



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Factors influencing fake news rebuttal acceptance during the COVID-19 pandemic and the moderating effect of cognitive ability

Xin Wang, Fan Chao, Guang Yu^{*}, Kaihang Zhang

School of Management, Harbin Institute of Technology, Harbin, China

ARTICLE INFO

Keywords:

COVID-19
Fake news
Elaboration likelihood model
Rebuttal acceptance
Cognitive ability
Social media

ABSTRACT

Fake news is spreading rapidly on social media and poses a serious threat to the COVID-19 outbreak response. This study thus aims to reveal the factors influencing the acceptance of fake news rebuttals on Sina Weibo. Drawing on the elaboration likelihood model (ELM), we used text mining and the econometrics method to investigate the relationships among the central route (rebuttal's information readability and argument quality), peripheral route (rebuttal's source credibility, including authority and influence), and rebuttal acceptance, as well as the moderating effect of receiver's cognitive ability on these relationships. Our findings suggest that source authority had a negative effect on rebuttal acceptance, while source influence had a positive effect. Second, both information readability and argument quality had positive effects on rebuttal acceptance. In addition, individuals with low cognitive abilities relied more on source credibility and argument quality to accept rebuttals, while individuals with high cognitive abilities relied more on information readability. This study can provide decision support for practitioners to establish more effective fake news rebuttal strategies; it is especially valuable to reduce the negative impact of fake news related to major public health emergencies and safeguard the implementation of anti-epidemic strategies.

1. Introduction

The global outbreak of Corona Virus Disease 2019 (COVID-19) in the age of social media has caused millions of people to contract the virus, and a significant number of them have lost their lives, resulting in a tremendous social and economic shock across the globe (Erku et al., 2021). Amid the growing burden of the COVID-19 pandemic, the parallel emergencies infodemic (information + epidemic) must be simultaneously tackled: the proliferation of fake news, false rumors, and misinformation surrounding COVID-19 (Gallotti et al., 2020; World Health Organisation, 2020a). At its worst, some health-related fake news has introduced people to ineffective or even potentially harmful remedies, which seriously disrupt the social order. For example, the fake news "The Chinese patent medicine Shuanghuanglian oral solution can prevent COVID-19" led to an overreaction as people rushed to buy the product, indirectly bringing about aggregation behavior. Moreover, the fake news "Smoking and drinking can kill COVID-19" led people to engage in risky behaviors that endangered their health (Song et al., 2021). Thus, the World Health Organization (WHO) warned the public that "we're not just fighting an epidemic; we're fighting an infodemic.

Fake news spreads faster and more easily than this virus, and is just as dangerous" (World Health Organisation, 2020b). Studies have shown that when fake news spreads as online rumors on social media, its consequences are amplified (Pal et al., 2020). An important reason is that what used to be propagated largely by word of mouth can now become viral in a short period of time with simple clicks (Chua & Banerjee, 2018). Therefore, how to effectively refute fake news on social media has become an important issue that has never been seen.

To curtail the effect of fake news, most researchers examine methods that might be effective in refuting fake news. There are two main types of methods (Li et al., 2021): blocking fake news by identifying influential users (Ji et al., 2014; Zhang et al., 2018) and a common approach that organizations adopt, which is to use fake news rebuttals (henceforth referred to as "rebuttals") to spread truths, a term referring to messages that debunk and clarify fake news (Pal et al., 2020; Wang & Qian, 2021). Notably, using rebuttals has better long-term performance than blocking, as the openness of the Internet has made it increasingly difficult to block fake news, and the more the fake news is blocked, the more concerned and skeptical the crowd becomes (Li et al., 2021; Wen et al., 2014). Therefore, on social media, effective methods increasingly tend to use fake news rebuttals.

^{*} Corresponding author.

E-mail addresses: wxrabbit@126.com (X. Wang), 2693231135@qq.com (F. Chao), yug@hit.edu.cn (G. Yu), 18800417216@163.com (K. Zhang).

<https://doi.org/10.1016/j.chb.2021.107174>

Received 28 June 2021; Received in revised form 23 December 2021; Accepted 29 December 2021

Available online 31 December 2021

0747-5632/© 2022 Elsevier Ltd. All rights reserved.

Nomenclature

COVID-19	Corona Virus Disease 2019
ELM	Elaboration Likelihood Model
LDA	Latent Dirichlet Allocation
WHO	World Health Organization
SDQC	Supporting, Denying, Querying, or Commenting
BERT	Bidirectional Encoder Representations from Transformers
RoBERTa-wwm-ext	Robustly Optimized Bidirectional Encoder Representations from Transformers Approach-Whole Word Masking-Extended Data
RBT3	Three-Layer Robustly Optimized Bidirectional Encoder Representations from Transformers Approach-Whole Word Masking-Extended Data
CCCCP	Central Committee of the Communist Party of China

The limited studies on the effectiveness of rebuttals have found mixed results. Some studies have aimed to evaluate the effectiveness of fake news rebuttal. For example, Li et al. (2021) developed a proper index to measure refutation effectiveness by identifying key factors through text characteristics, including content and contextual factors. Wang et al. (2021) explored the attitude-based echo chamber effect in users' response to rumor rebuttal for multiple topics on Sina Weibo in the early stage of the COVID-19 epidemic. Furthermore, some studies have shown that rebuttals help reduce people's belief in fake news (Bordia et al., 2000; Huang, 2017). These messages enhance individuals' critical thinking ability (Tanaka et al., 2013). Meanwhile, believing rebuttals helps correct individuals' misconceptions and reduce their anxiety (Bordia et al., 2005). However, other studies suggest that rebuttals are not always effective (DiFonzo et al., 1994; Nyhan & Reifler, 2010). Some psychology studies have shown that correction of fake news via the presentation of true information is not only helpful in reducing misperceptions but sometimes may even increase misperceptions (Nyhan & Reifler, 2010). In some cases, these rebuttals can backfire and, ironically, increase misbelief (Lewandowsky et al., 2012; Nyhan & Reifler, 2010).

