

Misinformation and Literacies in the Era of Generative Artificial Intelligence: A Brief Overview and a Call for Future Research

Emerging Media

1–16

© The Author(s) 2024

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/27523543241240285

journals.sagepub.com/home/emm**Chun Chu-Ke**

Mt. Dawu College, National Pingtung University, Pingtung, Taiwan

Yujie Dong 

School of Media and Communication, Shanghai Jiao Tong University, Shanghai, China

Abstract

Misinformation constitutes a societal practice and challenge that necessitates unwavering attention worldwide. In this essay, we discussed the theoretical advancement and empirical evidence in misinformation research, encompassing a review of definitions of misinformation, research orientations, research perspectives, and vulnerable groups. We then reviewed the misinformation fueled by generative artificial intelligence (AI) and the evolving conceptualization of literacy. To counter AI-fueled misinformation, we argue that the development of ethical AI necessitates regulations from AI practitioners and legislation, and ethical uses of AI require efforts in AI literacy education and research. The AI literacy should include (a) users' understanding and critical evaluation of knowledge, values, and cultures within which AI systems function, and their implications on the AI-generated content, (b) users' strategic interpretation and proper use of AI-generated content, and (c) users' utilization of feedback mechanisms to promote institutional management of the AI power.

Keywords

misinformation, artificial intelligence, literacy, AI literacy, cognition, human-AI interaction

Corresponding author:

Yujie Dong, School of Media and Communication, Shanghai Jiao Tong University, 800 Dongchuan Road, Minhang District, Shanghai, China, 200240.

Email: dorisdong@sjtu.edu.cn

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Introduction

With the emergence of social media and mobile media in the early 2000s, the spread of misinformation, a communication phenomenon that has existed for thousands of years, has accelerated at a speed faster than ever (Ha et al., 2021). Around the 2016 U.S. presidential election, misinformation and its news form, fake news, have been suggested to play a prominent role in influencing voters' choices and accelerating political polarization (Howard et al., 2018). Later, the COVID-19 pandemic facilitated a surge in misinformation and fake news worldwide, which produced devastating influences and (Roozenbeek et al., 2020). Such misinformation that is inextricably interwoven with social events and realities has widely attracted the public's attention as well as concerns. In other words, it has become an important public agenda around the globe.

Since 2017, academia has witnessed exponential growth in the research on misinformation and fake news. In light of the prevalence and the perseverance of misinformation and fake news, researchers examined not only individuals' exposure to, consumption of, and sharing of fake news during the U.S. election season (e.g. Allcott & Gentzkow, 2017; Grinberg et al., 2019) but also paid attention to the spread of fake news in other social domains, such as natural disasters, science, terrorism, finance, and health (e.g. Krishna & Thompson, 2021; Vosoughi et al., 2018) and in other non-western socio-cultural contexts, such as Asian societies (Chen et al., 2022). Following the investigation of components and communication of misinformation, there is a surge of research on misinformation correction. Studies probe the effectiveness of various message strategies in correcting misinformation, such as politeness (Malhotra & Pearce, 2022), debiasing (Dai et al., 2021), pre-bunking, and inoculation (Lewandowsky and van der Linden, 2021).

The importance of media literacy and relevant interventions in misinformation correction campaigns has been increasingly emphasized. Media literacy initially describes the skilled-based ability to access, analyze, evaluate, and communicate media information that normally has application to K-12 and higher education (Aufderheide, 1997; Christ & Potter, 1998). As the representation of knowledge and the conduct of communication are constantly changing due to the media and technology advances, media literacy can be broadly considered "part of a package of measures to lighten top-down media regulation by devolving responsibility for media use from the state to individuals, a move which can be interpreted either as 'empowering'" (Livingstone, 2004, p. 10). The prevalence of misinformation and fake news on social media poses challenges to media literary education and warrants research on digital media literacy programs "outside the traditional classroom" (Lee, 2018, p. 461).

Today, as generative artificial intelligence (AI) has boosted the creation and spread of misinformation (Galaz et al., 2023; Zhou et al., 2023), both misinformation and literacy needed in misinformation correction should be reconsidered. The present study first aims to provide a quick overview of theories and empirical research under the umbrella term misinformation. We then

argue how generative AI may influence the production and communication of misinformation and how the definition and interventions of literacy should adapt to the development of generative AI. A call for future research on misinformation and literacy is presented at the end.

An Overview of Misinformation: Theory and Empirical Evidence

Characterizing misinformation and related research

To avoid the blur of definitions and conceptualizations of misinformation-related terms, we in this essay primarily adopt the term *misinformation* to cover a wide array of information that is not accurate or false. Disinformation can be considered a specific type of misinformation that is intentionally false (Scheufele & Krause, 2019), and fake news is also deemed one type of misinformation presented in the form of news (Ha et al., 2021). Scholars may base their definitions of misinformation on different research contexts. For example, political science researchers define misinformation as false or misleading information that is believed to be true and related to one's political preferences (Nyhan & Reifler, 2010), or "incorrect, but confidently held, political beliefs" (Jerit & Zhao, 2020, p. 79). It is concerning that misinformed individuals may take political actions upon political misinformation (Jerit & Zhao, 2020). In the science and health domain, misinformation is defined as "information that is contrary to the epistemic consensus of the scientific community regarding a phenomenon" (Swire-Thompson & Lazer, 2019, p. 434). Studies on health misinformation mainly focus on misinformation related to vaccines and infectious diseases (Wang et al., 2019).

