time, less regulatory scrutiny may also encourage broader and faster use of chatbots in health communication, expanding their role as a frontline tool for addressing misinformation if developers maintain accuracy safeguards voluntarily or public health users employ LLMs to train health-specific models on credible information. Therefore, future research should replicate the current study to see if effects persist as LLMs themselves, the regulations around them, and their users continue to change.

In addition, while not a limitation, we want to highlight the tricky ethical nature of working with chatbots. This is particularly true when bots imitate humans without appropriate disclosure—as in the case of many traditional bots on social media and in a recent instance involving ethical concerns about the undisclosed use of AI to generate comments on a Reddit subforum (Retraction Watch, 2025). Given that this is an evolving area, there is no hard and fast rule about under what circumstances it is ethically appropriate to use a chatbot to persuade anyone about anything. However, we think the fact that our experiment demonstrated that even a chatbot that is explicitly disclosed—in this case, called "AI_checker"—can be effective at offering people new information offers a promising avenue for ethical and transparent use of chatbots to share accurate information. Finally, while we did manipulate the source (i.e., we know people saw different sources; O'Keefe, 2003), we acknowledge that, overall, the recall accuracy of the correction sources is modest: On average, between 46% and 60% of participants correctly identified the source of the first correction, and between 37% and 49% correctly recalled the second correction source.

[7] There is only one notable difference when examining the preregistered hypotheses and research questions among those with accurate correction source recall: In terms of belief accuracy for the first false claim, among those who accurately recall the correction source, those in the bot-responsive correction to the second false claim condition have higher belief accuracy than those in the pure control condition. These results must be interpreted with caution, given the reduced sample size and differences in recall among the experimental conditions that may introduce biases into these analyses. For more details on these analyses, see Appendix D.

However, this may reflect the ecological validity of the results: This is realistic on social media, where people often do not recall the source of information, and we need to know whether and how different sources of misinformation corrections actually affect or do not affect public opinion (Bode & Vraga, 2025).

Conclusions

Ultimately, this study found that witnessing extended exchanges of misinformation and correction produced few changes in accurate beliefs or vaccination attitudes, at least after people see a second false claim. Future research should examine how the balance of accurate versus inaccurate information affects beliefs and consider how to mobilize a wide range of allies to support corrective efforts (Tang, 2025). These allies may include not only human agents but also bots, which may make corrections more scalable and safer, provided that both the ethical risks and benefits of relying on them are carefully evaluated.

References

Aïmeur, E., Amri, S., & Brassard, G. (2023). Fake news, disinformation and misinformation in social media: A review. *Social Network Analysis and Mining, 13*(1), 30. doi:10.1007/s13278-023-01028-5

Allen, J., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2020). Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances, 6*(14), eaay3539. doi:10.1126/sciadv.aay3539

Araujo, T., Brosius, A., Goldberg, A. C., Möller, J., & de Vreese, C. (2023). Humans vs. AI: The role of trust, political attitudes, and individual characteristics on perceptions about automated decision making across Europe. *International Journal of Communication, 17*, 6222–6249. Retrieved from https://ijoc.org/index.php/ijoc/article/view/20612

Araujo, T., Helberger, N., Kruikemeier, S., & de Vreese, C. H. (2020). In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & SOCIETY, 35*(3), 611–623. doi:10.1007/s00146-019-00931-w

Armstrong, S., Neal, C., Tang, R., Rim, H., & Vraga, E. K. (2024). Bot versus humans: Who can challenge corporate hypocrisy on social media? *Social Media + Society, 10*(4), 20563051241292578. doi:10.1177/20563051241292578

Assenmacher, D., Clever, L., Frischlich, L., Quandt, T., Trautmann, H., & Grimme, C. (2020). Demystifying social bots: On the intelligence of automated social media actors. *Social Media + Society, 6*(3), 205630512093926. doi:10.1177/2056305120939264

Bautista, J. R., Zhang, Y., & Gwizdka, J. (2021). US physicians' and nurses' motivations, barriers, and recommendations for correcting health misinformation on social media: Qualitative interview study. *JMIR Public Health and Surveillance, 7*(9), e27715. doi:10.2196/27715

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*(4), 619–638. doi:10.1111/jcom.12166

Bode, L., & Vraga, E. K. (2018). See something, say something: Correction of global health misinformation on social media. *Health Communication, 33*(9), 1131–1140. doi:10.1080/10410236.2017.1331312

Bode, L., & Vraga, E. K. (2025). *Observed correction: How we can all respond to misinformation on social media*. Oxford, UK: Oxford University Press.

