Article

In Punishment We Trust: Analyzing Characteristics and Credibility of Rumor-Debunking Messages on Chinese Social Media

Bin Chen[1] and Shuning Lu[2] iD

Abstract

Research on fact-checking journalism has predominantly centered on Western countries, often overlooking how distinct political media systems in non-Western countries might influence its practices and effectiveness. This article addresses this gap by focusing on government-led fact-checking in China, referred to as "rumor-debunking," and we argue that compared to the Western mode of fact-checking, rumor debunking differs significantly with its unique motives, targets, and outcomes within the context of an authoritarian regime. Employing content analysis (Study 1) and a conjoint experiment (Study 2), we empirically examined the characteristics and credibility perceptions of rumor-debunking messages on Chinese social media Weibo during the COVID-19 pandemic. The findings of Study 1 reveal that rumor-debunking efforts in China were primarily driven by government-affiliated accounts and mainly targeting ordinary citizens. Study 2 shows that message specificity mattered in that detailing the rumors being debunked, providing evidence for the debunking, and including punitive actions against rumormongers significantly enhanced perceived credibility among the Chinese public. Additionally, we found messages that received higher social endorsement—reflected by a greater number of reposts, comments, and likes—boosted credibility. However, the presence of "recommended comments," which indicate that the account holders curated the user comments under their debunking messages, did not affect credibility. The social and political implications of rumor-debunking in authoritarian regimes were also discussed.

[1]Journalism and Media Studies Centre, The University of Hong Kong, Hong Kong
[2]Department of Communication, North Dakota State University, Fargo, ND, USA

Corresponding Author:
Shuning Lu, Department of Communication, North Dakota State University, PO Box 6050, Fargo, ND 58108-6050, USA.
Email: shuning.lu@ndsu.edu

Misinformation has become a major challenge to the global society. Although the overall prevalence of misinformation is found to be relatively low (e.g., Acerbi et al. 2022; Grinberg et al. 2019), its impact can be critical. Especially during times of global crisis such as the COVID-19 pandemic, even limited instances of misinformation can mask preventive health behaviors and promote erroneous practices (Lee et al. 2020). Fact-checking by both professional news organizations and independent fact-checkers is regarded as an indispensable force to combat misinformation. Yet, the emerging global trend of government-led fact-checking raises concerns about its potential damage to democracy and public access to truthful information.[1]

Although a growing body of literature has started to explore fact-checking in non-Western countries (e.g., Mare and Munoriyarwa 2022; Porter and Wood 2021; Schuldt 2021), we still lack the understanding of how fact-checking work is conceived and realized in authoritarian regimes where free media is lacking. As Madrid-Morales and Wasserman (2022) noted, the understanding of misinformation or disinformation is contextually driven and shaped by the power relations in the national information environment. Who are conducting fact-checks, whose claims are being checked, and how fact-checking is performed may appear different across Western, Educated, Industrialized, Rich and Democratic (WEIRD) and non-WEIRD countries. It is also worth noticing that fact-checking work in non-WEIRD countries often extends beyond verifying the accuracy of information and promoting factual reasoning and serves to maintain the legitimacy of the ruling government through undermining dissenting voices (Fang 2022).

In China, falsehoods or unsubstantiated information are termed as *rumors* (*yaoyan*, 谣言, literally means hearsay). Rumors could be destructive to authoritarian governments, because they constitute "an alternative form of media that directly competes with official information and mainstream media" (Huang 2017: 283). The traditional model of authoritarian information management centers on shaping public perception for political ends by distorting evidence (Roberts 2018). But this model has become increasingly infeasible in the online environment where rich information and diverse opinions, rumors included, are widely distributed. In response, the Chinese government has proactively adopted strategic information management by disseminating practical and clickbait information to compete for attention (Chang 2021; Lu and Pan 2021). Rumor-debunking is one such effort that delivers practical information to counter falsehoods while extending the reach of government information. Note that the information addressed by government-led rumor debunking efforts may not necessarily be factually false, but it may be considered undesirable by the government.

The article sought to examine rumor-debunking on Chinese social media with a focus on the COVID-19 pandemic. Since its onset in China in December 2019, the pandemic has been rife with numerous rumors distributed on Chinese social media. These rumors not only fueled the uncertainty and anxiety among the public but also challenged the authorities' ability to govern effectively. It is not surprising that both government entities and official media made considerable efforts to debunk rumors during the pandemic (Fang 2022). To provide an updated account of the features and effects of rumor-debunking practices on Chinese social media, we conducted two studies. Study 1 involved a content-analysis of the characteristics of rumor-debunking messages on Weibo, one of the most popular social media platforms in China. Study 2 used a conjoint experiment to test how the Chinese public evaluate the credibility of rumor-debunking messages with features identified in Study 1.

Study 1: Exploring Features of Rumor-Debunking Messages

From Fact-Checking to Rumor-Debunking

Fact-checking journalism has become increasingly popular in the globe amid critical events filled with disinformation campaigns and misinformation (Graves 2018). Fact-checking encompasses journalistic practices that prioritize the validation of accuracy and truth in claims made by politicians, journalists, or other public figures (Graves 2016). However, fact-checking can vary across different political and/or media systems. For example, fact-checkers in different countries may have varying understandings of what should be checked. In democratic countries, the claims selected for fact-checking are often those made by political figures. Conversely, in authoritarian countries, fact-checking more frequently concerns claims that may tarnish the government's image, which is often called "rumor-debunking." "Rumor-debunking" is a direct translation of the Chinese term *piyao (辟谣)*, widely used by the Chinese government and official media to denote the act of fact-checking rumors. Although its English variations such as "rumor-refutation" or "rumor-correction" are adopted in the literature (Chen et al. 2021; Song et al. 2021), we contend that "rumor-debunking" is particularly apt for capturing the dynamic nature of the government-led antirumor practices.[2] This term has been recognized and adopted in previous work in the field (e.g., Fang 2022; Gao et al. 2022).

It is worth noting that rumor-debunking is inherently and contextually different from fact-checking or misinformation correction. First, we need to distinguish between misinformation and rumors. Some research uses the two terms interchangeably (e.g., Fang 2022; Xiang and Neo 2024); however, we chose to adopt "*rumors*" rather than "misinformation" to emphasize the political purpose and ramifications of the government-led rumor-debunking movement in China. Although both misinformation and rumor refer to inaccurate or misleading information, rumors convey a social psychological significance. Researchers have argued that rumors oftentimes reflect public fear, anger, or anxiety especially amid social crises (Rosnow 1988), thereby creating social instability and public distrust in government. Rumors tend to be more prevalent in authoritarian

regimes due to the absence of credible public information sources, a situation precipitated by a lack of independent news media (Difonzo and Bordia 2007; Kapferer 1990). Rumors can be particularly destructive to authoritarian governments. In regimes where the media is tightly controlled, rumors emerge as an alternative form of media, directly competing with official information and mainstream media, and thereby constitute a counterpower against authority (Kapferer 1990). Indeed, past research has shown that rumors may have mobilizing power in authoritarian countries (Zhao 2004).

Second, the term "fact-checking" is often used interchangeably with "rumor-debunking," as both involve calling out false claims. However, it is crucial to distinguish between these two terms within authoritarian regimes. Fang (2022) conceived authoritarian rumor-debunking as "a co-optation of fact-checking" by political powers. Along this line, we argue that authoritarian rumor-debunking differs significantly from journalistic fact-checking in terms of *motive, target*, and *outcome*. While the motive behind fact-checking is to verify the accuracy or truthfulness of claims made by public figures so as to uphold the truth (Graves et al. 2016), debunking rumors often aims to reduce public expressions of negative sentiment such as fear and anxiety, especially during times of crises (Rosnow 1988). Rumor-debunking in authoritarian regimes serves as a propaganda strategy to uphold the legitimacy of the ruling party and the government by dismissing dissenting voices as mere "rumors" (Fang 2022). Next, the primary targets of fact-checking in the Western democratic contexts are claims made by public figures, such as politicians (Graves 2016), whereas rumor-debunking in authoritarian regimes focuses more on foreign or domestic dissidents (Fang 2022). Lastly, the conclusions that fact-checking reaches could range from (partially) true to false, but rumor-debunking predominantly involves denouncing statements as "rumors," indicating that they are completely false.