The orientations and topics of misinformation research can differ across regions. The authors conducted a brief literature review of articles related to "fake news" published in prestigious communication journals in mainland China, Taiwan, and the United States. The purpose was to explore differences between articles published in different regions in their research orientations. Considering the 2016 U.S. presidential election constitutes a catalyst for the spike of research on misinformation (Li, 2020), the time frame of this research was set between 2017 and 2021. The preliminary analyses (see Figure 1) showed that a large proportion of articles published in mainland China and Taiwan were (often non-empirical) review articles, case analyses, and essays, and they put great emphasis on the institutional-level factors centering around fake news and legislation against fake news. Comparatively, most articles published in the United States explored the influences of fake news on individuals' attitudes, behaviors, and social interactions particularly.

Misinformation, media bias, and psychological determinants

Research on the production of misinformation can be traced back to the investigation of media bias, which is anchored in the perspective of structural influences. As the intensity of mediated experiences, institutions, events, and processes beyond one's reach have increased, media has become a vital channel for information exchange and communication and breeds mediated politics

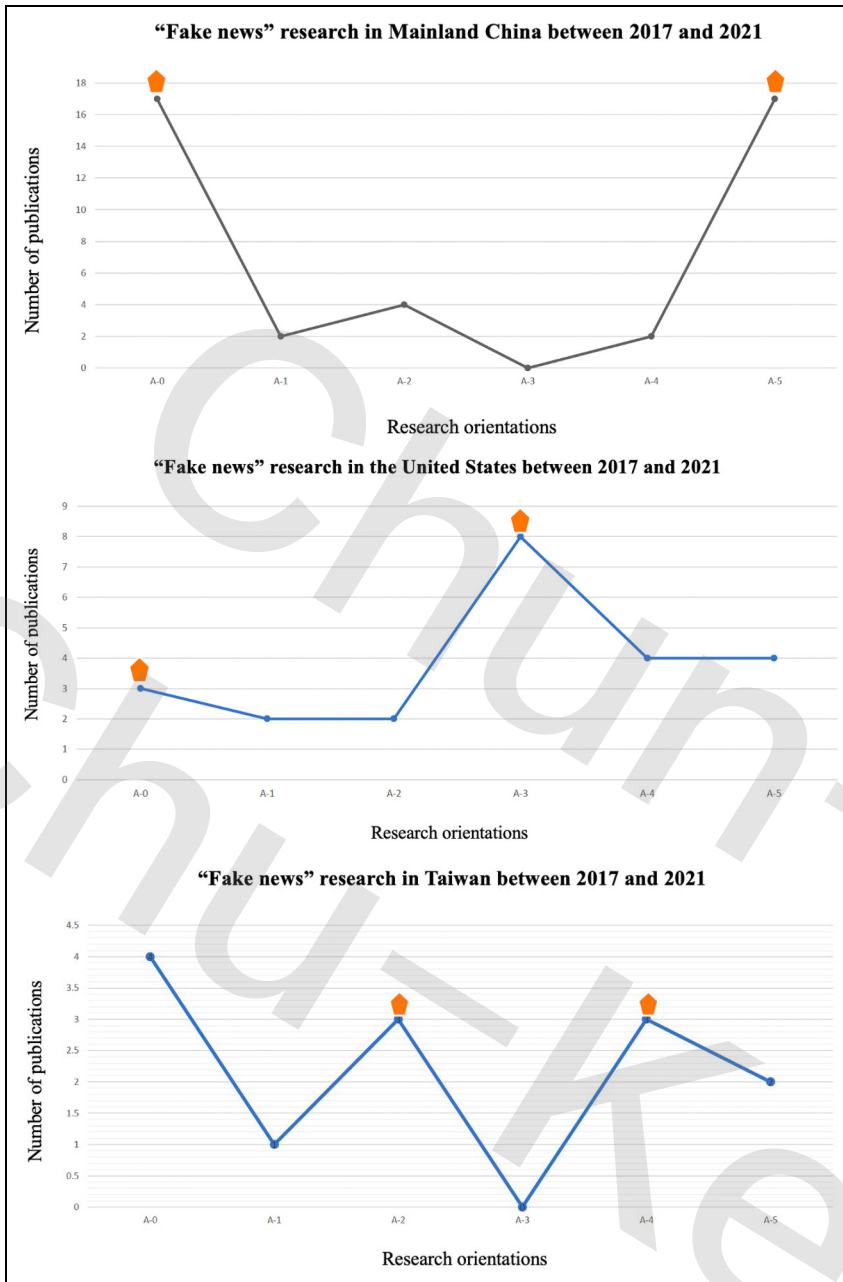


Figure 1. Research orientations in three regions between 2017 and 2021.

Note. A-0 represents (often non-empirical) review articles; A-1, A-2, and A-3 are articles exploring conceptualization, the psychological mechanism underlying fake news consumption, and influences of fake news on individuals' attitudes and behaviors, respectively; A-4 focuses on news fact-checking; A-5 predominantly utilizes cases analyses and elaborates on trending fake news cases related to social issues.