Bond, R. M., & Garrett, R. K. (2023). Engagement with fact-checked posts on Reddit. *PNAS Nexus, 2*(3), pgad018. doi:10.1093/pnasnexus/pgad018

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*(3), 624–652. doi:10.1037/0033-295X.108.3.624

Calo, W. A., Gilkey, M. B., Shah, P. D., Dyer, A.-M., Margolis, M. A., Dailey, S. A., & Brewer, N. T. (2021). Misinformation and other elements in HPV vaccine tweets: An experimental comparison. *Journal of Behavioral Medicine, 44*(3), 310–319. doi:10.1007/s10865-021-00203-3

Carrieri, V., Madio, L., & Principe, F. (2019). Vaccine hesitancy and (fake) news: Quasi-experimental evidence from Italy. *Health Economics, 28*(11), 1377–1382. doi:10.1002/hec.3937

Centers for Disease Control and Prevention. (2024, August 20). *HPV vaccination*. Retrieved December 18, 2024, from https://www.cdc.gov/hpv/vaccines/index.html

Chan, K., Liedtke, M., O'Brien, M., Ortutay, B., & Parvini, S. (2024). What does Big Tech hope to gain from warming up to Trump? *Associated Press.* Retrieved from https://apnews.com/article/trump-big-tech-ceo-donations-visits-1ddf9464047b519abe5593244be488d9

Chang, H.-C. H., Chen, E., Zhang, M., Muric, G., & Ferrara, E. (2021). Social bots and social media manipulation in 2020. In U. Engel, A. Quan-Haase, S. X. Liu, & L. Lyberg (Eds.), *Handbook of Computational Social Science, Volume 1* (1st ed., pp. 304–323). London, UK: Routledge. doi:10.4324/9781003024583-21

Costello, T. H., Pennycook, G., & Rand, D. G. (2024). Durably reducing conspiracy beliefs through dialogues with AI. *Science, 385*(6714), eadq1814. doi:10.1126/science.adq1814

Dahlstrom, M. F. (2021). The narrative truth about scientific misinformation. *Proceedings of the National Academy of Sciences, 118*(15), e1914085117. doi:10.1073/pnas.1914085117

DiResta, R., & Willner, D. (2023). White House AI executive order takes on complexity of content integrity issues. *Tech Policy Press.* Retrieved from https://www.techpolicy.press/white-house-ai-executive-order-takes-on-complexity-of-content-integrity-issues/

Edwards, C., Edwards, A., Spence, P. R., & Shelton, A. K. (2014). Is that a bot running the social media feed? Testing the differences in perceptions of communication quality for a human agent and a bot agent on Twitter. *Computers in Human Behavior, 33*, 372–376. doi:10.1016/j.chb.2013.08.013

Farooq, A., Adlam, A., & Rutland, A. (2024). Rejecting ingroup loyalty for the truth: Children's and adolescents' evaluations of deviant peers within a misinformation intergroup context. *Journal of Experimental Child Psychology, 243*, 105923. doi:10.1016/j.jecp.2024.105923

Farooq, A., Ketzitzidou Argyri, E., Adlam, A., & Rutland, A. (2022). Children and adolescents' ingroup biases and developmental differences in evaluations of peers who misinform. *Frontiers in Psychology, 13*. doi:10.3389/fpsyg.2022.835695

Forestal, J. (2021). *Designing for democracy: How to build community in digital environments*. Oxford, UK: Oxford University Press.

Gottfried, J. (2024, January 31). *Americans' social media use*. Pew Research Center. Retrieved from https://www.pewresearch.org/internet/2024/01/31/americans-social-media-use/

Graefe, A., & Bohlken, N. (2020). Automated journalism: A meta-analysis of readers' perceptions of human-written in comparison to automated news. *Media and Communication, 8*(3), 50–59. doi:10.17645/mac.v8i3.3019

Grimme, C., Preuss, M., Adam, L., & Trautmann, H. (2017). Social bots: Human-like by means of human control? *Big Data, 5*(4), 279–293. doi:10.1089/big.2017.0044

Hajli, N., Saeed, U., Tajvidi, M., & Shirazi, F. (2022). Social bots and the spread of disinformation in social media: The challenges of artificial intelligence. *British Journal of Management, 33*(3), 1238–1253. doi:10.1111/1467-8551.12554

Hameleers, M., Brosius, A., & de Vreese, C. H. (2022). Whom to trust? Media exposure patterns of citizens with perceptions of misinformation and disinformation related to the news media. *European Journal of Communication, 37*(3), 237–268. doi:10.1177/02673231211072667

Hendrix, J., & Lima-Strong, C. (2025, May 22). US House passes 10-year moratorium on State AI laws | TechPolicy.Press. *Tech Policy Press*. Retrieved from https://techpolicy.press/us-house-passes-10year-moratorium-on-state-ai-laws