In sum, we used "rumor-debunking" instead of "fact-checking" in this article to describe the efforts by the Chinese government and media outlets to combat "rumors" during the COVID-19 pandemic.

Rumor-Debunking During the COVID-19 Pandemic

The COVID-19 pandemic has emerged as the most significant global challenge since World War II. Concurrently, the proliferation of disinformation and conspiracy theories surrounding COVID-19 has led to the so-called "infodemic."[3] In China, the debunking of COVID-19 rumors emerged and evolved alongside the pandemic itself. As early as when cases of pneumonia of unknown origin were reported, several individuals, including Dr. Li Wenliang, who later died of the disease, were reprimanded by the police for what was initially deemed as spreading false information.[4] As the pandemic developed, the scope of debunked rumors expanded from those concerning the number of new cases and deaths, to lockdowns, and later, vaccine-related misinformation. Rumor-debunking remains an important strategy in the Chinese propaganda system during the pandemic (Fang 2022). Although we argue that rumor-debunking is conceptually distinct from fact-checking, the two share several key components: fact-checkers or rumor debunkers, selection of claims/rumors, and debunking approaches.

Rumor-Debunkers. Rumor-debunkers refer to social media accounts that regularly publish rumor-debunking messages. These accounts often belong to government bodies, media outlets, or fact-checking organizations. Their work typically involves selecting rumors deemed harmful or crucial, researching and verifying the accuracy of these rumors, and crafting debunking messages for dissemination to the public.

Studying rumor-debunkers is essential within the context of Chinese information governance. The act of defining selected claims as "rumors" demonstrates a contest of power and authority to speak (Kapferer 1990). Empirical studies show that multiple actors have participated in rumor-debunking in China, including state media, government institutions, and social media platforms (Deng et al. 2020). Identifying the primary rumor debunkers offers critical insights into the distribution of power dynamics within the country's stringent information governance structure and the broader strategies of information control within authoritarian regimes. Therefore, we first ask:

RQ1: Who were the primary rumor-debunkers of COVID-19 related rumors on Chinese social media?

Selection of Rumors. Selecting claims for verification is one vital step in rumor-debunking. The question of which claims should be selected is not merely about feasibility but more about the perceived significance of the claims, or "rumors" (Graves 2016). These decisions are largely influenced by the perceived roles of the entities. For example, news organizations that embrace a monitorial role—seeing their mission as monitoring the government and holding those in power accountable—would be more likely to select claims from politicians and those related to the general public's interest (Steensen et al. 2023). However, Chinese news organizations often adopt a "collaborative role" with a mission to support the government and uphold social stability and harmony (Chen and Koo 2022). Therefore, news media in China are more likely to select rumors that may harm the government's image or disrupt social order.

The selection of claims on Chinese social media unveils government's priorities and biases. This study focuses on the types of rumors that were frequently debunked during the pandemic on Chinese social media. We specifically examine two attributes of debunked rumors: topics and sources of rumors. The topics of selected rumors indicate which domains are considered more susceptible to the harms brought by rumors, which are suggestive of government resource allocation regarding rumor debunking. The sources of selected rumors represent the primary targets of Chinese rumor-debunking efforts, indicating the potential biases or unfair treatment in the government information management. Therefore, we ask:

RQ2: What types of COVID-19 related rumors were more likely to be selected for debunking regarding (a) topics and (b) sources of rumors on Chinese social media?

Debunking Approaches. In addition, our analysis extends to the debunking approaches employed by rumor-debunkers. Researchers identify three primary approaches to debunking rumors: denial, refutation, and attack (Paek and Hove 2019). *Denial* is

simply a direct counterstatement to the original rumors without supporting evidence, whereas *refutation* involves countering rumors with factual evidence (Yang and Paek 2018). The *attack* approach includes responses that contain threatening information or potential punishments directed at those who spread the rumors. These different response types significantly impact the effectiveness of rumor control efforts (Paek and Hove 2019; Gao et al. 2022).

Scholars have revealed some patterns of the debunking approaches used by Chinese rumor-debunkers. In their comparative study of fact-checking work between the United States and China, Liu and Zhou (2022) found that the Chinese fact-checkers were less likely to provide either data-based or narrative evidence than their U.S. counterparts. It is also worth noting that Chinese social media such as Weibo has implemented stringent regulatory strategies to inhibit rumor spreading (Zeng et al. 2017). Both the platform and the government have debunked rumors by declaring punishment of the rumormongers (Gao et al. 2022). Taken together, we ask:

> **RQ3**: What debunking approaches were used regarding (a) factual evidence and (b) punishment information for COVID-19 related rumors on Chinese social media?

Method

Data. We collected a total of 5,677 posts published and reposted by the official rumor-debunking account @WeiboPiyao[5] from December 1, 2019 to August 1, 2022. Since no established Weibo API is available, we used a web scraping tool "Weibo Spider" to collect data.[6] We selected this account based on the following rationales: @WeiboPiyao is an official account of Chinese social media Weibo, with more than 2.3 million followers. @WeiboPiyao is a prominent rumor-debunking source, which routinely re-posts rumor-debunking content originally published on other accounts.[7] Hence, the collected data encompasses rumor-debunking messages originally posted by a wide array of accounts on Weibo, including government accounts such as the police and Internet information office at various administrative levels, news media outlets, and accounts of other stakeholders. Subsequently, we used a dictionary[8] lookup method to remove irrelevant posts. As a result, this winnowing process yielded a valid sample of 2,938 posts for content analysis.

Coding. The codebook was developed based on previous research (e.g., Carnahan et al. 2022; Kim et al. 2021) and insights gained from qualitative analysis of the subsample. We conducted a qualitative content analysis of 290 randomly sampled posts (roughly 10% of the entire sample). We employed open-coding techniques to identify themes and patterns relevant to the RQs. "Rumor-debunkers" are "source of the rumor-debunking message," including various types of communicators such as government agencies, media outlets, or other entities. "Topic of rumors" encompasses various domains, such as everyday life, new infections/deaths, lockdown/policies, and so forth. "Sources of rumors" focus on the origin of the rumors, such as social media, individuals, and unspecified sources. "Evidence" refers to whether the message includes clear, verifiable information that supports the debunking of the rumor.
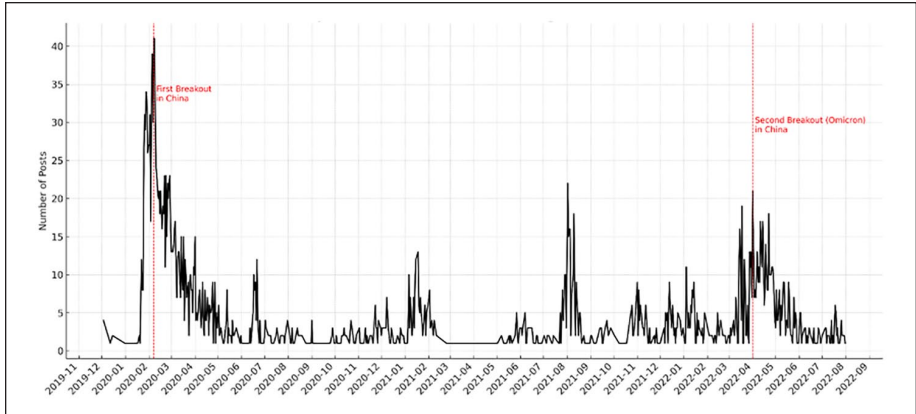
Figure 1. Daily counts of COVID-19-related rumor-debunking posts on Weibo over time.