(Strömbäck, 2008). This mediated communication often goes hand in hand with bias in media coverage, characterized by the lack of accuracy and realism, the blur between fact and opinion, and slant (Hackett, 1984; McQuail, 1996). In political communication, media bias is reflected in partisanship and biased reporting that favors one political actor/position/stance over another. More specifically, the bias in political media coverage may manifest itself in the visibility of political actors/parties (i.e. visibility bias), evaluation of political actors/parties (i.e. tonality bias), and the association between political actors/parties and certain issues (i.e. agenda bias) (Eberl et al., 2017). Nevertheless, it's noteworthy that misinformation in other social domains (e.g. health-related misinformation) is found to be created mostly by individuals with no official or institutional affiliations (Wang et al., 2019), which shifts a wide bunch of researchers' attention from the production of misinformation to individuals' reception and formation of misinformation at a micro-level.

Misinformation research at a micro-level emphasizes people's false and unsubstantiated belief formation process. According to Nyhan and Reifler (2010), misinformation refers to a situation where "people's beliefs about factual matters are not supported by clear evidence and expert opinion" (p. 305). This definition is open to discussion on the formation of such beliefs from psychological perspectives. Accordingly, a general pathway is that individuals who are exposed to misinformation will evaluate or assess the information first, during which rational thinking and heuristics may intertwine, and emotions enter into the decision process at the same time; consequently, attitudinal changes and decision bias will occur, and the behavioral decision will be made (see Figure 2) (Chu-Ke, 2022). Amazeen (2023) in a recent paper proposed a misinformation recognition and response model and treated the misinformation reception process in much detail. For instance, problem identification and issue motivation are important prerequisites of the activation of misinformation recognition, after which individuals may adopt different cognitive coping strategies and display ensuing responses.

Psychological decision theories on the role of heuristics may shed light on the individuals' reception and formation of misinformation. As behavioral decision theory suggests, humans use heuristics to simplify their decision-making under uncertainty. Three heuristics that are often used to assess probabilities and values of uncertain information include representativeness heuristics (i.e. assessing the probability by considering the object or event to what extent is representative of existing knowledge or stereotypes), availability heuristics (i.e. assessing the probability by considering the instances or occurrences of an object or event can be brought to mind), and adjustment and anchoring heuristics (i.e. assessing the probabilities biased toward the initial values that are either provided by others or produced by insufficient computation) (Slovic et al., 1977; Tversky & Kahneman, 1974). Despite being very economical and beneficial sometimes (Bilek et al., 2018), these judgmental heuristics can lead to serious cognitive biases and facilitate the reception

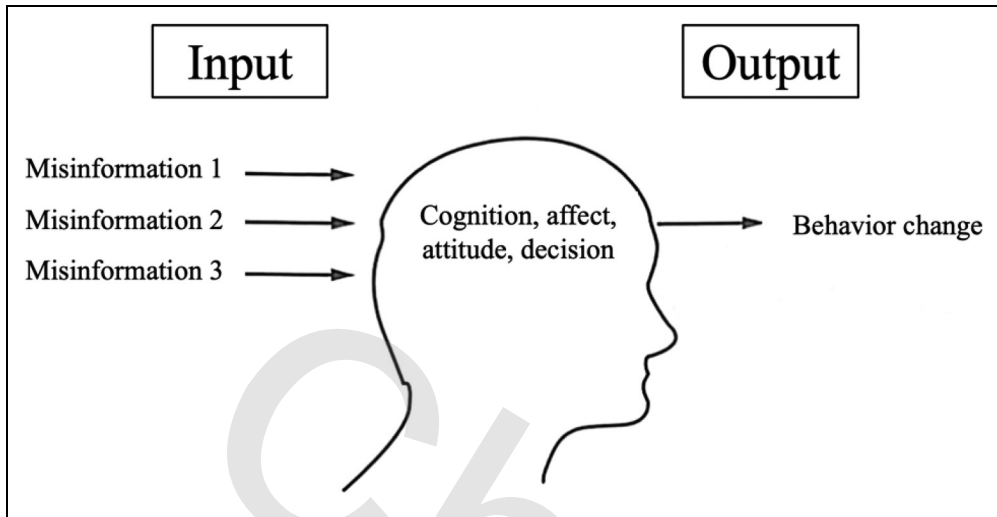


Figure 2. Misinformation and individual information processing.

of political and health misinformation among citizens (Dancey & Sheagley, 2013; Hwang & Jeong, 2023).

Emotions may produce bias when individuals identify misinformation and motivate individuals to spread misinformation. According to a review of research on emotion and decision-making by Lerner et al. (2015), emotions, including incidental and integral ones, can “powerfully, predictably, and pervasively influence decision-making” (p. 802) such that they shape the content of thought, depth of thought, goal activation, etc. Anger, specifically, is associated with approach motivation (Harmon-Jones & Sigelman, 2001) and, together with counterarguments, goes against typically beneficial advocacies and drives unwise decisions (Brehm, 1966; Rosenberg & Siegel, 2018). A systematic review explored the association between emotions and misinformation, and it concluded that moderate levels of stress and short-term distress can decrease susceptibility to misinformation, while anger may increase susceptibility (Sharma et al., 2023).

Misinformation and vulnerable groups

Empirical research has suggested that senior citizens are susceptible to misinformation. Loos and Nijenhuis (2020) in their study on political fake news articles in the form of advertisements on Facebook revealed that older adults, compared with younger adults, are more likely to be exposed to political fake news articles. Yousuf et al. (2021) also found that senior citizens in the Netherlands generally believed misinformation related to COVID-19 vaccines. Some reasons behind the positive relationship between age and susceptibility to misinformation have been summarized by Brashier and Schacter (2020): they pointed out that senior citizens may have greater

misplaced trust, more difficulty in detecting misinformation, and less emphasis on information accuracy when consuming information, which facilitates the reception of misinformation among senior citizens. With a cross-national survey, Jo et al. (2022) found that senior citizens are less likely to check on suspicious content. These may also be related to the digital literacy of senior citizens. Some of the senior citizens may be newcomers to the Internet and have insufficient ability to detect sponsored or manipulated content. Besides, the reception of misinformation among senior citizens can be attributed to both media exposure and social influence. For instance, family members and friends who are exposed to misinformation can actively relay it to senior citizens (Chia et al., 2023).