In addition to the mixed findings presented above, two research gaps can be identified in the related literature. First, there is a lack of research on the acceptance of fake news rebuttals during public health emergencies, and which properties of rebuttals may work better as well as how they do so have not been extensively investigated. Previous research on the factors that influence fake news rebuttal has focused on the characteristics of the source subjects (Bordia et al., 2005; DiFonzo & Bordia, 2000), content (Bordia et al., 2005), and dissemination channels (Jang et al., 2019), which are primarily based on qualitative or empirical studies (e.g., surveys or questionnaires). At the same time, the advent of the big data era has brought a large and growing volume of data generated from numerous sources; however, the big data generated by social media have not been well utilized, and few studies have been conducted on enhancing fake news refutation effectiveness (Li et al., 2021). As statements of truth, effective rebuttals are essentially a powerful, persuasive message. However, previous research has neglected the factors affecting the routes of persuasive strategies for fake news rebuttal acceptance. One promising theoretical framework using which we can understand fake news rebuttal acceptance is the elaboration likelihood model (ELM) of persuasive theory. The ELM is well constructed and clearly and simply articulates the persuasion process (Kitchen et al., 2014). In addition, the model is so descriptive that it can accommodate several different outcomes and hence can be used to support many situations (Kitchen et al., 2014). The ELM argues that information processing can act via a central or peripheral route (Petty & Cacioppo, 1986, pp. 1–24), which has been applied to detect fake news

in recent studies (Osatuyi & Hughes, 2018; Singh et al., 2020; Zhao et al., 2021), such as Zhao et al. (2021) drew on the ELM to reveal the effective features to detect misinformation in online health communities. But there has been limited research on the factors influencing the rebuttal acceptance of fake news based on the ELM. Therefore, based on the abundant research data available on social media platforms and drawing from the literature on the ELM, this study argues that peripheral cues (source credibility, including authority and influence) and central clues (information readability and argument quality) may influence the acceptance of fake news rebuttal on social media in China.

Second, to some extent, previous rebuttal research has ignored the individual differences in information receivers' responses to fake news rebuttal, particularly differences in psychological factors, such as individuals' cognitive abilities. Recent studies have begun to focus on the psychological factors that may make some individuals less likely to fall for fake news, such as emotional intelligence (Preston et al., 2021) and critical thinking ability (Tanaka et al., 2013). However, although the number of rebuttals has also risen in response to fake news, it appears that corrections to incorrect information only work on some individuals. In addition, previous research has noted that without a nuanced understanding of the cognitive effects of fake news on individuals, the array of possibilities against which society must guard itself becomes insuperable (Bastick, 2021). Thus, to successfully correct users' cognitive errors, rebuttals must convince them (Sui & Zhang, 2021). The ELM posits that the ability and motivation to process the information determine which route individuals will employ—central or peripheral—and those individual attributes determine the relative effectiveness of these processes (Angst & Agarwal, 2009). For example, Bordia et al. (2005) test the effects of central cues (denial message quality) and peripheral cues (source credibility), as well as the moderating effects of personal relevance. Prior works also suggest that ability, as used in the ELM literature, is conceptualized as the cognitive ability of a recipient to process the information presented in the message; it is thus an important component of the information processing act (Angst & Agarwal, 2009). Cognitive ability refers to the ability of the human brain to process, store, and extract information (Lubinski, 2009). A previous study investigates the role of political interest, cognitive ability, and social network size in fake information sharing and indicates that those with lower cognitive ability are more likely to share fake information inadvertently (Ahmed, 2021). However, based on the central and peripheral routes in the ELM, how individuals with different cognitive abilities make judgments and further accept fake news rebuttals on social media has not been sufficiently investigated hitherto, particularly regarding the COVID-19 pandemic.

For these reasons, this study poses four research questions (RQs):

RQ1. How is the source credibility (authority and influence) of rebuttals related to rebuttal acceptance?

RQ2. How is the information readability of rebuttals related to rebuttal acceptance?

RQ3. How is the argument quality of rebuttals related to rebuttal acceptance?

RQ4. How are the relationships between source credibility (authority and influence) and rebuttal acceptance, and between information readability and argument quality and rebuttal acceptance moderated by an individual's cognitive ability?

Thus, to resolve the objectives, we collected data on five fake news events related to the COVID-19 pandemic from Sina Weibo in China from January 20, 2020, to June 28, 2020, and investigated rebuttal postings and their comments to reveal the underlying mechanisms of user's rebuttal acceptance. We used deep learning algorithms, the Latent Dirichlet Allocation (LDA) topic model, and the econometrics method to investigate the relationships among rebuttal's source credibility (authority and influence), information readability, argument quality, and rebuttal acceptance, as well as the moderating effect of user's cognitive ability on these relationships. First, we found that source credibility influenced rebuttal acceptance. Specifically, source authority had a

negative effect on rebuttal acceptance, while source influence could enhance rebuttal acceptance; the positive effect was slightly stronger. Second, we found that both the information readability and argument quality had a positive effect on rebuttal acceptance. More insightful was the discovery that the effect of argument quality on rebuttal acceptance was more significant than that of information readability. Third, we also found that higher cognitive ability weakened the links among source authority, source influence, argument quality, and rebuttal acceptance, while they strengthened the link between information readability and rebuttal acceptance. Thus, we found that individuals with low cognitive abilities relied more heavily on source credibility in peripheral cues and argument quality in central cues to accept rebuttals, while individuals with high cognitive abilities relied more on information readability in central cues to accept rebuttals.

This study is significant in both theory and practice. Theoretically, this study provides two insights. First, this research enriches the literature related to rebuttal by revealing the underlying mechanisms of rebuttal acceptance on social media, particularly on the Sina Weibo platform in China. To the best of our knowledge, this is one of the first studies to explore the underlying mechanism of rebuttal acceptance during the COVID-19 pandemic from a large volume of text information based on text mining (e.g., deep learning algorithms and LDA topic model) and econometric analysis technology on Chinese social media platforms. Second, the majority of applied research on ELM has been conducted in the fields of social psychology and marketing (e.g., advertising and consumer behavior) to describe how individuals process information (Kitchen et al., 2014; Morris et al., 2005). However, using the ELM theory, we propose that source credibility (authority and influence), information readability, and argument quality are factors that affect the acceptance of fake news rebuttals. This study extends the applicability of the ELM by measuring factors from a text-mining perspective and identifies pathways through which individuals with different cognitive abilities accept fake news rebuttal, which enriches the application of the ELM theory in the field of fake news rebuttal on social media.