"Punishment" refers to any form of punitive action imposed on those rumormongers (see Supplemental Information file, Appendix A for example posts).

To ensure the reliability of coding, two trained coders independently coded the 290 posts. After several rounds of coding and discussion, the intercoder reliability reached an average *Kohen's K* value of 0.90 across all codes (see Supplemental Information file, Appendix B). The same two coders coded the remaining posts independently.

Results

To grasp the overall pattern of the rumor-debunking posts on Weibo during the COVID-19 pandemic, we begin with a descriptive analysis of the collected data from December 1, 2019, to August 1, 2022 (see Figures 1). The results revealed distinct periods of heightened activity, with two notable peaks in the volume of rumor-debunking posts. In February 2020, there was a significant surge with daily posts reaching as high as 40, coinciding with the initial outbreak of COVID-19 in China. While post volumes varied in the subsequent months, another pronounced spike was observed in April 2022, with daily posts climbing to 20. This aligns with the emergence of the Omicron variant in China.

For **RQ1**, the result in Table 1 shows that *government accounts* are the predominant sources of the rumor-debunking information on Weibo, accounting for 61.33 percent of the sample. These accounts include government departments at various administrative levels, such as the national police department (China Police Online: @中国警方在线), regional governments (Hangzhou Release: @杭州发布), and local police departments (Safe Beijing: @平安北京). Following government accounts, media accounts contribute 21.03% of the debunking posts, which are primarily from both national and regional mainstream media including party media (e.g., People's Net: @人民网, Zhejiang Daily: @浙江日报) and market-driven media (e.g., Shanghai Morning Post; @新闻晨报, Yangtze Evening Post: @扬子晚报). The remaining posts (17.63%)

come from individual users, organizations, and independent fact-checking accounts (e.g., Catching Rumors: @捉谣记). These results indicate that the government and mainstream media hold the authority and power in rumor-debunking in China.

For the topics of rumors (*RQ2a*) being selected for debunking, it shows that *everyday life* (52.45%) ranked the top, followed by *new infections/deaths* (31.59%), *lockdowns* (10.96%), *treatments* (2.62%), and *vaccines* (2.38%). The "everyday life" category includes rumors that directly impacted the daily routines and well-being of the public, such as rumors about food supply disruptions during lockdowns or false claims that hospitals were refusing to accept new patients. The distribution of rumor topics highlights the emphasis on addressing rumors that could potentially cause widespread disruption and panic among the general public. This ensured that accurate and practical information was maintained throughout the crisis.

For the sources of rumors (*RQ2b*), it reveals that the majority of the debunking messages cited *social media* platforms as the origin (53.13%), which underscores social media platforms like Weibo and WeChat as central arenas for rumor spread in China. Due to their wide reach and rapid dissemination capabilities, rumors on these platforms can quickly gain traction, potentially causing widespread panic among the public. In addition, *individuals*, ordinary people sharing unverified or speculative information, were responsible for 8.13 percent of the rumors. A big proportion of the rumors (38.73%) came from *unspecified sources*. This highlights a critical aspect of rumor-debunking in China: lacking specified information about the origin of the rumors; instead, they simply call out the claims as false.

Regarding the use of evidence (*RQ3a*), more than half of the rumor-debunking messages do not contain evidence (55.00%). For the use of punishment (*RQ3b*), about one-fifth (19.37%) report punishment for rumormongers. Our post-hoc qualitative analysis of debunking messages that include information about punishment has further demonstrated the patterns of punitive actions taken against those spreading rumors. For minor infractions, such as less harmful or less widespread rumors, the typical punishments include account suspensions or reprimands. For instance, the *@Yulin Cyber Police* issued a reprimand to a woman who falsely claimed that the town of Bobai would go into lockdown the following day.[9] However, for more severe cases, particularly those involving rumors that could potentially cause significant public disorder or harm, the punishments are notably harsher. An example involves a 38-year-old man from Sichuan Province who was detained for four days for falsely claiming that three customers at a hotpot restaurant in Shiyang Town had contracted the coronavirus.[10] He spread this false information via WeChat in an attempt to gain attention, which prompted swift police action under the Public Security Administration Punishment Law. These examples demonstrate the varying degrees of punitive measures, illustrating how responses escalate based on the severity and potential impact of the rumors (see Supplemental Information, Appendix A punishment section).

Conclusion

In Study 1, we conducted a content analysis of the rumor-debunking posts on Weibo during COVID-19 and yielded several intriguing findings. Government-affiliated

Table 1. Summary of Attributes of Rumor-Debunking Posts.

	$N = 2,938$	%
Source		
Government accounts	1,802	61.33
Media accounts	618	21.03
Others	518	17.63
Topics of Rumor		
Everyday life	1,541	52.45
New infections/deaths	928	31.59
Lockdown/policies	322	10.96
Treatments	77	2.62
Vaccines	70	2.38
Rumor Source		
Social media	1,561	53.13
Individual	239	8.13
Unspecified	1,138	38.73
Evidence		
Present	1,322	45.00
Not present	1,616	55.00
Punishment		
Present	569	19.37
Not present	2,369	80.63

accounts were the primary rumor-debunkers on Weibo. The majority of rumormongers were ordinary citizens. Regarding debunking strategies, while more than half of the posts did not contain evidence, about one-fifth of the posts included information about penalties imposed on rumormongers. In addition, we noticed that the comment sections under contentious rumor-debunking posts were censored, as shown in either closed comment sections or "recommended comments by the account holder." To examine how these factors may influence credibility perceptions of the rumor-debunking messages among the Chinese public, we conducted a conjoint experiment in Study 2.

Study 2 Determinants of Perceived Credibility of Rumor-debunking Messages

Credibility perceptions rest on people's judgments of the veracity of a message (Appelman and Sundar 2016). It indicates how well people approve the message they receive. People's perceptions of message credibility could affect the way they develop subsequent judgments and consumption behavior (Appelman and Sundar 2016). Credibility perceptions have been used as a key measure to indicate the effectiveness of misinformation correction efforts. Studies show that credibility perception is the first step that subsequently reduces misperceptions of and prompts engagement behavior with corrective messages (Kim et al. 2021; Lu and Zhong 2022).

Credibility perception has distinct implications in different social and political environments. In a society with free information flow, information that is deemed credible champions in the marketplace of ideas by encouraging consumption behavior and bringing economic benefits. In authoritarian regimes like China, credibility judgment is not only an indicator of audience preferences but also a parameter of effective governance. Given that the public mostly lacks the resources to verify online information for themselves, rumor-debunking serves to entice the public to consume and buy in official discourses. As online rumors could decrease public trust in the government, evidence-based and vivid rumor rebuttals have greater potential to recover political trust in the governments than simple rebuttals (Huang 2017).

Building on Study 1, we examine people's credibility perceptions of rumor-debunking messages through a multidimensional framework that encompasses the debunking sources, strategies, and context in Study 2.

Debunking Sources. Debunking sources, or rumor-debunkers, are an important component that determines the credibility of rumor-debunking messages. In social psychology, source cue information forms the basis for credibility perception (Reinhard and Sporer 2010). According to the dual processing theory (Chaiken and Trope 1999), most people do not have the motivation for systematic information processing; instead, they rely on simple judgment rules (i.e., heuristics) to evaluate messages. Source serves as an easy-to-use heuristic in information processing, because it is readily available in one's memory. Studies show that source characteristics such as attractiveness, competence, and expertise can influence credibility perceptions and subsequent decision-making (DeBono and Harnish 1988; Yoon et al. 1998).