In addition to senior citizens, it is important to recognize that adolescents are also at risk of falling victim to misinformation. Baumel et al. (2021) suggest that adolescents may be frequently exposed to misinformation and “at a substantial risk of misinformation” because of their deep engagement in social media (p. 1021). Individuals with certain political stances can also be vulnerable to misinformation. For instance, conservative political orientation was found to be associated with higher acceptance of COVID-19 misinformation compared with liberal political orientation (Xiao, 2022). When communicating information about stigmatized health topics, those stigma-aware individuals are found to display information avoidance, which encourages the belief in misinformation (Dong et al., 2024). Besides, misinformation acceptance may depend on various information-processing styles and states. Those who experienced message fatigue and information overload were found to have a greater acceptance of COVID-19 misinformation via increased information avoidance (Hwang et al., 2022; Tandoc & Kim, 2022). For those who engaged in heuristic processing, the effect of COVID-19 misinformation exposure on COVID-19 misinformation acceptance was greater (Hwang & Jeong, 2023).

A case analysis of misinformation: pesticides into the air via aerial spraying

The emission of pesticides into the air via aerial spraying used to play an important role in preventing the spread of diseases and insect attacks in many parts of the world. In May 2016, a piece of misinformation about “the emission of pesticides into the air via aerial spraying for the control of fall webworms” started to circulate in group chats on WeChat, the most popular social media in China. The misinformation informed the citizens that the local Forestry and Park Bureau planned to use aircraft to spray pesticides to prevent and control the fall webworms and suggested citizens avoid outdoor activities during the spraying period. The misinformation spread from the east of the Chinese mainland to the whole Chinese mainland and Taiwan area at a very fast speed, which drew the attention of mainstream media and local government. The media coverage followed very soon after the virality of misinformation and confirmed that the local government would not release pesticides into the air via aerial spraying. Every year social media in China

witnesses the viral spread of similar misinformation on “the emission of pesticides into the air via aerial spraying” and the correction to it as follows.

First, it’s noteworthy that this misinformation heavily relies on authority heuristics. In the misinformation, either the Forestry and Park Bureau or the Department of Agriculture as the unfounded information source attracted public attention and significantly increased the information’s credibility via heuristics. Second, as time went by, the content of misinformation evolved and was inextricably interwoven with trending social events. In 2019, the misinformation on “the emission of pesticides into the air via aerial spraying for the control of fall webworms” included more details than in 2016, such as associating the outbreaks of fall webworms with swine fever. Since 2020, the misinformation combined pesticide spraying rumors with COVID-19 information.

Senior citizens are found to play an important role in spreading and accepting this misinformation. In eight counties and cities in Taiwan area where the misinformation about “the emission of pesticides into the air via aerial spraying” gained the greatest popularity, senior citizens occupy over 14% of the local population. As pesticide misinformation is tightly associated with peoples’ health and environment, it may draw much attention from senior citizens and raise their rumor anxiety, which consequently leads to a greater likelihood of spreading misinformation (He et al., 2019).

Misinformation and Literacies in the Age of Generative Ai

Generative AI and misinformation

The debate about AI flooding the public sphere can be traced back to social bots and deepfakes. Social bots, “a computer algorithm that automatically produces content and interacts with humans on social media, trying to emulate and possibly alter their behavior” (Ferrara et al., 2016, p. 96), have been criticized for posing great risks to the digital environment as well as our society. They amplify interactions with low-credibility content and help circulate it to influential users so that low-credibility content can go viral (Shao et al., 2018; Shu et al., 2020). The social networks consisting of bots help diffuse misinformation and fake news (Miró-Llinares and Aguerri, 2023). Deepfake is another debatable generative AI technique that has been exploited for creating fake media. In a study on the discussion about deepfakes in Reddit communities, Gamage et al. (2022) revealed that niche social media platforms such as Reddit provide an incubator for deepfake generators/creators, and especially for expert users to support the creation of customized pornography.

Generative AI is defined as “AI based on advanced neural networks with the ability to produce highly realistic synthetic text, images, video and audio (including fictional stories, poems, and programming code)” (Galaz et al., 2023, p. 5). They are characterized by increased accessibility

(i.e. being easily available), sophistication (i.e. the capacity to create longer pieces of synthetic text, video, and voice with specific prompts), and capabilities for persuasion (i.e. the capacity to elaborate the content persuasively). With these characteristics, generative AI can now help create and spread false, seemingly credible information quickly and effortlessly (Zhou et al., 2023). For instance, ChatGPT, a chatbot developed by OpenAI and launched on 30 November 2022, has been found to produce convincing text that comprises conspiracy theories and misleading narratives (Hsu & Thompson, 2023) and manufacture fake news articles that claim trusted news organizations and journalists as the source (Moran, 2023). Meanwhile, due to the easy access with minimal programmatic knowledge, human users who deploy generative AI originally as assistive technologies can unintentionally become creators and spreaders of misinformation.