Furthermore, this study has two practical implications. First, it highlights the source credibility (authority and influence), information readability, and argument quality features of rebuttals and shows how rebuttals can be crafted effectively for refuting fake news on social media. Our findings offer insights for news outlet professionals, social platform managers, and Chinese government regulators regarding the use of rebuttals in combating online fake news during the COVID-19 pandemic. Second, this study shows the extent to which individuals with different cognitive abilities accept fake news rebuttals in terms of source credibility, information readability, and argument quality. These results provide guidance and a reference point for fake news rebuttal practitioners to provide refuting strategies targeted at individuals with different cognitive abilities.

The remainder of this paper is organized as follows. **Section 2** presents a brief overview of the prior research and its theoretical foundation. The development of the research hypotheses and the proposed model are presented in **Section 3**. The research methodology used in our model is explained in **Section 4**. The data analysis and results are presented in **Section 5**. Finally, in **Section 6**, key findings are discussed, and the paper concludes with theoretical contributions, practical implications, limitations, and future directions.

2. Theoretical background

2.1. Rumors, fake news, and rebuttals

There are numerous terms related to false information on social media; fake news, misinformation, disinformation, and rumors are the four most prominent examples (Jung et al., 2020). Regarding the definition and connotation of fake news, mis- and disinformation and rumors, scholars have different understandings, which are often used

interchangeably in academic research (Wang et al., 2018; R. Wang, He, et al., 2020). Allport and Postman (1946) provided an early definition, stating that a rumor is “a specific (or topical) proposition for belief, passed along from person to person, usually by word-of-mouth, without secure standards of evidence being present”. Later, rumors were defined as information spread widely in uncertain or dangerous situations to alleviate fear and anxiety (DiFonzo & Bordia, 2007). In the Chinese context, the fake news problem often refers to online rumors (*yao yan* in Chinese), a type of unverified information statement widely distributed online (DiFonzo & Bordia, 2007). Similar to the online spread of rumors in Western democracies, rumors concerning politics, health care, food safety, and the environment are rampant on China’s Internet (Guo, 2020). Zhang and Ghorbani (2020) state that “fake news refers to all kinds of false stories or news that are mainly published and distributed on the Internet, in order to purposely mislead, befool or lure readers for financial, political or other gains”.

Therefore, in this paper, owing to its useful scientific meaning and construction, we have retained the term “fake news” to represent our research object, which refers to false news or false rumors that authorities’ (e.g., government agencies, state media, and other authoritative organizations) statements determine to be false; the term has been used by most researchers (Grinberg et al., 2019; Lazer et al., 2018). Notably, first, we decided to disregard the politicized nature the term has developed, especially since the 2016 U.S. presidential elections. Second, we adopt a broad definition of news: any information (text, emoticons, links, etc.) posted on social media (Vosoughi et al., 2018). Third, our study does not consider the intentions behind the posters of information, such as their political, economic, or other interests (Zhang & Ghorbani, 2020), or the differences between social robots and people. This is because it has been confirmed that humans, not robots, are more likely to spread fake news (Vosoughi et al., 2018), which implies that containment policies should emphasize human behavioral interventions rather than interventions targeted at automated robots. In summary, we decided to use the term “fake news” in this study.

Most scholarly attention has focused on fake news on social media in terms of theoretical models of fake news diffusion (Tambuscio et al., 2015), simulation models of fake news propagation (Hui et al., 2020; Shrivastava et al., 2020; King et al., 2021), methods for fake news detection (Ozbay & Alatas, 2020; Zhang & Ghorbani, 2020), credibility evaluation (Ciampaglia et al., 2015; Sui & Zhang, 2021), and interventions to curtail the spread of fake news (Hui et al., 2020; Pennycook et al., 2020). Rebuttals have been a subject of growing scholarly attention as a possible approach to mitigating online fake news (Pal et al., 2020). Some studies have suggested that rebutting fake news is effective for the acceptance of rebuttals (Li et al., 2021), while others have suggested that rebuttals may not always achieve their intended effect (DiFonzo et al., 1994) and may even bring about backfire effects that reinforce the belief in fake news (Nyhan & Reifler, 2010). Due to confirmation bias from individuals’ preexisting beliefs (Nickerson, 1998), rebuttals only work on some individuals, which is related to the level of individuals’ cognitive ability.

Taken together, these mixed findings suggest that further research on the underlying mechanism of rebuttal acceptance is warranted. In addition, given that the purpose of rebuttals is to refute fake news, the two antithetical messages often create complex dynamics during fake news refutation (Pal et al., 2020). However, this research focuses on exploring factors that influence fake news rebuttals, and whether acceptance of fake news rebuttal depends on an individual’s level of cognitive ability. Thus, this research focuses on enhancing the competitive advantage of rebuttals by considering how to improve rebuttal acceptance, focusing on the exploration of factors that improve the effectiveness of rebuttal strategies to better guide and suggest fake news rebuttals in today’s era of post-factual truths.

2.2. The elaboration likelihood model and rebuttal acceptance

The rebuttal of fake news belongs to the category of persuasion information, which aims to correct the public's misconceptions and provide knowledge about the truth of matters (Sui & Zhang, 2021). The ELM is among the most authoritative dual-process theories and describes the change in attitudes in the field of knowledge persuasion (Petty & Cacioppo, 1986, 2012), with "attitude" defined as an evaluation of a target object, person, behavior, or event on a scale reflecting some degree of good or bad, such as favor/disfavor, approval/disapproval, liking/disliking, or similar reactions (Eagly & Chaiken, 1993; Griffith et al., 2018). The ELM posits that the ability and motivation to process information determine which route individuals will follow: central routes or peripheral routes, which differ in the amount of thoughtful information processing or "elaboration" demanded of the information receiver (Bhattacharjee & Sanford, 2006; Petty & Cacioppo, 1986, 2012).