Hong, J.-W., Chang, H.-C. H., & Tewksbury, D. (2024). Can AI become Walter Cronkite? Testing the machine heuristic, the hostile media effect, and political news written by artificial intelligence. *Digital Journalism, 13*(4), 845–868. doi:10.1080/21670811.2024.2323000

Huang, Y., & Wang, W. (2022). When a story contradicts: Correcting health misinformation on social media through different message formats and mechanisms. *Information, Communication & Society, 25*(8), 1192–1209. doi:10.1080/1369118X.2020.1851390

Jin, Y., van der Meer, T. G. L. A., Lee, Y.-I., & Lu, X. (2020). The effects of corrective communication and employee backup on the effectiveness of fighting crisis misinformation. *Public Relations Review, 46*(3), 101910. doi:10.1016/j.pubrev.2020.101910

Kent, M. L., & Taylor, M. (2002). Toward a dialogic theory of public relations. *Public Relations Review, 28*(1), 21–37. doi:10.1016/S0363-8111(02)00108-X

Kim, C., & Shin, W. (2022). Harassment of journalists and its aftermath: Anti-press violence, psychological suffering, and an internal chilling effect. *Digital Journalism, 13*(2), 232–248. doi:10.1080/21670811.2022.2034027

Kornides, M. L., McRee, A. L., & Gilkey, M. B. (2018). Parents who decline HPV vaccination: Who later accepts and why? *Academic Pediatrics, 18*(2), S37–S43. doi:10.1016/j.acap.2017.06.008

Kreps, S., McCain, R. M., & Brundage, M. (2022). All the news that's fit to fabricate: AI-generated text as a tool of media misinformation. *Journal of Experimental Political Science, 9*(1), 104–117. doi:10.1017/XPS.2020.37

Lee, S. K., Sun, J., Jang, S., & Connelly, S. (2022). Misinformation of COVID-19 vaccines and vaccine hesitancy. *Scientific Reports, 12*(1), 13681. doi:10.1038/s41598-022-17430-6

Li, J., & Yang, X. (2024). Does exposure necessarily lead to misbelief? A meta-analysis of susceptibility to health misinformation. *Public Understanding of Science, 34*(2), 222–242. doi:10.1177/09636625241266150

Liu, B., & Wei, L. (2019). Machine authorship in Situ. *Digital Journalism, 7*(5), 635–657. doi:10.1080/21670811.2018.1510740

Martini, F., Samula, P., Keller, T. R., & Klinger, U. (2021). Bot, or not? Comparing three methods for detecting social bots in five political discourses. *Big Data & Society, 8*(2), 20539517211033566. doi:10.1177/20539517211033566

Møller, L. A., Skovsgaard, M., & de Vreese, C. (2024). Reinforce, readjust, reclaim: How artificial intelligence impacts journalism's professional claim. *Journalism, 26*(7), 1373–1390. doi:10.1177/14648849241269300

Mourali, M., & Drake, C. (2022). The challenge of debunking health misinformation in dynamic social media conversations: Online randomized study of public masking during COVID-19. *Journal of Medical Internet Research, 24*(3), e34831. doi:10.2196/34831

Murthy, D., Powell, A. B., Tinati, R., Anstead, N., Carr, L., Halford, S. J., & Weal, M. (2016). Automation, algorithms, and politics| bots and political influence: A sociotechnical investigation of social network capital. *International Journal of Communication, 10*, 4952–4971. Retrieved from https://ijoc.org/index.php/ijoc/article/view/6271

National Cancer Institute. (2024, March). *HPV vaccination*. Retrieved October 17, 2024, from https://progressreport.cancer.gov/prevention/hpv_immunization

Nyhan, B., Reifler, J., Richey, S., & Freed, G. L. (2014). Effective messages in vaccine promotion: A randomized trial. *Pediatrics, 133*(4), 835–842. doi:10.1542/peds.2013-2365

O'Keefe, D. J. (2003). Message properties, mediating states, and manipulation checks: Claims, evidence, and data analysis in experimental persuasive message effects research. *Communication Theory, 13*(3), 251–274. doi:10.1111/j.1468-2885.2003.tb00292.x

Retraction Watch. (2025). *Experiment using AI-generated posts on Reddit draws fire for ethics concerns*. Retrieved from https://retractionwatch.com/2025/04/28/experiment-using-ai-generated-posts-on-Reddit-draws-fire-for-ethics-concerns/

Royster, J., Meyer, J. A., Cunningham, M. C., Hall, K., Patel, K., McCall, T. C., & Alford, A. A. (2024). Local public health under threat: Harassment faced by local health department leaders during the COVID-19 pandemic. *Public Health in Practice, 7*, 100468. doi:10.1016/j.puhip.2024.100468

Schmid, P., Altay, S., & Scherer, L. D. (2023). The psychological impacts and message features of health misinformation: A systematic review of randomized controlled trials. *European Psychologist, 28*(3), 162–172. doi:10.1027/1016-9040/a000494