Given the capabilities of generative AI and its potential impact on misinformation, it's necessary to consider the following three dimensions for the development of ethical AI and ethical uses of AI in counteracting misinformation, the first two of which concern AI practitioners and legislation, and the rest at the user end will be discussed with a focus on its implications for the literacy program and communication studies.

Making ethical AI: regulations From inside and outside

Ethical AI for combating misinformation necessitates efforts from AI practitioners, especially those who design and develop models, platforms, and APIs. First, transparent AI development practices should be encouraged. It is critical that AI practitioners, platforms, and organizations come together to create a culture that nurtures values of collaboration and openness. Specifically, AI algorithms, data, and methodologies should be openly accessible while maintaining ethical bounds to promote transparency and collective understanding. Moreover, it's important to deploy interpretable models and explainable user interfaces. With enhanced interpretability and explainability of generative AI, users can better understand how generative AI arrives at specific decisions and predictions and detect potential falsehoods. Second, there's a need for user guidance and monitoring of misuse (Zhou et al., 2023). Despite that the current ChatGPT notifies the users that "ChatGPT can make mistakes. Consider checking important information" at the bottom of its interface, such a warning cue may not be salient enough to raise users' awareness of the potentially flawed information. Cues appearing right after copy-and-paste action may also assist in reminding the users to do fact-checking if necessary. To help users verify information, it can provide useful links to fact-checking websites, accredited news organizations, and relevant governmental agencies.

Ethical AI for combating misinformation also entails the intervention of laws. On December 8, 2023, European Parliament and Council negotiators successfully reached a provisional agreement on the Artificial Intelligence Act (i.e. the "AI Act"). This is the inaugural comprehensive law on AI enacted by a major regulatory body, aiming to "ensure that fundamental rights, democracy, the rule

of law and environmental sustainability are protected from high risk AI, while boosting innovation and making Europe a leader in the field” (Artificial Intelligence Act, 2023). While the AI Act does not directly target misinformation, its provisions regarding prohibited AI practices and requirements for high-risk AI can contribute to the development and deployment of more reliable AI systems for combatting misinformation. For instance, in line with the interpretability and explainability, we previously mentioned, the AI act emphasizes the rights of individuals who may be subject to decisions produced by high-risk AI or adversely impacting their health, safety, etc., to request from the AI deployer clear and meaningful explanations about the role of the AI system plays in the decision-making process. Some general-purpose models that are often deployed to produce misinformation and disinformation, such as generative models that create images, code, and video, have also been regulated accordingly (Gibney, 2024). Enforcement mechanisms and penalties for non-compliance to provisions regarding data governance, transparency obligations, monitoring, and surveillance are established.

Promoting ethical use of AI: constructing AI literacies

Faced with the changing media landscape and pervasive misinformation, a variety of literacy has been conceptualized and investigated. For instance, Khan and Idris (2019) contribute to the conceptualization of information literacy by specifying information-seeking, sharing, and verification skills. Specifically, information verification skills refer to the capacity to verify online information using an online tool and to use a search engine. Carmi et al. (2020) argued about the significance of data literacies in an increasingly datafied society. They emphasized what data literacies encompass in addition to the previous literacies, e.g. networked literacies, critical thinking about the online ecosystem, and literacies assisting in active citizenship, to help citizens identify and combat the misinformation crisis. In light of the development of generative AI, Long and Magerko (2020) further devised a conceptual framework for AI literacy composed of five overarching themes: *What is AI?*; *What can AI do?*; *How does AI work?*; *How should AI be used?*; and *How do people perceive AI?*. They also outlined competencies in each theme, which includes the aforementioned data literacy as a competency needed for the understanding of AI’s working mechanism.

Here, we’d like to engage in a dialogue with the conceptualization of literacy by Livingstone (2004) and explore a few points that may offer directions for future literacy education and research regarding AI-fueled misinformation. According to Livingstone (2004), literacy is rooted in the historically and culturally conditioned relationship among three processes:

The symbolic and material representation of knowledge, values, and cultures. AI literacy can be considered the ability to understand and critically evaluate knowledge, values, and cultures in which AI systems operate and their impacts on AI-generated content. Specifically, AI literacy requires an understanding of specific cultural, social, and political contexts in which AI systems

operate, and the potentially prejudiced decisions based on demographic features (e.g. race, sex, etc.) (e.g. Ntoutsis et al., 2020). This can help individuals identify how AI-fueled misinformation may exploit or perpetuate societal biases and narratives. This perspective in education programs can incorporate this understanding into curricula to encourage critical reflection on the societal implications of AI. Social science research may delve into AI as a cultural artifact in different tasks and its socio-psychological implications for misinformation detection and correction.

The dissemination of interpretative skills and abilities among the population. As Livingstone (2004) emphasized, literacy should be conceived as “a coproduction of the interactive engagement between technology and user” (p. 11). The skilled and strategic interpretation of AI-generated content in countering misinformation should include hands-on experience with AI tools and platforms and exploration of the underlying principles and mechanisms of AI systems. During the information-seeking phase, users can utilize source verification prompts, fact-checking suggestions, and debunking resources. When using AI-generated content, individuals should realize that this content is not inherently original. When available, individuals should acknowledge the source of the content. The education program should target vulnerable groups (e.g. senior citizens and adolescents), using accessible, easy-to-understand, and tailored materials. Research should be dedicated to the effectiveness of cues and prompt suggestions in encouraging proper use of AI-generated content and verification.