Rumor-debunking messages can be complex for people to process, as it often involves multiple components, including the rumor, evidence used for debunking, and additional information cues like user comments adjacent to the original message. As such, debunking sources constitute a crucial heuristic for people to evaluate the credibility of the message. Among the different types of correction sources, both government entities and news media prove to be effective at reducing misperceptions (van der Meer and Jin 2020; Vraga and Bode 2017). In China, more than three-fourths of corrective messages on health-related topics were sent by organizational accounts (Zheng et al. 2021). The Chinese public bestows more trust in government sources than in news media and other sources when reading rumor-debunking messages on emerging science topics (Yu et al. 2022). Note that Study 1 identified government and media as two primary rumor-debunkers on Chinese social media during COVID-19. Given that previous research indicates people trust government sources more than media for rumor-debunking, we hypothesize:

> **H1**: People will perceive rumor-debunking messages posted by government accounts as more credible than those posted by news media accounts on Chinese social media.

Debunking Strategies. Regarding debunking strategies, factual claims used by debunkers can vary in specificity: they may be broad, lacking contextual detail (low specificity),

or specific, providing concrete information (high specificity). Message specificity originates from the vividness hypothesis, which posits that vivid information, which is "(a) emotionally interesting, (b) concrete and imagery provoking, and (c) proximate in a sensory, temporal, or spatial way," is more likely to be stored and remembered than pallid information (Nisbett and Ross 1980: 45). Message specificity taps into the degree of concreteness, such as the amount of detailed information provided. In public relations and advertising literature, message specificity is found to increase perceived credibility of claims made by companies (Ganz and Grimes 2018) and lead to positive evaluations of the companies (Robinson and Eilert 2018). Most recently, Carnahan et al (2022) extended this concept to studying fact-checking messages but found that message specificity did not impact credibility assessments or intent to share such messages.

While the differentiation between low and high specificity regarding factual elaboration provides insights on crafting credible rumor-debunking messages, few attempts have delved into what aspects of the stated information should be specified. A rumor-debunking message involves components such as rumor identification, refutation, and intervention. When identifying the rumor, debunkers can enhance its specificity by stating the details, such as the actor and location, pertinent to the rumor. As for refutation, debunkers can counter the rumor by either offering reliable evidence and reasoning or simply rebutting the rumor, the former of which are favored by the public (Gao et al. 2022). Both rumor identification and refutation strategies indicate the transparency of the debunking processes. If they are less ambiguous, such a message is deemed as more credible, because production transparency is a precursor of people's credibility perceptions of journalistic content (Koliska and Chadha 2018; Lu and Zhen 2024). Lastly, debunkers can announce specific interventions to offer closure. Punishing the rumormongers, also known as the attack approach (Paek and Hove 2019), is a common information control strategy used by the Chinese government to inject political and social values into citizenry (Ruan et al. 2021). Tracing and punishing the rumormongers carries a certain level of deterrent power (Gao et al. 2022), which could be translated to people's credibility perception, effectively curbing the spread of rumors over time. Taken together, we propose:

> **H2**: People will perceive rumor-debunking messages with high specificity regarding (a) rumor identification, (b) refutation, and (c) intervention as more credible than those with low specificity on Chinese social media.

(Non)participatory Rumor-Debunking. Social media platforms provide users with abundant opportunities to engage with rumor-debunking messages through liking, commenting, and sharing. User engagement with rumor-debunking messages is crucial, because such engagement increases the visibility of rumor-debunking efforts in the online ecosystem and helps such messages distribute widely to different online communities. Furthermore, user engagement with rumor-debunking messages can shape the way that later users perceive and engage with these messages. Here, we explore two types of cues pertinent to the participatory aspects of rumor debunking–social endorsement and recommended comments on Weibo.

Social endorsement in the form of likes, shares, and comments can shape the credibility perceptions of the message content. From an information processing perspective, social endorsement cues serve as mental shortcuts for people to make simple and effortless judgment based on other users' favorability of the content in an information-abundant environment. According to Sundar (2012), social endorsement cues trigger bandwagon heuristic, referring to the tendency that people adopt certain behaviors or attitudes simply because others are doing so. Empirical research shows that a high number of likes, shares, and comments on social media can increase people's credibility perceptions of news headlines and other types of online content (Luo et al. 2022; Metzger et al. 2010; Xu 2013). Thus, we hypothesize:

H3: People will perceive rumor-debunking messages with high levels of social endorsement as more credible than those with low levels of social endorsement on Chinese social media.

Another important aspect of the participatory nature of online rumor-debunking is whether the rumor-debunkers allow other users to leave comments on the original posts. Social media platforms such as Twitter and Facebook provide their users with functionalities to prioritize or selectively show some of the comments on their posts. Initially, this function is designed to highlight relevant and engaging content while reducing spam and low-quality content so as to improve user experience and encourage user engagement. This practice will likely influence credibility perceptions in light of the *warranting theory*, which holds that the credibility of information shared online is shaped by the difficulty of manipulating or controlling that information (Walther and Parks 2002). In other words, information that is hard to manipulate or control tends to be perceived as more credible, whereas information that is subject to control is deemed as less credible. User-generated content, such as publicly available user comments under a specific user's profile, can function as an important warrant, signaling that the information about this specific user is valid and reliable. In the case of selectively presented user comments, such information is perceived to be susceptible to manipulation by this specific user, which further suggests that these user comments may not be trustworthy. Based on this notion, Carnahan et al. (2022) argued that hidden comments under a social media post lack the warranting value and, therefore, reduce perceived credibility of the post, which, however, were not supported by empirical evidence.

In China, rumor-debunking accounts on Weibo often utilize the "recommended comments by the account holder" (博主精选评论) to curate user comments on their posts. This functionality enables the account holders to show positive comments and hide negative ones. For account holders, they can leverage the warranting value of "recommended comments" to bolster public trust in their content. Indeed, the comment section that is presented in the form of "recommended comments" on Weibo are predominantly pro-government, which could greatly promote supportive discourses on government policies from the grassroots (Wong and Liang 2023). However, from the public's side, such efforts may be viewed as a way to manipulate public opinion, thus reducing the credibility of the original messages. Wong and Liang (2023) revealed

that the Chinese public developed an appetite toward alternative views on social and political issues. As such, the "recommended comments by the account holder," by constraining the market of ideas on Chinese social media, may be viewed as a deception cue, thus dampening people's credibility perceptions. Therefore, we posit:

H4: People will perceive rumor-debunking messages without recommended comments as more credible than those with recommended comments on Chinese social media.

Method

Participants. The study protocol was approved by North Dakota State University's Institutional Review Board on December 8, 2022. From December 13, 2022 to January 5, 2023, we recruited participants from the WJX online panel,[11] which consists of diverse internet users based in China. The participants were compensated ten *yuan* for completion. After deleting those failing to correctly answer the attention check questions (n=95), we retained 493 respondents in the final sample. Among them, 213 were male, with an average age of 30.95 years old ($SD=6.85$). The median educational level was college degree. The median monthly household income was 8,001–15,001 *yuan*. Note that the sample is not representative of the general Chinese population. Therefore, Study 2 focuses on identifying the causal effects of different attributes in rumor-debunking messages instead of generalizing the findings to the Chinese population.

Procedure. The study adopted conjoint experimental design. Conjoint experiments are frequently used in marketing research (Green and Srinivasan 1990) and have been used in communication research to study selective exposure behavior (Lu et al. 2023; Mukerjee and Yang 2021). Unlike a simple random experiment where participants are exposed to one randomized stimulus, conjoint experiments expose participants to multiple stimuli for comparison. Conjoint experiments have several advantages over simple random assignment. It allows researchers to (a) simultaneously test the causal effects of a large number of factors with a relatively small sample size, (b) grasp people's sensitivity to small variations in attributes presented in the stimuli, thus enhancing internal validity, and (c) create a scenario that mimics the high-choice information environment, where exposure to and comparison between multiple pieces of information are common.