Petty and Cacioppo (1986) define elaboration as "the extent to which a person thinks about the content (e.g., issue-relevant arguments) contained in a message." The ELM states that in different situations, different message recipients will vary in the extent to which they cognitively elaborate on a particular message, and these variations in the likelihood of elaboration affect the success of an influence attempt, along with other factors (Sussman & Siegal, 2003). Because of the cognitive effort involved, receivers do not elaborate on every message that they receive, and some receivers elaborate on fewer messages than others (Sussman & Siegal, 2003). According to the ELM, informational influence can occur at any degree of receiver elaboration, but it occurs as a result of very different influence processes: high levels of elaboration represent a central route to influence, while low levels result in a peripheral route (Sussman & Siegal, 2003).

Some researchers have used argument quality and source credibility to represent the central and peripheral routes in the ELM (Cheung et al., 2008; Stephenson et al., 2001; Sussman & Siegal, 2003). Mak et al. (1997) proposed that source credibility has been regarded as a major peripheral cue, while argument strength is a critical factor for central route cues. Zha et al. (2018) considered information quality to be the central route and source credibility and reputation to be the peripheral route. Central and peripheral routes are operationalized using source credibility and argument quality, respectively, and do not consider whether source credibility was manipulated before participants read the argument (Stephenson et al., 2001); these appear to be two of the most frequently referenced constructs (Li, 2013; Rucker & Petty, 2006). Source credibility is defined as a person who provides expertise or whom information recipients perceive to be believable, competent, and trustworthy to deliver persuasive messages (Petty & Cacioppo, 1986, pp. 1–24). Argument quality indicates argument strength, which refers to a message that provides other messages embedded with strong, persuasive arguments (Eagly & Chaiken, 1993; Li, 2013). In addition, the ELM states that elaboration involves attending to the content of the message and scrutinizing and assessing its content (Sussman & Siegal, 2003). Meanwhile, in the online context, the central factors primarily consist of text-related attributes (Srivastava & Kalro, 2019). When the message is processed via the central route, in addition to the argument itself, the receivers evaluate the information embedded within the text (Srivastava & Kalro, 2019). Online text is less likely to be narrated in a "standard" form; rather, it is narrated in a form with a certain level of informativeness descriptions (e.g., description of fake news events) as well as in persuasive language with expressions (e.g., persuasive evidence to refute fake news; K.Z. Zhang, Zhao, et al., 2014). The online text is more or less unstructured, thus presenting challenges for people who read and interpret it (Cao et al., 2011). In particular, fake news and the rebuttal messages related to COVID-19 involve newer fields and specialized vocabulary, as previous studies have shown that online information about COVID-19 is difficult for the general population to read and comprehend (Szmuda et al., 2020). Thus, for the general public, the readability of this

online information has an important impact on the understanding of disease (Worrall et al., 2020). Extant literature indicates that readability refers to the level of effort required to understand a text (DuBay, 2004), and it is a textual property that can be considered the central route (Agnihotri & Bhattacharya, 2016; Srivastava & Kalro, 2019). Previous results have validated the usefulness of the readability feature extracted from the linguistic model for fake news detection on social media (Choudhary & Arora, 2021). Studies have also validated it as an important influential factor characterizing the writing style related to fake news in empirical research (Barrón-Cedeno et al., 2019; Deng et al., 2021). Since the Internet is one of the most popular sources of information in recent times, it is crucial that people are provided with easily understandable information. Health-related news especially, being difficult to read and understand may cause spread of misinformation (Daraz et al., 2018), and public panic resulting from lack of accessible and credible information can pose a greater burden on a country's healthcare system (Szmuda et al., 2020). Therefore, in the online fake news rebuttal field during the COVID-19 pandemic, information readability is an important factor in the central route (Agnihotri & Bhattacharya, 2016; Srivastava & Kalro, 2019; Zhou et al., 2021).

Therefore, to simplify the research model, this study operationalized the central route (argument quality and information readability) and peripheral route (source credibility) constructs as two differential influencing routes to fake news rebuttal acceptance.

Another important aspect of the ELM is the study of the moderating role of "elaboration likelihood," which relates to an individual's abilities and motivations (Shahab et al., 2021). In the ELM, motivation refers to an individual's desire to exert a high level of mental effort, and capacity refers to an individual's ability and opportunity to think (Griffith et al., 2018). When individuals are encouraged by various factors to have the motivation and/or the ability to process arguments thoughtfully and carefully, elaboration likelihood is on the high end, and the central route to persuasion becomes salient (Zha et al., 2018). When elaboration anchors at the low end, the peripheral route to persuasion becomes salient (Zha et al., 2018).

The ELM posits that to change someone's understanding and attitude, the receiver must be motivated to consider the message and must have the ability to process (i.e., devote cognitive effort to) the message (Robert & Dennis, 2005). When using social media to communicate, the receiver must commit some portion of his or her attention to the message and will, therefore, be more likely to be motivated to hear it, and the attention and motivation required on the part of the receiver, which is the first step toward elaboration in the ELM. If individuals are motivated, they are inclined to begin the elaboration process, and the degree to which they can elaborate will depend on their ability to process the message received (Robert & Dennis, 2005). Hence, the next hurdle to elaboration is to ensure that the receiver has the ability to process the information, the final step toward elaboration (Robert & Dennis, 2005). Therefore, scholars posit that the influence of motivation in determining information processing routes may be limited as users are already consciously engaging in deeper cognitive activity when they actively go online to search for specific details (Goh & Chi, 2017). Therefore, in this study, we focus more on the moderating role of cognitive ability.

Prior work has also suggested that ability, as used in ELM literature, is conceptualized as the cognitive ability of a recipient to process the information presented in the message and is an important component of the act of information processing (Angst & Agarwal, 2009). The ELM explains why a given influence process may lead to different outcomes (Li, 2013). Most theories focus on either higher or lower cognitive processing separately, but the ELM is a unified model that deals with both aspects simultaneously (Shahab et al., 2021); that is, information processing can act through both a central and a peripheral route simultaneously (Eagly & Chaiken, 1993; Sussman & Siegal, 2003). Both routes signify that one's attitudes take shape or vary according to intrinsic information processing capabilities—that is, they can evaluate information based on their past experiences and knowledge (Sussman &

Siegal, 2003). When individuals have sufficient cognitive abilities to process information, they accept the persuasion of central routes, such as paying more attention to the argument and message; in contrast, in the peripheral route, individuals exert little cognitive abilities and rely on simple cues to evaluate information, such as focusing on source credibility (Petty & Cacioppo, 1986, 2012). Thus, these perspectives suggest that there are two routes of information processing based on the level of an individual’s cognitive ability.