The institutional use and management of the power that grants skills and knowledge to people. Literacy is often institutionalized through educational systems, libraries, and other organizations that disseminate knowledge and skills. Similarly, AI literacy can be institutionalized through the cooperation between aforementioned AI practitioners, deployers, and legislation that promotes responsible use of AI technologies. From the users’ end, users should develop an awareness of and skills in the feedback mechanisms of the AI systems and report potentially misleading or false information encountered through AI use. This feedback can help improve the system’s accuracy and effectiveness in countering misinformation over time and empower users to facilitate institutional management of the AI power.

Conclusion

Misinformation is a social practice and challenge that demands our substantial attention. With the proliferation of falsehoods, the erosion of trust and deepening societal divisions are inevitable and undesired consequences. Therefore, misinformation research is not only a scholarly investigation but also a social inquiry. In this essay, we briefly reviewed theoretical advancements and empirical evidence regarding misinformation. We suggest that the definition and study of misinformation can vary across disciplines and regions due to differences in research orientations, methodologies, and cultural contexts. Misinformation research at the macro level primarily touches upon media bias

and institutional inadequacy that facilitate the information crisis, whereas micro-level studies mainly examine the (conditioned) process of individuals' reception and formation of misinformation. Populations that are specifically vulnerable to misinformation and a case of misinformation about "pesticides into the air via aerial spraying" are discussed.

In the age of generative AI, misinformation and AI literacy targeting misinformation should be reconsidered. It's imperative that we embark on ethical AI development and ethical AI use to combat the rising tide of misinformation. The call to action should and has been increasingly made for AI practitioners and lawmakers alike to embrace ethical guidelines and regulations. Furthermore, the empowerment of users through the implementation of AI literacy programs and the study of communication is of necessity. We suggest that at least three aspects of AI literacy should be incorporated into the current literacy education and research: (a) The understanding and critical evaluation of knowledge, values, and cultures within which AI systems function, and their implications on the content they generate; (b) The strategic interpretation and proper use of AI-generated content; (c) The utilization of feedback mechanisms to promote institutional management of the AI power.

Acknowledgment

All authors have agreed to the submission and that the article is not currently being considered for publication by any other print or electronic journal.

Declaration of Conflicting Interests

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

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Yujie Dong  <https://orcid.org/0000-0001-8146-6571>

References

- 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>
- Amazeen, M. A. (2023). The misinformation recognition and response model: An emerging theoretical framework for investigating antecedents to and consequences of misinformation recognition. *Human Communication Research*, hqad040, <https://doi.org/10.1093/hcr/hqad040>
- Artificial Intelligence Act: Deal on comprehensive rules for trustworthy AI. (2023). European Parliament Press Room. <https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai>.

- Aufderheide, P. (1997). Media Literacy: From a Report of the National Leadership Conference on Media Literacy. In *Media Literacy Around the World*. Routledge.
- Baumel, N. M., Spatharakis, J. K., Karitsiotis, S. T., & Sellas, E. I. (2021). Dissemination of mask effectiveness misinformation using TikTok as a medium. *Journal of Adolescent Health, 68*(5), 1021–1022. <https://doi.org/10.1016/j.jadohealth.2021.01.029>
- Bílek, J., Nedoma, J., & Jirásek, M. (2018). Representativeness heuristics: A literature review of its impacts on the quality of decision-making. *Scientific Papers of the University of Pardubice. Series D, Faculty of Economics and Administration, 43*: 29–38.
- Brashier, N. M., & Schacter, D. L. (2020). Aging in an era of fake news. *Current Directions in Psychological Science, 29*(3), 316–323. <https://doi.org/10.1177/0963721420915872>
- Brehm, J. W. (1966). *A theory of psychological reactance*. Academic Press.
- Carmi, E., Yates, S. J., & Lockley, E., & Pawluczuk, A. (2020). Data citizenship: Rethinking data literacy in the age of disinformation, misinformation, and malinformation. *Internet Policy Review, 9*(2), 1–22. <https://doi.org/10.14763/2020.2.1481>
- Chen, K., Jin, Y., & Shao, A. (2022). Science factionalism: How group identity language affects public engagement with misinformation and debunking narratives on a popular Q&A platform in China. *Social Media + Society, 8*(1), 20563051221077019. <https://doi.org/10.1177/20563051221077019>
- Chia, S. C., Lu, F., & Sun, Y. (2023). Tracking the influence of misinformation on elderly people's perceptions and intention to accept COVID-19 vaccines. *Health Communication, 38*(5), 855–865. <https://doi.org/10.1080/10410236.2021.1980251>
- Christ, W. G., & Potter, W. J. (1998). Media literacy, media education, and the academy. *Journal of Communication, 48*(1), 5–15. <https://doi.org/10.1111/j.1460-2466.1998.tb02733.x>
- Chu-Ke, C. (2022). Psychological processes and behavioral responses of misinformation. In L. Kuo, & Y. Chen (Eds.), *Emerging media in communication and media literacy* (1st ed, pp. 65–85). Yeh Yeh Book Gallery.
- Dai, Y., Yu, W., & Shen, F. (2021). The effects of message order and debiasing information in misinformation correction. *International Journal of Communication, 15*: 1039–1059.
- Dancey, L., & Sheagley, G. (2013). Heuristics behaving badly: Party cues and voter knowledge. *American Journal of Political Science, 57*(2), 312–325. <https://doi.org/10.1111/j.1540-5907.2012.00621.x>
- Dong, Y., Zhang, L., Lam, C., & Huang, Z. (2024). Counteracting sexual and reproductive health misperceptions: Investigating the roles of stigma, misinformation exposure, and information overload. *Patient Education and Counseling, 120*, 108098. <https://doi.org/10.1016/j.pec.2023.108098>
- Eberl, J.M., Boomgaarden, H. G., & Wagner, M. (2017). One bias fits all? Three types of media bias and their effects on party preferences. *Communication Research, 44*(8), 1125–1148. <https://doi.org/10.1177/0093650215614364>
- Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM, 59*(7), 96–104. <https://doi.org/10.1145/2818717>
- Galaz, V., Metzler, H., Daume, S., Olsson, A., Lindström, B., & Marklund, A. (2023). *AI could create a perfect storm of climate misinformation* (arXiv:2306.12807). arXiv. <https://doi.org/10.48550/arXiv.2306.12807>
- Gamage, D., Ghasiya, P., Bonagiri, V. K., Whiting, M. E., & Sasahara, K. (2022). Are deepfakes concerning? Analyzing conversations of deepfakes on reddit and exploring societal implications. *CHI Conference on Human Factors in Computing Systems*, 1–19. <https://doi.org/10.1145/3491102.3517446>