After consenting, participants answered social demographics questions and were instructed to complete four tasks, each involving reading one pair of randomized COVID-19 related rumor-debunking posts on Weibo and rating the credibility of them. The posts varied with source, rumor identification, refutation, intervention, social endorsement, and recommended comments (for more information, see stimuli construction section below). We used credibility ratings instead of selection because the former bears more granularity when tapping into people's preferences. We embedded attention check questions in three of the four tasks for quality control.

Stimuli Construction. We used an online tool[12] to fabricate Weibo posts that looked realistic in appearance. We selected "new reports of local COVID-19 cases" as the

topic, because it remains one of the most common rumors distributed online during the study period. We chose Lanzhou, a northwestern city in China, as the location mentioned in the rumor. Unlike Shanghai, Beijing, Wuhan, and other major cities, Lanzhou bore relatively low media attention, which allowed us to reduce the noises from potential confounders.

The stimuli varied along six independent variables under study. For *source*, we differentiated between local government and local media, to which the post was attributed. For debunking strategies, we varied *rumor identification, refutation*, and *intervention* with high and low levels of message specificity, respectively. We also manipulated the two participatory cues: *social endorsement* was varied between high and low levels; *recommended comments* were manipulated into the availability of commenting or none. In sum, each pair the participants read was two distinct permutations of these six attributes (see Figure 2 for sample stimuli).
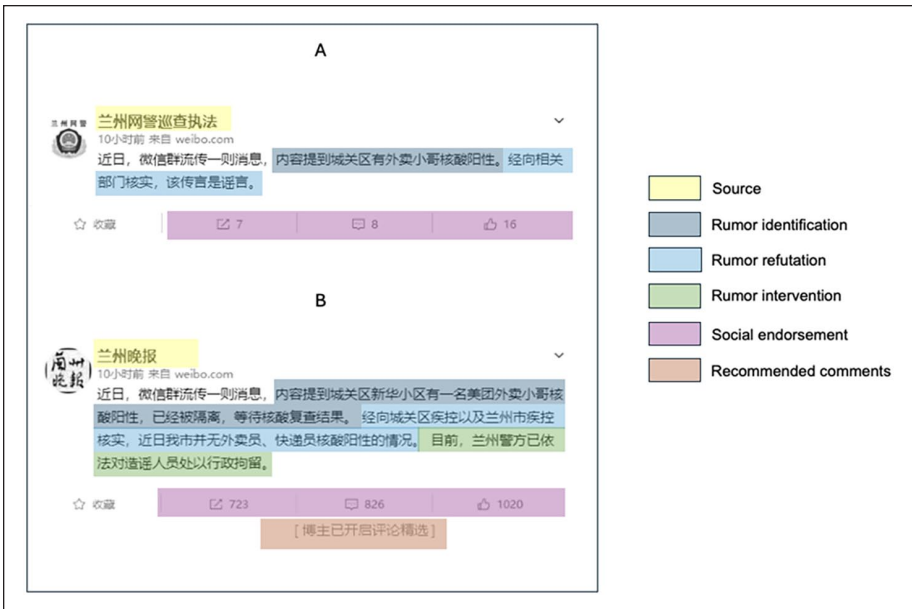

Figure 2. Sample stimuli.
Note. There are a total of 64 different permutations of the attributes. For simplicity, we illustrated two examples here. The full sample of stimuli is available upon request. Post (A) illustrates a rumor-debunking post with local government as the source, rumor identification with low specificity, rumor refutation with low specificity, intervention with low specificity, low social endorsement, and no recommended comments. Post (B) illustrates a rumor-debunking post with local media as the source, rumor identification with high specificity, rumor refutation with high specificity, intervention with high specificity, high social endorsement, and recommended comments. For English translation of the two posts, please see Supplemental Information, Appendix D.

Analytical Approach. We had eight data points per participant, each of which corresponds to one of the eight posts in the four pairs they had read. The six attributes of the Weibo posts were dummy-coded and used as the independent variables. The dependent variable—credibility perception was measured with a seven-point scale (1=describes very poorly, 7=describes very well) on three descriptors: accurate, authentic, and believable (Appelman and Sundar 2016; $M=4.81$–5.14, $SD=1.31$–1.56, $\alpha=0.92$–0.95).

For the analysis, we used ordinary least squares linear regression to estimate the average marginal component effects (Hainmueller et al. 2014). Given each participant was involved in four tasks, we corrected the standard errors to address within-respondent clustering.

Results. Figure 3 shows the results of the marginal effect of each attribute averaged over the joint distribution of other attributes (Supplemental Information file, Appendix C for concrete results). It reveals no significant difference between local government and local media as debunkers on credibility perceptions of the messages ($B=-0.06$, *Robust SE*=0.07, *CI*=[−0.20, 0.08], $p=.37$), not supporting H1.

Regarding debunking strategies, messages that identified the rumor with details (high specificity in rumor identification) were perceived as more credible than those without ($B=0.25$, *Robust SE*=0.09, *CI*=[0.07, 0.42], $p=.005$). Messages that refuted the rumor with evidence (high specificity in rumor refutation) were perceived as more credible than simple rebuttal ($B=0.30$, *Robust SE*=0.05, *CI*=[0.19, 0.40], $p<.001$). In addition, messages stating punishment toward rumormongers (high specificity in rumor intervention) were perceived as more credible than those without such information ($B=0.54$, *Robust SE*=0.06, *CI*=[0.42, 0.67], $p<.001$). Hence, H2a, H2b, and H2c were supported.

Finally, while levels of social endorsement positively predicted credibility perceptions ($B=0.12$, *Robust SE*=0.06, *CI*=[0.004, 0.24], $p=.04$), the availability of "recommended comments by account holder" did not significantly affect credibility perceptions ($B=-0.09$, *Robust SE*=0.06, *CI*=[−0.20, 0.03], $p=.13$). Therefore, H3 was supported, but H4 was not supported.

Conclusion

Study 2 used a conjoint experiment to examine how various features of rumor-debunking messages on Weibo influenced people's credibility perceptions. The results show no significant difference between local government and local media as debunkers in shaping credibility perceptions. However, message specificity mattered—rumor identification with details, evidence-based refutation, and punishment of rumormongers as intervention—significantly boosted credibility perception. Regarding participatory cues, social endorsement enhanced credibility perception, but the availability of "recommended comments" did not.
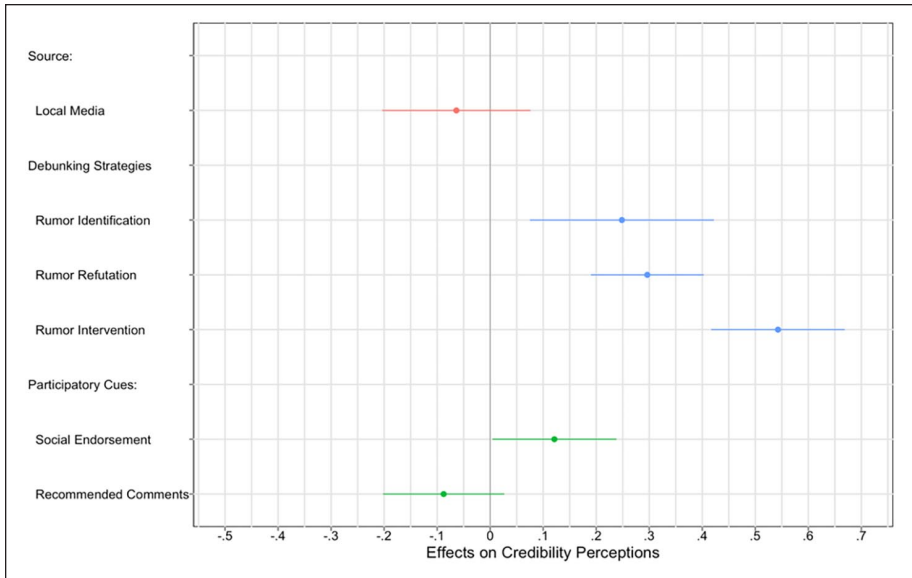
Figure 3. Effects of rumor-debunking message attributes on credibility perceptions.
Note. The figure illustrates estimates of the effects of rumor-debunking message attributes on credibility perceptions. Estimates are based on the ordinary least squares (OLS) regression model with clustered standard errors. The bars denote 95% confidence intervals for each attribute; reference of each attribute is not presented. For source, local government was the reference group; for debunking strategies, low specificity were the reference categories for rumor identification, refutation, and intervention, respectively; for participatory cues, low social endorsement level and no availability of recommended comments were the reference categories, respectively.