Accordingly, this study identifies two modes of information processing: central route processing, which requires higher cognitive ability to make judgments regarding related information, and peripheral route processing, which requires lower cognitive ability to make judgments involving the use of simple rules or cues (Petty & Cacioppo, 1986, 2012; Yu, 2021). The most frequently used dependent variable for the ELM is attitude (Bhattacharjee & Sanford, 2006; Chang et al., 2015). In this study, we consider using rebuttal acceptance as a measure of users’ acceptance of rebuttals when COVID-19-related fake news rebuttals have been published. Thus, this study considers both routes to explain how individuals with different cognitive abilities accept fake news rebuttals on social media.

With the ELM as its foundation, this study aims to develop and empirically validate a research model of rebuttal acceptance. The ELM was deemed appropriate because it provides a versatile theoretical lens to examine the underlying mechanism of rebuttal acceptance. That is, when individuals with different cognitive abilities encounter rebuttals, the model demonstrates how central or peripheral cue processing may influence individuals’ internal states of information processing and the extent to which they accept fake news rebuttals.

3. Research model and hypotheses

The research model is presented in Fig. 1. This research relies on the theoretical lens of ELM to propose the construction of a research model of the factors affecting fake news rebuttal acceptance from two aspects: the peripheral route and the central route. The central route includes information readability and argument quality; while the peripheral route includes source credibility (authority and influence). To be specific, we constructed a theoretical model for the source credibility, information readability, and argument quality determinants refuting fake news based on the ELM, focusing primarily on the impact of a rebuttal’s source credibility, information readability, and argument quality on rebuttal acceptance. In addition, we also explored whether higher or lower cognitive abilities would have moderating effects on a rebuttal’s source credibility, or information readability, argument quality, and rebuttal acceptance. In this vein, the research model was formed and the

hypotheses are presented in the following.

3.1. Effects of source credibility

Source credibility refers to a message recipient’s perception of the credibility of a message source, reflecting nothing about the message itself (Sussman & Siegal, 2003). Different sets of conceptual dimensions of source credibility have been identified in prior ELM-based research. The source credibility model provides a three-dimension conceptual framework to operationalize source credibility that influences recipients’ acceptance of a message; these three dimensions are expertise, trustworthiness, and attractiveness (Ohanian, 1990). Recent research generally suggests that the dimensions of source credibility can be narrowed down to expertise and attractiveness (Zhu et al., 2014).

In an online environment in which communication is computer-mediated, verifying source credibility is not an easy task (Li et al., 2017). On social media, a key element of source credibility is the characteristic of the users of posting (L. Zhang, Peng, et al., 2014). For instance, (L. Zhang, Peng, et al., 2014) used users’ authority to indicate source credibility; Lin et al. (2016) suggested that source credibility is strongly influenced by authority. The ability to identify authority has been widely recognized as a key component of source credibility in early persuasion research (Lin et al., 2016). The concept of authority may vary across cultures and expertise fields; however, in general, people with authority are those who have acquired status through education, experience, expertise, talent, or other means (Cialdini, 2001). For instance, in the public health sphere, expert sources (e.g., government health agencies), news media, and social peers may have the potential to correct fake news about public health crises (van der Meer & Jin, 2020), and they may be regarded as authority figures or organizations, but they possess different levels of authority. Authority can assist in the evaluation of source credibility, particularly in risk and health issues, given the inequality of information flow between content producers and consumers (Lin et al., 2016). Therefore, we adopted authority as one dimension of source credibility in social media. In addition, previous literature has also shown the positive influence of a communicator’s physical attractiveness on message persuasiveness (Zhu et al., 2014). This effect is independent of argument quality and source authority, regardless of whether the message receiver has read the message or whether the communicator is an authority figure or organization (Pallak et al., 1983). In the absence of traditional social ties and face-to-face interactions, social structural information such as the number of on-line fans/followers can emulate physical attractiveness to enhance source credibility and influence users’ attitudes toward fake news rebuttals (Zhu et al., 2014). A user’s number of followers indicates the

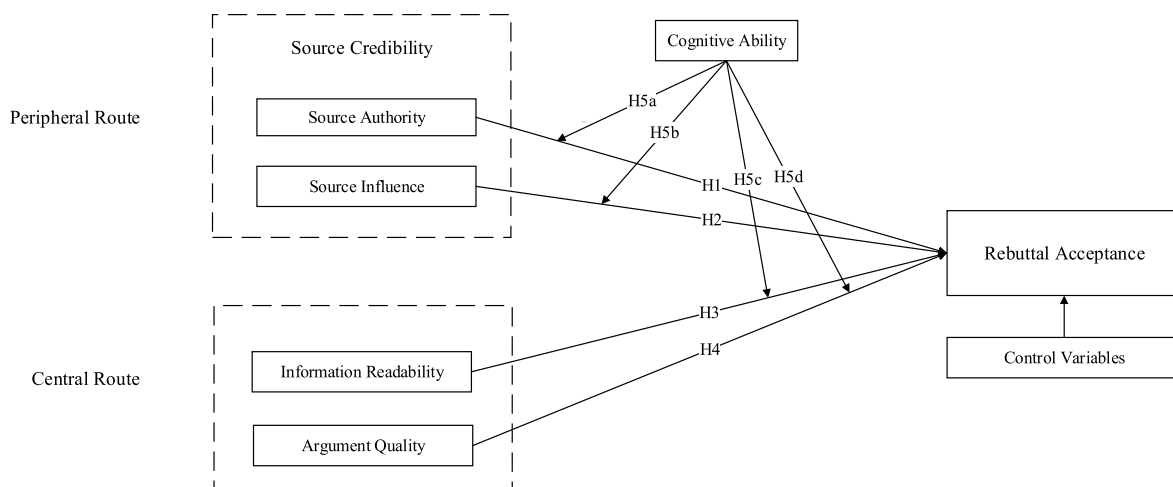


Fig. 1. Research model.

user's influence on Sina Weibo, and users with large numbers of followers are identified as users with great influence (L. Zhang, Peng, et al., 2014). Thus, we adopted influence to indicate the attractiveness of a rebuttal source on social media as one dimension of source credibility. Therefore, in this study context, we measure source credibility through two dimensions: source authority and source influence.