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. doi:10.1038/s41467-018-06930-7

Smith, C. N., & Seitz, H. H. (2019). Correcting misinformation about neuroscience via social media. *Science Communication, 41*(6), 790–819. doi:10.1177/1075547019890073

Suarez-Lledo, V., & Alvarez-Galvez, J. (2021). Prevalence of health misinformation on social media: Systematic review. *Journal of Medical Internet Research, 23*(1), e17187. doi:10.2196/17187

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285. doi:10.1207/s15516709cog1202_4

Tang, R. (2025). Effects of problem-recognition messages from different sources and cues-to-action on promoting corrective efforts on social media. *Journal of Communication Management*. Advance online publication. doi:10.1108/JCOM-09-2024-0181

Tang, R., Fang, Y., Sun, J., Bode, L., & Vraga, E. K. (2025). Automating accuracy: Scalable approaches to correcting disinformation with artificial intelligence on social media. *Journalism & Mass Communication Quarterly*, 10776990251359660. doi:10.1177/10776990251359660

Tang, R., Tully, M., Bode, L., & Vraga, E. K. (2025). Effects of a news literacy video on news literacy perceptions and misinformation evaluation. *Media and Communication, 13*, 8983. doi:10.17645/mac.8983

Tang, R., Vraga, E. K., Bode, L., & Boulianne, S. (2024). Who reports witnessing and performing corrections on social media in the United States, United Kingdom, Canada, and France? *Harvard Kennedy School Misinformation Review*. doi:10.37016/mr-2020-145

Taumberger, N., Joura, E. A., Arbyn, M., Kyrgiou, M., Sehouli, J., & Gultekin, M. (2022). Myths and fake messages about human papillomavirus (HPV) vaccination: Answers from the ESGO Prevention Committee. *International Journal of Gynecologic Cancer, 32*(10), 1316–1320. doi:10.1136/ijgc-2022-003685

Thorson, E., Wicks, R., & Leshner, G. (2012). Experimental methodology in journalism and mass communication research. *Journalism & Mass Communication Quarterly, 89*(1), 112–124. doi:10.1177/1077699011430066

van der Meer, T. G. L. A., & Jin, Y. (2020). Seeking formula for misinformation treatment in public health crises: The effects of corrective information type and source. *Health Communication, 35*(5), 560–575. doi:10.1080/10410236.2019.1573295

Vraga, E. K., & Bode, L. (2017). Using expert sources to correct health misinformation in social media. *Science Communication, 39*(5), 621–645. doi:10.1177/1075547017731776

Vraga, E. K., & Bode, L. (2018). I do not believe you: How providing a source corrects health misperceptions across social media platforms. *Information, Communication & Society, 21*(10), 1337–1353. doi:10.1080/1369118X.2017.1313883

Vraga, E. K., & Bode, L. (2020). Defining misinformation and understanding its bounded nature: Using expertise and evidence for describing misinformation. *Political Communication, 37*(1), 136–144. doi:10.1080/10584609.2020.1716500

Vraga, E. K., & Bode, L. (2021). Addressing COVID-19 misinformation on social media preemptively and responsively. *Emerging Infectious Diseases, 27*(2), 396–403. doi:10.3201/eid2702.203139

Waddell, T. F. (2018). A robot wrote this? *Digital Journalism, 6*(2), 236–255. doi:10.1080/21670811.2017.1384319

Walter, N., Brooks, J. J., Saucier, C. J., & Suresh, S. (2021). Evaluating the impact of attempts to correct health misinformation on social media: A meta-analysis. *Health Communication, 36*(13), 1776–1784. doi:10.1080/10410236.2020.1794553

World Health Organization. (2024a, July 15). *Percentage of 15 years old girls received the recommended doses of HPV vaccine*. Datadot. Retrieved December 18, 2024, from https://data.who.int/indicators/i/A7398F0/287D1D2

World Health Organization. (2024b, October 4). *WHO adds an HPV vaccine for single-dose use*. Retrieved October 17, 2024, from https://www.who.int/news/item/04-10-2024-who-adds-an-hpv-vaccine-for-single-dose-use

Zaller, J. R. (1992). *The nature and origins of mass opinion*. Cambridge, UK: Cambridge University Press. Retrieved from https://www.cambridge.org/core/books/nature-and-origins-of-mass-opinion/70B1485D3A9CFF55ADCCDD42FC7E926A

Zhang, J., Featherstone, J. D., Calabrese, C., & Wojcieszak, M. (2021). Effects of fact-checking social media vaccine misinformation on attitudes toward vaccines. *Preventive Medicine, 145*, 106408. doi:10.1016/j.ypmed.2020.106408