General Discussion

This research joins the growing efforts that examine the dynamics and implications of rumor-debunking on social media in non-WEIRD countries (Mare and Munoriyarwa 2022; Porter and Wood 2021; Song et al. 2021). Using the case of COVID-19 rumor-debunking on Chinese Weibo, we identified key features of rumor-debunking messages (Study 1) and their effects on credibility perceptions among Chinese internet users (Study 2). The findings contribute to our understanding of authoritarian information management in general and rumor-debunking efforts on Chinese social media in particular. In what follows, we elaborate on the theoretical implications.

First, given that rumors are prevalent and politically sensitive in authoritarian countries, the governments have taken a proactive role to manage this information risk to safeguard their legitimacy and social stability. Unlike its Western counterparts where fact-checkers are typically news media and independent fact-checkers (Graves 2016), our findings show that rumor-debunking on Chinese social media is primarily driven by the government. This indicates that the power of selecting claims and deciding what is truth largely rests with the government, which align with scholarly

accounts that the Chinese government uses fact-checking as a way to propagate official discourses and downplay the unwanted information (Fang 2022). For the pandemic where social uncertainties and anxiety prevailed, the government sought to maintain its primitive authority in effective governance through rumor-debunking. This centralized control over information highlights the lack of compatibility between government-led information governance and democratic ideals of a free press, thus raising questions about the impartiality and transparency within "fact-checking journalism" emerged in authoritarian regimes (Liu and Zhou 2022).

This primitive authority in information management by the Chinese government also manifests in the types of rumors selected and the debunking strategies employed. The results from Study 1 demonstrate that the topics of rumors are mostly about everyday life at the grassroot level. Government- or institution-related rumors such as lockdown policy and medical treatment of COVID-19 were only a small portion. This draws a sharp difference from the WEIRD counterparts, which often fact-check claims made by politicians and other public figures (Graves 2016). Rumor-debunking in China, even those performed by news media, predominantly targets statements originating from ordinary citizens, which signify the collaborative role that media play in the country (Chen and Koo 2022). In other words, rumor-debunking in authoritarian regimes serves to support the government by curbing factors that threaten social stability. Grassroots rumormongers became the common enemy that both the government and the public could blame. In this regard, the punishment on rumormongers exerted by the government could serve as a powerful tool to pacify public anxiety and rebuild social orders during the pandemic.

Crucially, our post-hoc qualitative analysis of rumor-debunking messages that include information about punishment reveals different levels of punitive actions imposed on rumormongers. These actions ranged from temporary bans or permanent account suspensions on Weibo to more severe legal consequences including administrative detention, depending on the severity of the rumor and perceived malicious intent of the rumormonger. Enforcement of these punishments typically involves a collaborative effort between social media platforms and state authorities, including police departments at various administrative levels. Platforms like Weibo are tasked with the initial flagging or removal of detected rumors. Once identified, these instances can be escalated to relevant governmental bodies and subject to further evaluation and administrative penalties. This integration of state power with platform-based enforcement significantly distinguishes Chinese rumor-debunking practices from fact-checking in the Western democracies, where platforms issue warnings or reduce the visibility of misinformation but government bodies are seldom involved in imposing legal sanctions. In China, the potential for state-imposed punishments adds a layer of gravity to the rumor-debunking process, reinforcing the legitimacy and authority of initiatives like @WeiboPiyao. This approach not only underscores the government's determination to crack down online rumors but also serves as a deterrent against future attempts to spread anti-regime information, reflecting a broader "stability maintenance" strategy in a nationwide law enforcement system that curtails government critics before they can take actions (Zeng et al. 2017).

Another characteristic of authoritarian rumor-debunking is the limited involvement in fact-based rebuttal, with fewer than half of the rumor-debunking messages containing detailed evidence (Study 1). One plausible explanation is that factual elaboration may open the room for interpretation of the presented facts. Given the evolving nature of the COVID-19 pandemic, these elaborations could be insufficient or need to be updated at any given time. To correct the rumor-debunking messages presents a potential backlash to the authorities—in our case, the Chinese government. So simple rebuttal seems to be a viable and efficient strategy for the government to close the file and uphold its primitive authority in defining what is truth. In sum, government-led rumor-debunking initiatives in China have drawn a clear boundary on what can be debunked and what cannot and developed a set of tacit rules about how to debunk rumors.

Despite the limited efforts in fact-based rumor-debunking, our experiment demonstrates that the Chinese public perceived rumor-debunking messages featuring high specificity more credible than those with low specificity (Study 2). This sounds an alarm to rumor-debunkers in China that simple rebuttal, albeit its seemingly effectiveness in closure, may not win the hearts and minds of the public. It is crucial to provide reasonable details when debunking rumors to boost credibility, because this credibility will further encourage the public to disseminate rumor-debunking messages more widely (Lu and Zhong 2022). Another important finding is that punishment imposed on rumormongers remained a strong predictor of credibility perceptions. Simulation research shows that government punishment could curb the diffusion of online rumors (see Li and Ma 2017). Our study pinpoints the possible mechanism whereby governmental punishment of rumormongers may work to curtail misinformation by strengthening people's perception of credibility of rumor-debunking efforts. This increased credibility could be closely tied to the deterrent power signified through those punishments (Gao et al. 2022). This finding also aligns with the argument that the Chinese public acts more as consumers in search of guidance, rather than as citizens engaged with political and social issues (see Liu and Zhou 2022). The Chinese public discerning rumor-debunking messages containing punishment information more credible suggests that they rely on authoritative sources like @WeiboPiyao and other rumor-debunking accounts not merely for factual accuracy but also for cues on permissible speech and behavior under government policies.

Finally, our research offers two important insights into the participatory nature of rumor-debunking on Chinese social media. For one, the results underscore that social endorsement in the form of likes and shares could boost credibility perception of rumor-debunking messages, corroborating prior research (Luo et al. 2002; Metzger et al. 2010; Xu 2013). This indicates that the Chinese public largely relied on this bandwagon heuristic to compare and make credibility judgments of multiple messages on Weibo. For the other, we observed that rumor debunkers had taken advantage of the "recommended comments" function on Weibo (Study 1). This function can be either manually activated by the account holders or automatically triggered by the platform, especially with the surge of politically sensitive comments on Weibo. However, turning this function on did not sway credibility perceptions (Study 2), which is in line with Carnahan et al. (2022)'s findings in the United States. Our interpretation is that

people may have different understanding about the warranting value of "recommended comments." Some may perceive this as a censorship cue used for deceiving the readers through hiding criticism toward the original posts; others may regard it as being used to reduce uncivil and unruly content under the main posts. Future research should explore people's perceived functionalities of "recommended comments" on social media to address the null effects in our study.

Limitations and Future Research Directions

The study is not without limitations. Although @weibopiyao is a hub of extensive and diverse rumor-debunking information, our findings cannot be generalized to the broad rumor-debunking practices on different social media platforms. We wish future studies could go beyond the pandemic and the platform Weibo to offer additional evidence on the patterns and effects of rumor-debunking efforts in China and other non-WEIRD countries.