- Gibney, E. (2024). What the EU's tough AI law means for research and ChatGPT. *Nature* 626: 938–939. <https://doi.org/10.1038/d41586-024-00497-8>.
- Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, 363(6425), 374–378. <https://doi.org/10.1126/science.aau2706>
- Ha, L., Andreu Perez, L., & Ray, R. (2021). Mapping recent development in scholarship on fake news and misinformation, 2008 to 2017: Disciplinary contribution, topics, and impact. *American Behavioral Scientist*, 65(2), 290–315. <https://doi.org/10.1177/0002764219869402>
- Hackett, R. A. (1984). Decline of a paradigm? Bias and objectivity in news media studies. *Critical Studies in Mass Communication*, 1(3), 229–259. <https://doi.org/10.1080/15295038409360036>
- Harmon-Jones, E., & Sigelman, J. (2001). State anger and prefrontal brain activity: Evidence that insult-related relative left-prefrontal activation is associated with experienced anger and aggression. *Journal of Personality and Social Psychology*, 80(5), 797–803. <https://psycnet.apa.org/record/2001-17232-008>
<https://doi.org/10.1037/0022-3514.80.5.797>
- He, L., Yang, H., & Xiong, X., & Lai, K. (2019). Online rumor transmission among younger and older adults. *SAGE Open*, 9(3): 1–9. <https://doi.org/10.1177/2158244019876273>
- Howard, P. N., Kollanyi, B., Bradshaw, S., & Neudert, L.M. (2018). Social media, news and political information during the US election: Was polarizing content concentrated in swing states? *arXiv Preprint arXiv:1802.03573*.
- Hsu, T., & Thompson, S. A. (2023). Disinformation Researchers Raise Alarms About A.I. Chatbots. *The New York Times*. <https://www.nytimes.com/2023/02/08/technology/ai-chatbots-disinformation.html>.
- Hwang, Y., & Jeong, S.H. (2023). Misinformation exposure and acceptance: The role of information seeking and processing. *Health Communication*, 38(3), 585–593. <https://doi.org/10.1080/10410236.2021.1964187>
- Hwang, Y., So, J., & Jeong, S.H. (2023). Does COVID-19 message fatigue lead to misinformation acceptance? An extension of the risk information seeking and processing model. *Health Communication*, 38(12), 2742–2749. <https://doi.org/10.1080/10410236.2022.2111636>
- Jerit, J., & Zhao, Y. (2020). Political misinformation. *Annual Review of Political Science*, 23(1), 77–94. <https://doi.org/10.1146/annurev-polisci-050718-032814>
- Jo, H., Yang, F., & Yan, Q. (2022). Spreaders vs victims: The nuanced relationship between age and misinformation via FoMO and digital literacy in different cultures. *New Media & Society*, 14614448221130476. <https://doi.org/10.1177/14614448221130476>
- Khan, M. L., & Idris, I. K. (2019). Recognise misinformation and verify before sharing: A reasoned action and information literacy perspective. *Behaviour & Information Technology*, 38(12), 1194–1212. <https://doi.org/10.1080/0144929X.2019.1578828>
- Krishna, A., & Thompson, T. L. (2021). Misinformation about health: A review of health communication and misinformation scholarship. *American Behavioral Scientist*, 65(2), 316–332. <https://doi.org/10.1177/0002764219878223>
- Lee, N. M. (2018). Fake news, phishing, and fraud: A call for research on digital media literacy education beyond the classroom. *Communication Education*, 67(4), 460–466. <https://doi.org/10.1080/03634523.2018.1503313>
- Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual Review of Psychology*, 66(1), 799–823. <https://doi.org/10.1146/annurev-psych-010213-115043>