On the one hand, the "authority principle" refers to the tendency of individuals to comply with the recommendations or directives of authority figures, and is a fundamental social influence principle (Cialdini, 2001). In general, the authority principle functions such that the higher the level of authority, the more positive attitudes (e.g., toward fake news rebuttal acceptance) should be (Sui & Zhang, 2021). For instance, previous research suggests that individuals are more influenced by a medical expert or official authority, such as tweets generated by the American Heart Association (Westerman et al., 2014). However, under certain circumstances, authority may exert a negative influence (Jung & Kellaris, 2006). For example, individuals who resent or distrust authority react negatively to authority-based persuasion attempts (Jung & Kellaris, 2006). Previous research has indicated that methods of health persuasion often cause rejection of health messages, a phenomenon known as "reactance" (S. S. Brehm & J. W. Brehm, 2013; Hachaturyan et al., 2021). For example, Lewandowsky et al. (2012) highlighted one cognitive reason that could explain the persistence of erroneous beliefs based on misinformation: because of psychological reactance (S. S. Brehm & J. W. Brehm, 2013), people generally do not like to be told what to think and how to act and may thus reject particularly authoritative fake news rebuttals, especially if they concern complex issues such as public health crises. The COVID-19 infodemic is full of false claims, half-baked conspiracy theories, and pseudoscientific therapies regarding the diagnosis, treatment, prevention, origin, and spread of the disease (Naeem et al., 2021). However, in responding to the COVID-19 pandemic, considering the nature of the pandemic and the effectiveness of measures against the disease (e.g., fake news rebuttals regarding vaccines and preventive measurements), citizens worldwide do not trust their governments or health authorities; in other words, they do not consider the recommendations prescribed by figures in authority or organizations as useful or trustworthy (Houdek et al., 2021, pp. 1–4). In addition, psychological reactance tends to be more pronounced when communicators are perceived as having greater authority (Invernizzi et al., 2003; Pavey et al., 2021). For example, Invernizzi et al. (2003) concluded that when participants with a strong identity as smokers were exposed to a message that targeted smokers, the intention to quit decreased when the message was attributed to a health institute, but increased when it was attributed to a neighborhood association. The neighborhood association (low authority) tended to lead smokers to make a positive change, whereas a health institute (high authority) did not. Similarly, in Gans's (2014) study on HPV vaccination, findings supported the choice of a nonexpert young girl rather than an authoritative expert as the spokesperson to deliver a pro-HPV vaccine message. In this vein, the refuter with low authority bears a greater social risk in making a recommendation and thus may elicit more positive attitudes (Jung & Kellaris, 2006), which suggests that receivers may exhibit a reverse authority effect (Coppola & Girandola, 2018; Invernizzi et al., 2003; Jung et al., 2009), that is, the higher the level of authority of the rebuttal source, the less positive receivers' attitudes toward accepting fake news rebuttals.

On the other hand, as mentioned before, users who have large numbers of followers are identified as users with great influence on social media (L. Zhang, Peng, et al., 2014). A study found that a user's number of followers had significantly positive effects on their number of postings, comments, and posting popularity (L. Zhang, Peng, et al., 2014), and influenced others' perceptions and persuading others to follow their choices (Susarla et al., 2012). The influence of the source can enhance source credibility and positively influence users' attitudes toward fake news rebuttals (Zhu et al., 2014); that is, the recipient's decision-making process will be more affected by the provider if the

rebuttal provider has a high level of influence.

In summary, as a peripheral cue, source credibility may play a crucial role in fake news debunking on social media, and a rebuttal source's authority and influence are important factors that affect the acceptance of fake news rebuttals. Therefore, with respect to the first research question (RQ1), we hypothesize the following:

H1. The authority of a rebuttal source has a negative effect on the acceptance of fake news rebuttals.

H2. The influence of a rebuttal source has a positive effect on the acceptance of fake news rebuttals.

3.2. Effects of information readability

In general terms, the concept of readability describes the effort and educational level required for a person to understand and comprehend a segment of text (DuBay, 2004). The purpose of a readability test is to provide a scale-based indication of how difficult a segment of text is for readers to comprehend based on the linguistic characteristics of that text (Korfiatis et al., 2012). Thus, a readability test can indicate how understandable a piece of rebuttal text is based on its syntactical elements and style (Korfiatis et al., 2012).

Descriptions that are easy to read should be more helpful and influential than other descriptions that are hard to read, and easy-reading text improves comprehension, retention, and reading speed (Han et al., 2018). The extent to which a reader can understand and comprehend the information embedded within a message drives its adoption and effectiveness (Srivastava & Kalro, 2019). Thus, the readability of a message is an important factor that determines how well the receiver is able to understand and adopt the information (Liu & Park, 2015). An online text that is easily readable in terms of content and presentation is deemed more helpful (Srivastava & Kalro, 2019) and draws more utility as its comprehensibility increases (Agnihotri & Bhattacharya, 2016).

In a rebuttal, the provision of information that is easy to read is necessary to facilitate a favorable response in readers to the refutation (Sui & Zhang, 2021), which indicates that providing information with high readability will effectively promote the acceptance of rebuttal messages. Thus, as a central cue, information readability is an important factor (Agnihotri & Bhattacharya, 2016; Srivastava & Kalro, 2019; Zhou et al., 2021) that can affect the acceptance of fake news rebuttals; regarding the second research question (RQ2), we hypothesize the following:

H3. The information readability of a rebuttal has a positive effect on the acceptance of fake news rebuttals.