Second, it should be acknowledged that we only focused on a limited set of message features, leaving such attributes as emotional appeals, linguistic features, and expert endorsement untended (Lee and Britt 2023; Lu and Zhong 2022; Vraga and Bode 2017). Future scholars could delve into these various features of rumor-debunking messages and estimate their relative effects on credibility perceptions and engagement intentions among the public.

Declaration of Conflicting Interests

Funding

ORCID iD

Shuning Lu https://orcid.org/0000-0001-8259-7987

Supplemental Material

Supplemental material for this article is available online.

Notes

1. Meseret, E. "A global rise in government-led fact-checking initiatives cause concern, worries of misuse." *Poynter*. Feb. 8, 2024. https://www.poynter.org/ifcn/2024/government-fact-checking-initiatives-worldwide/
2. Rumor refutation focuses on using structured and evidence-based arguments to correct rumors. Rumor correction involves providing accurate information to correct falsehood. Rumor debunking may or may not have arguments and accurate information when

disproving rumors. That said, rumor debunking captures the broad spectrum of corrective efforts led by the Chinese government.

3. Stolberg, S. G. and Weiland, N. "Study Finds 'Single Largest Driver' of Coronavirus Misinformation: Trump." *The New York Times*, Sep. 30, 2020. https://www.nytimes.com /2020/09/30/us/politics/trump-coronavirus-misinformation.html

4. Dong, J. "Two years after his death, the Chinese doctor who warned of the virus is remembered." *The New York Times*, Feb. 7, 2022. https://www.nytimes.com/2022/02/07/world/ asia/chinese-doctor-li-wenliang-covid-warning.html

5. Piyao translates to "debunking rumor."

6. https://github.com/dataabc/weiboSpider

7 The rules by the Cyberspace Administration of China only permit private online media, such as Weibo, to recirculate and re-edit information pieces by state-owned media and government (Liu and Zhou 2022).

8. Posts that do not contain any of these words are removed: '新冠', '肺炎', '疫情', '新增', '确诊', '隔离', '封控', '阳性', '密接', '疫苗', '核酸', '停运', '传染', '感染', '封城', '封校', '静默', '绿码', '疫', '物资'.

9. See post: http://m.weibo.cn/status/4465296183893067?

10. See post: http://m.weibo.cn/status/4701671500287696?

11. https://www.wjx.cn/

12. https://tools.miku.ac/meme-weibo/

References

Acerbi, A., S. Altay, and H. Mercier. 2022. "Research Note: Fighting Misinformation or Fighting for Information?" *Harvard Kennedy School Misinformation Review*. Published online January 12. doi: 10.37016/mr-2020-87.

Appelman, A., and S. S. Sundar. 2016. "Measuring Message Credibility: Construction and Validation of an Exclusive Scale." *Journalism & Mass Communication Quarterly* 93(1):59–79.

Carnahan, D, E. Ulusoy, R. Barry, J. McGraw, I. Virtue, and D. E. Bergan. 2022. "What Should I Believe? A Conjoint Analysis of the Influence of Message Characteristics on Belief in, Perceived Credibility of, and Intent to Share Political Posts." *Journal of Communication* 72(5):592–603.

Chaiken, S., and Y. Trope. 1999. *Dual-Process Theories in Social Psychology*. New York: Guilford Press.

Chang, C. 2021. "Information Credibility Under Authoritarian Rule: Evidence from China." *Political Communication* 38(6):793–813.

Chen, B., X. Chen, J. Pan, K. Liu, B. Xie, W. Wang, Y. Peng, F. Wang, N. Li, and J. Jiang. 2021. "Dissemination and Refutation of Rumors During the COVID-19 Outbreak in China: Infodemiology Study." *Journal of Medical Internet Research* 23(2):e22427.

Chen, B., and G. H. Koo. 2022. "Journalistic Roles and News Framing: A Comparative Framing Analysis of COVID-19 Pandemic Across China, South Korea, and the United States." *International Journal of Communication* 16:4254–74.

DeBono, K. G., and R. J. Harnish. 1988. "Source Expertise, Source Attractiveness, and the Processing of Persuasive Information: A Functional Approach." *Journal of Personality and Social Psychology* 55(4): 541–46.

Deng, Z., Q. Meng, Z. Huang, Z. Kang, and X. Liu. 2020. "Harmonic Resonance: Research on Public Opinion Guidance and Collaboration of Government Microblog about COVID-19." *Information Science* 38(8):79–87 [in Chinese].

Difonzo, N., and P. Bordia. 2007. "Rumors Influence: Toward a Dynamic Social Impact Theory of Rumor." In *The Science of Social Influence: Advances and Future Progress, ed. A. R. Pratkanis. Frontiers of Social Psychology*, 271–95. New York: Psychology Press.

Fang, K. 2022. "'*Rumor-Debunking' as a Propaganda and Censorship Strategy in China*." In Disinformation in the Global South, eds. H. Wasserman and D. Madrid-Morales, 108–22. Hoboken, NJ: John Wiley & Sons, Ltd.

Ganz, B., and Grimes A. 2018. "How Claim Specificity Can Improve Claim Credibility in Green Advertising." *Journal of Advertising Research* 58(4):476–86.

Gao, H., D. Guo, H. Yin, J. Wu, Z. Cao, and L. Li. 2022. "Strategies and Effectiveness of the Chinese Government Debunking COVID-19 Rumors on Sina Weibo: Evaluating from Emotions." *Journal of Applied Communication Research* 50(6):632–54.

Graves, L. 2016. *Deciding What's True: The Rise of Political Fact-Checking in American Journalism*. New York: Columbia University Press.

Graves, L. 2018. "Boundaries Not Drawn." *Journalism Studies* 19(5):613–31.

Graves, L., B. Nyhan, and J. Reifler. 2016. "Understanding Innovations in Journalistic Practice: A Field Experiment Examining Motivations for Fact-Checking." *Journal of Communication* 66(1):102–38.

Green, P. E., and Srinivasan, V. (1990). Conjoint Analysis in Marketing: New Developments With Implications for Research and Practice. *Journal of Marketing*, 54(4):3–19.

Grinberg, N., K. Joseph, L. Friedland, B. Swire-Thompson, and D. Lazer. 2019. "Fake news on Twitter during the 2016 U.S. presidential election." *Science* 363(6425):374–8.

Hainmueller, J., Hopkins, D. J., & Yamamoto, T. (2014). Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments. *Political analysis*, 22(1):1–30.

Huang, H. 2017. "A War of (mis) Information: The Political Effects of Rumors and Rumor Rebuttals in an Authoritarian Country." *British Journal of Political Science* 47(2):283–311.

Kapferer, J.-N. 1990. *Rumors: Uses, Interpretation and Necessity*. New York: Routledge.

Kim, S. C., E. K. Vraga, and J. Cook. 2021. "An Eye Tracking Approach to Understanding Misinformation and Correction Strategies on Social Media: The Mediating Role of Attention and Credibility to Reduce HPV Vaccine Misperceptions." *Health Communication* 36(13):1687–96.

Koliska, M., and K. Chadha. 2018. "Transparency in German newsrooms: Diffusion of a New Journalistic Norm?." *Journalism Studies* 19(16):2400–16.

Lee, J. J., K. A. Kang, M. P. Wang, S. Zhi Zhao, J. Yuen Ha Wong, S. O'Connor, S. Ching Yang, and S. Shin. 2020. "Associations between COVID-19 Misinformation Exposure and Belief with COVID-19 Knowledge and Preventive Behaviors: Cross-Sectional Online Study." *Journal of Medical Internet Research* 22(11):p.e22205.

Lee, J., and B. C. Britt. 2023. "Factbait: Emotionality of Fact-Checking Tweets and Users' Engagement during the 2020 U.S. Presidential Election and the COVID-19 Pandemic." *Digital Journalism*, 1–25. Published online September 07. doi:10.1080/21670811.2023.2253859.