- Lewandowsky, S., & van der Linden, S. (2021). Countering misinformation and fake news through inoculation and prebunking. *European Review of Social Psychology*, 32(2), 348–384. <https://doi.org/10.1080/10463283.2021.1876983>
- Li, J. (2020). Toward a research agenda on political misinformation and corrective information. *Political Communication*, 37(1), 125–135. <https://doi.org/10.1080/10584609.2020.1716499>
- Livingstone, S. (2004). Media literacy and the challenge of new information and communication technologies. *The Communication Review*, 7(1), 3–14. <https://doi.org/10.1080/10714420490280152>
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–16. <https://doi.org/10.1145/3313831.3376727>
- Loos, E., & Nijenhuis, J. (2020). Consuming fake news: A matter of age? The perception of political fake news stories in Facebook ads. In Q. Gao, & J. Zhou (Eds.), *Human aspects of IT for the aged population. Technology and society* (pp. 69–88). Springer International Publishing. https://doi.org/10.1007/978-3-030-50232-4_6
- Malhotra, P., & Pearce, K. (2022). Facing falsehoods: Strategies for polite misinformation correction. *International Journal of Communication*, 16: 2303–2324.
- McQuail, D. (1996). Media performance, mass communication and the public interest: The future of a research tradition. In C. Honneger (Ed.), *Gesellschaften im umbau* (pp. 235–244). Seismo. <https://dare.uva.nl/search?identifier=7a618c36-4d50-4732-98c1-5413662865fa>
- Miró-Llinares, F., & Aguerri, J. C. (2023). Misinformation about fake news: A systematic critical review of empirical studies on the phenomenon and its status as a ‘threat’. *European Journal of Criminology*, 20(1), 356–374. <https://doi.org/10.1177/1477370821994059>
- Moran, C. (2023). ChatGPT is making up fake Guardian articles. Here’s how we’re responding. *The Guardian*. <https://www.theguardian.com/commentisfree/2023/apr/06/ai-chatgpt-guardian-technology-risks-fake-article>.
- Ntoutsis, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejd, W., Vidal, M. E., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., Kompatsiaris, I., Kinder-Kurlanda, K., Wagner, C., Karimi, F., Fernandez, M., Alani, H., Berendt, B., Kruegel, T., Heinze, C., & Staab, S. (2020). Bias in data-driven artificial intelligence systems—an introductory survey. *WIREs Data Mining and Knowledge Discovery*, 10(3), e1356. <https://doi.org/10.1002/widm.1356>
- Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32(2), 303–330. <https://doi.org/10.1007/s11109-010-9112-2>
- Roozenbeek, J., Schneider, C. R., Dryhurst, S., Kerr, J., Freeman, A. L. J., Recchia, G., van der Bles, A. M., & van der Linden, S. (2020). Susceptibility to misinformation about COVID-19 around the world. *Royal Society Open Science*, 7(10), 201199. <https://doi.org/10.1098/rsos.201199>
- Rosenberg, B. D., & Siegel, J. T. (2018). A 50-year review of psychological reactance theory: Do not read this article. *Motivation Science*, 4(4), 281–300. <https://doi.org/10.1037/mot0000091>
- Scheufele, D. A., & Krause, N. M. (2019). Science audiences, misinformation, and fake news. *Proceedings of the National Academy of Sciences*, 116(16), 7662–7669. <https://doi.org/10.1073/pnas.1805871115>
- 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), Article 1. <https://doi.org/10.1038/s41467-017-02088-w>

- Sharma, P. R., Wade, K. A., & Jobson, L. (2023). A systematic review of the relationship between emotion and susceptibility to misinformation. *Memory*, 31(1), 1–21. <https://doi.org/10.1080/09658211.2022.2120623>
- Shu, K., Bhattacharjee, A., Alatawi, F., Nazer, T. H., Ding, K., Karami, M., & Liu, H. (2020). Combating disinformation in a social media age. *WIREs Data Mining and Knowledge Discovery*, 10(6), e1385. <https://doi.org/10.1002/widm.1385>
- Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. *Annual Review of Psychology*, 28(1), 1–39. <https://doi.org/10.1146/annurev.ps.28.020177.000245>
- Strömbäck, J. (2008). Four phases of mediatization: An analysis of the mediatization of politics. *The International Journal of Press/Politics*, 13(3), 228–246. <https://doi.org/10.1177/1940161208319097>
- Swire-Thompson, B., & Lazer, D. (2019). Public health and online misinformation: Challenges and recommendations. *Annual Review of Public Health*, 41, 433–451. <https://doi.org/10.1146/annurev-publhealth-040119-094127>
- Tandoc, E. C., & Kim, H. K. (2022). Avoiding real news, believing in fake news? Investigating pathways from information overload to misbelief. *Journalism* 24(6): 1174–1192. <https://doi.org/10.1177/14648849221090744>.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
- Wang, Y., McKee, M., Torbica, A., & Stuckler, D. (2019). Systematic literature review on the spread of health-related misinformation on social media. *Social Science & Medicine*, 240, 112552. <https://doi.org/10.1016/j.socscimed.2019.112552>
- Xiao, X. (2022). Not doomed: Examining the path from misinformation exposure to verification and correction in the context of COVID-19 pandemic. *Telematics and Informatics*, 74, 101890. <https://doi.org/10.1016/j.tele.2022.101890>
- Yousuf, H., van der Linden, S., Bredius, L., Essen, G. A. V., Sweep, G., Preminger, Z., van Gorp, E., Scherder, E., Narula, J., & Hofstra, L. (2021). A media intervention applying debunking versus non-debunking content to combat vaccine misinformation in elderly in the Netherlands: A digital randomised trial. *eClinicalMedicine*, 35, <https://doi.org/10.1016/j.eclinm.2021.100881>
- Zhou, J., Zhang, Y., Luo, Q., Parker, A. G., & De Choudhury, M. (2023). Synthetic lies: Understanding AI-generated misinformation and evaluating algorithmic and human solutions. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–20. <https://doi.org/10.1145/3544548.3581318>