3.3. Effects of argument quality

Although argument quality is commonly used as a central route in prior literature, its conceptualization and operationalization remain inconsistent among scholars (Angst & Agarwal, 2009). Eagly and Chaiken (2003) defined argument quality as "the strength or plausibility of persuasive argumentation", which was similar to Petty and Cacioppo's definition (Petty & Cacioppo, 1986, pp. 1–24), which indicated that the concept pertains to perceptions of strong and convincing arguments rather than weak and unreal ones. Following this line of definitions, Bhattacharjee and Sanford (2006) operationalized argument quality by emphasizing the strength of arguments, whereas Cheung et al. (2009) employed argument strength to highlight whether received information can persuade a person to believe something or to perform a behavior. Other scholars examined argument quality from a slightly different angle, such as measuring argument quality by examining whether the information is complete, consistent, accurate, or adequate (Sussman & Siegal, 2003) or its relevance, timeliness, accuracy, and comprehensiveness (Cheung et al., 2008). In this study, we argue that argument

quality is the strength of arguments, which refers to the extent to which the argument is convincing or valid in refuting fake news and shows how well a refuting position is justified based on available evidence or a set of reasons (Musuva et al., 2019).

Research demonstrates that argument strength or quality is positively related to attitude change (Stephenson et al., 2001), particularly in online environments (Cheung et al., 2009). The persuasion literature shows that high argument quality has been found to contribute to favorable decision outcomes (Angst & Agarwal, 2009; Srivastava & Kalro, 2019). If the received information is perceived to have valid arguments, that is, if the message contains powerful arguments, good evidence, sound reasoning, and so on, the receiver will develop a positive attitude toward the information; conversely, if the received information appears to have invalid arguments, that is, the message's argument reveals weak evidence, slipshod reasoning, and the like, the receiver will adopt a negative attitude toward the information (Cheung et al., 2009).

If the argument regarding fake news rebuttals is strong, it can generate favorable cognitive responses regarding arguments (Chang et al., 2015). Thus, as a central cue, argument quality may play a crucial role in fake news rebuttal on social media; regarding the third research question (RQ3), we hypothesize the following:

H4. The argument quality of a rebuttal has a positive effect on the acceptance of fake news rebuttals.

3.4. Effects of cognitive ability

The ELM outlines a process-oriented approach to persuasion, and the same variable can induce different attitude changes at different degrees of elaboration (Chen & Ku, 2013). In other words, the informational influence on attitude change is likely to occur at any degree of elaboration, but the influence processes are very different. According to the ELM, the effects of central and peripheral cues on users' decision-making are moderated by the users' ability and motivation to elaborate on the information (Petty & Cacioppo, 1986, pp. 1–24). Many prior studies have tested elaboration likelihood's positive moderating effects on the central route and its negative moderating effects on the peripheral route. For example, Chen and Ku (2013) found that personal relevance and user expertise have positive moderating effects on the central route (argument quality) and negative moderating effects on the peripheral route (source credibility). Zha et al. (2018) found that focused immersion positively moderates the effect of information quality (central route) and negatively moderates the effect of reputation (peripheral route) on informational fit-to-task.

Cognitive ability refers to the ability of the human brain to process, store, and extract information (Lubinski, 2009), including the capability to execute higher cognitive processes of reasoning, remembering, understanding, and problem-solving (Bernstein et al., 2012). As previously mentioned, ability, as used in ELM literature, is conceptualized as the cognitive ability of a subject to process the information presented in the message (Angst & Agarwal, 2009), which refers to the extent to which one possesses prior knowledge or expertise regarding the topic at hand (Bhattacharjee & Sanford, 2006). The current study examines individuals' cognitive abilities regarding fake news rebuttals to understand the different informational influence processes underlying the effectiveness of accepting fake news rebuttals under different levels of elaboration likelihood.

Previous studies have noted that it is also probable that cognitive ability can moderate the relationship between political interest and inadvertent fake information-sharing behavior, and higher news exposure may increase the likelihood of exposure to fake information (Ahmed, 2021). As such, the importance of an individual's information-processing ability is heightened under such conditions. If cognitive ability acts as a buffer against the manipulative nature of fake news, it is plausible that social media users with higher cognitive

abilities of fake news rebuttals would accept such fake news rebuttals. Individuals with higher cognitive abilities are known to make sound judgments when they engage with fake news (Ahmed, 2021); for example, they are skilled at discerning facts from disinformation (Pennycook & Rand, 2019) and adjust their attitudes while engaging with fake news (Roets, 2017). Previous findings also support the thesis that cognitive ability is positively related to efficient information processing (Lodge & Hamill, 1986) and better decision-making (Ahmed, 2021; Gonzalez et al., 2005).

According to ELM, individuals with higher information-related motivations and abilities are more likely to engage in central route processing (Metzger, 2007; Yu, 2021); they may thus be more likely to accept fake news rebuttal through central cues. In contrast, prior work shows that those with lower information-related motivations and abilities are more likely to engage in peripheral route processing (Metzger, 2007; Yu, 2021); they may thus be more likely to accept fake news rebuttal through peripheral cues.

Therefore, it is hypothesized that users with higher cognitive abilities will prefer to make judgments based on central cues, while those with lower cognitive abilities may be more likely to make judgments based on peripheral cues. Thus, with respect to the fourth research question (RQ4), to examine the above associations, we hypothesize the following:

H5a. The effect of source authority on the acceptance of fake news rebuttals will be negatively moderated by individuals' cognitive ability regarding fake news rebuttals such that this effect will be weakened by individuals with higher cognitive abilities.

H5b. The effect of source influence on the acceptance of fake news rebuttals will be negatively moderated by individuals' cognitive ability regarding fake news rebuttals such that this effect will be weakened by individuals with higher cognitive abilities.

H5c. The effect of information readability on the acceptance of fake news rebuttals will be positively moderated by individuals' cognitive ability regarding fake news rebuttals such that this effect will be strengthened by individuals with higher cognitive abilities.

H5d. The effect of argument quality on the acceptance of fake news rebuttals will be positively moderated by individuals' cognitive ability regarding fake news rebuttals such that this effect will be strengthened by individuals with higher cognitive abilities.

4. Methodology

4.1. Data collection and preprocessing

4.1.1. Study context and data collection

Sina Weibo (Weibo, Beijing, China, <http://www.weibo.com/>), often referred to as "Chinese Twitter" (Twitter, San Francisco, USA, <http://www.twitter.com/>), is one of the most influential social network platforms in China (Rodríguez et al., 2020). As in some other microblogging systems (e.g., Twitter), users of Sina Weibo can post short messages, termed "postings." The process of information diffusion on Sina Weibo can be traced by the analysis of original postings, reposting, comments, replies, and likes of postings. These features make this social platform highly interactive and allow for rapid and broad dissemination of information. Furthermore, in contrast to other social networks used in China such as WeChat (WeChat, Shenzhen, China, <http://weixin.qq.com/>), communication on Sina Weibo is almost entirely public, which generates a large amount of unstructured text data and provides data support for our research. Using the Zhiwei Data Sharing Platform (Zhiwei Data, Beijing, China, <http://university.zhiweidata.com/>), a well-known social media data service company in China, we collected data on COVID-19-related fake news events through the Business Application Programming Interface (API) of Sina Weibo, as outlined below.