Li, D., and J. Ma. 2017. "How the Government's Punishment and Individual's Sensitivity Affect the Rumor Spreading in Online Social Networks." *Physica A: Statistical Mechanics and its Applications* 469:284–92.

Liu, Y., and R. Zhou. 2022. "'Let's Check it Seriously': Localizing Fact-Checking Practice in China." *International Journal of Communication* 16:23.

Lu, S., H. Liang, and G. M. Masullo. 2023. "Selective Avoidance: Understanding how Position and Proportion of Online Incivility Influence News Engagement." *Communication Research* 50(4):387–409.

Lu, S., and L. Zhen. 2024. "From Transparency to Transactive Memory System: How do Newsrooms' GitHub Pages Shape News Outlet Credibility?." *Digital Journalism* 12(7): 945–64.

Lu, S., and L. Zhong. 2022. "From Believing to Sharing: Examining the Effects of Partisan Media's Correction of COVID-19 Vaccine Misinformation." *International Journal of Communication* 16(0):1–22.

Lu, Y., and J. Pan. 2021. "Capturing Clicks: How the Chinese Government Uses Clickbait to Compete for Visibility." *Political Communication* 38(1–2):23–54.

Luo, M., J. T. Hancock, and D. M. Markowitz. 2022. "Credibility Perceptions and Detection Accuracy of Fake News Headlines on Social Media: Effects of Truth-Bias and Endorsement Cues." *Communication Research* 49(2):171–95.

Madrid-Morales, D., and H. Wasserman. 2022. "Research Methods in Comparative Disinformation Studies." In *Disinformation in the Global South*, eds. H. Wasserman & D. Madrid-Morales, 41–57. Hoboken, NJ: Wiley-Blackwell.

Mare, A., and A. Munoriyarwa. 2022. "Guardians of Truth? Fact-Checking the 'Disinfodemic' in Southern Africa during the COVID-19 Pandemic." *Journal of African Media Studies* 14:63–79.

Metzger, M., A. Flanagin, and R. Medders. 2010. "Social and Heuristic Approaches to Credibility Evaluation Online." *Journal of Communication* 60:413–39.

Mukerjee, S., & Yang, T. (2021). Choosing to Avoid? A Conjoint Experimental Study to Understand Selective Exposure and Avoidance on Social Media. *Political Communication* 38(3):222–40.

Nisbett, R. E., and L. Ross. 1980. *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, NJ: Prentice-Hall.

Paek, H.-J., and T. Hove. 2019. "Mediating and Moderating Roles of Trust in Government in Effective Risk Rumor Management: A Test Case of Radiation-Contaminated Seafood in South Korea." *Risk Analysis* 39(12):2653–67.

Porter, E., and T. J. Wood. 2021. "The Global Effectiveness of Fact-Checking: Evidence from Simultaneous Experiments in Argentina, Nigeria, South Africa, and the United Kingdom." *Proceedings of the National Academy of Sciences* 118(37):e2104235118.

Reinhard, M-A., and S. L. Sporer. 2010. "Content Versus Source Cue Information as a Basis for Credibility Judgments: The Impact of Task Involvement." *Social Psychology* 41(2):93–104.

Robinson, S., and M. Eilert. 2018. "The Role of Message Specificity in Corporate Social Responsibility Communication." *Journal of Business Research* 90:260–8.

Roberts, M. 2018. *Censored: Distraction and Diversion Inside China's Great Firewall*. New Jersey: Princeton University Press.

Rosnow, R. L. 1988. "Rumor as Communication: A Contextualist Approach." *Journal of Communication* 38(1):12–28.

Ruan, L., J. Knockel, and M. Crete-Nishihata. 2021. "Information Control by Public Punishment: The Logic of Signalling Repression in China." *China Information* 35(2):133–57.

Schuldt, L. 2021. "Official Truths in a War on Fake News: Governmental Fact-Checking in Malaysia, Singapore, and Thailand." *Journal of Current Southeast Asian Affairs* 40(2): 340–71.

Steensen, S., B. Kalsnes, and O. Westlund. 2023. "The Limits of Live Fact-Checking: Epistemological Consequences of Introducing a Breaking News Logic to Political Fact-Checking." *New Media & Society*. Published online February 12. doi:10.1177/14614448231151436.

Song, Y., K. H. Kwon, L. Yin, Y. Fan, and B. Li. 2021. "The 'Parallel Pandemic' in the Context of China: The Spread of Rumors and Rumor-Corrections During COVID-19 in Chinese Social Media." *American Behavioral Scientist* 65(14):2014–36.

Sundar, S. S. 2012. "Social Psychology of Interactivity in Human-Website Interaction." In *Oxford Handbook of Internet Psychology*, eds. A. N. Joinson, K. Y. A. McKenna, T. Postmes and U.-D. Reips, 89–104. Oxford, England: Oxford University Press.

van der Meer, T. G. L. A., and Y. Jin. 2020. "Seeking Formula for Misinformation Treatment in Public Health Crises: The Effects of Corrective Information Type and Source." *Health Communication* 35(5):560–75.

Vraga, E. K., and L. Bode. 2017. "Using Expert Sources to Correct Health Misinformation in Social Media." *Science Communication* 39(5):621–45.

Walther, J., and M. Parks. 2002. "*Cues Filtered Out, Cues Filtered In*: Computer-Mediated Communication and Relationships. " In G. R. Miller (Ed.), *The Handbook of Interpersonal Communication*, 529–63. Thousand Oaks, CA: Sage Publications.

Wong, S. H-W., and J. Liang. 2023. "Attraction or Distraction? Impacts of Pro-regime Social Media Comments on Chinese Netizens." *Political Behavior* 45(3):1071–95.

Xiang, C., and R. Neo. 2024. "Political Fact-Checking and Its Effects on Public Attitudes: Experimental Evidence from China." *The China Quarterly* 2024: 1–12.

Xu, Q. 2013. "Social Recommendation, Source Credibility, and Recency: Effects of News Cues in a Social Bookmarking Website." *Journalism & Mass Communication Quarterly* 90(4):757–75.

Yang, J. H., and H. J. Paek. 2018. "A Content Analysis of 10-year Media Coverage on Risk-Related Rumors." *Korean Journal of Journalism & Communication Studies* 62(3):345–82.

Yoon, K., C. H. Kim, and M-S. Kim. 1998. "A Cross-Cultural Comparison of the Effects of Source Credibility on Attitudes and Behavioral Intentions." *Mass Communication and Society* 1(3–4):153–73.

Yu, W., F. Shen, and C. Min. 2022. "Correcting Science Misinformation in an Authoritarian Country: An Experiment from China." *Telematics and Informatics* 66:101749.

Zeng, J., C. H. Chan, and K. W. Fu. 2017. "How Social Media Construct 'Truth' Around Crisis Events: Weibo's Rumor Management Strategies After the 2015 Tianjin Blasts." *Policy & Internet* 9(3):297–320.

Zhao, D. 2004. *The Power of Tiananmen: State-Society Relations and the 1989 Beijing Student Movement*. Chicago, IL: University of Chicago Press.

Zheng, X., S. Wu, and D. Nie. 2021. "Online Health Misinformation and Corrective Messages in China: A Comparison of Message Features." *Communication Studies* 72(3):474–89.

Author Biographies

Bin Chen is an assistant professor at the Journalism and Media Studies Centre, the University of Hong Kong. His research interests include digital platform and politics, international and comparative media studies, and computational social science. His work has appeared in the *Journal of Communication, Social Media + Society*, and *International Journal of Communication*.

Shuning Lu is an assistant professor in the Department of Communication at North Dakota State University. Her research interests revolve around digital journalism, political communication, and media psychology. Her work has appeared in *Communication Research, Digital Journalism*, and *News Media & Society*, among others.