- (1) We collected original postings (i.e., non-reposted postings) regarding COVID-19-related fake news events. First, based on the fake news event list dating from January 20, 2020, to June 28, 2020, for the Zhiwei Data Sharing Platform, we defined keyword combinations related to fake news events that occurred in China using multiple logical relationships (e.g., AND and OR). Then, we collected fake news event data related to the COVID-19 pandemic dating from January 20, 2020, to June 28, 2020 (these stories had been confirmed as “fake” by authoritative statements) by obtaining the original postings and corresponding comments of each event. Second, to avoid data selection bias and according to the event classification criteria of the Zhiwei Data Sharing Platform, according to the number of comments, we sampled five fake news events that involved three fields: authority, society, and politics. These fake news events are the five fake news stories with the highest number of comments from January 20, 2020, to June 28, 2020, which have the characteristics of the most widespread dissemination and influence on social media platforms in China during the COVID-19 pandemic. The annotation scheme of three fields of fake news was developed through an iterative process of rounds of annotation and evaluation involving three researchers (two Ph.D. students and one expert from Zhiwei Data Sharing Platform, who are experienced in fake news research amid public health crises). Table 1 shows the annotation scheme for the three fields of fake news. Subsequently, after manual data denoising and cleaning by three expert annotators, we randomly selected 2053 postings related to fake news events. For brevity, the descriptions of the five fake news events are shown in Appendix A.
- (2) We also collected all comments that replied to 2053 original postings. From the original postings obtained, as mentioned earlier, we collected comment conversations associated with each original posting. To collect comments, we scraped the webpage of each original posting to retrieve the URL. However, if the URL was missing or invalid, we chose MIDs (a unique identification code for each posting on Sina Weibo) to retrieve the comments. We collected a total of 100,348 comments (either direct replies or nested replies to the original postings) that replied to the 2053 original postings.

Table 2 presents the basic attributes, including the type of event and the number of original postings and comments, of each fake news event. All of the information obtained on the web was written in simplified Chinese and released publicly by Sina Weibo. Because we used publicly available data, we only referred to the summarized results and did not collect any sensitive data.

4.1.2. Filtering of postings for fake news rebuttals

We used manual annotation to filter out postings that refuted fake news. A three-person expert panel of social media researchers also labeled the refuting postings out of the 2053 original postings. Referring

Table 1

The annotation scheme for three fields of fake news.

Category of fake news	Annotation scheme for category of fake news
Authority	Those slandered by the fake news are authoritative individuals or organizations in expertise fields related to COVID-19, such as Nanshan Zhong (Event 1) and the Wuhan Institute of Virology (Event 3).
Society	Fake news regarding social events, social problems, and social style involving people’s daily lives in relation to COVID-19, especially reflecting social morality and ethics.
Politics	The fake news related to COVID-19 regarding politics, i.e., the activities of classes, parties, social groups, and individuals in domestic and international relations.

Table 2

Five fake news events (translated into English from Chinese).

No.	Event	Category	Original Postings	Comments
1	Yansong Bai dialogued with Nanshan Zhong.	Authority	478	11,207
2	Materials of Jiangsu Province aided medical team of Hubei Province were detained.	Society	114	8006
3	A postgraduate from the Wuhan Institute of Virology was the “Patient Zero.”	Authority	676	49,763
4	80 Chinese citizens were abused while being quarantined in Russian Federation.	Politics	258	15,524
5	A car owner in Hubei Province died of COVID-19.	Society	527	15,848

Note: a. The data collection for the postings and comments ended on September 12, 2020, at 23:59:59. b. The main objects affected by fake news in Events 1 and 3 were authoritative individuals and organizations in expertise fields, respectively.

to Tian et al. (2016), the specific steps are as follows:

First, two annotators with a detailed understanding of COVID-19-related fake news independently labeled postings of fake news rebuttals among all 2053 original postings.

Second, to eliminate the differences due to human factors, the two members discussed all the annotation results and re-annotated the postings to reach an agreement on the differences.

Third, the third annotator annotated 2053 original postings to calculate inter-rater reliability. Cohen’s kappa (κ) for the annotators was 0.92 ($p \sim 0.0$), indicating a good agreement among them (Cohen, 1960).

Finally, from the 2053 original postings, we obtained 1721 original postings of labeled fake news rebuttals.

4.1.3. Stance classification

Utilizing the rich information in comments that replied to the original postings on Sina Weibo, we employed SDQC stance classification (Derczynski et al., 2017) to classify the stance in user comments on fake news as one of the following: supporting [S], denying [D], querying [Q], or commenting [C]. However, most prior studies have examined SDQC stances on fake news rather than stances on fake news rebuttals (Zubiaga et al., 2018). To follow the existing criteria for classifying SDQC stances, the four stances in the comments on refuting postings studied in this paper are stances on fake news. The comments of each user were categorized according to the following four stances:

- Supporting (S): Users who commented believing that the fake news is true; that is, they think the fake news rebuttal is false.
- Denying (D): Users who commented believing that the fake news is false; that is, they think the fake news rebuttal is true.
- Querying (Q): Users who commented on the veracity of the fake news while asking for additional evidence.
- Commenting (C): Users who commented without clearly expressing whether they wanted to assess the veracity of the fake news.

To improve classification efficiency and obtain satisfactory results in the case of massive text data, we designed text classifiers with supervised learning methods to automatically identify and classify the four types of SDQC stances observed in the 100,348 comments.

Similarly, first, we asked two members of our team to label 12,000 comments (randomly selected from 100,348 comments based on the proportion of the number of comments in each event) with the corresponding SDQC stances.

Second, the third member randomly selected 2000 comments to validate the inter-annotator agreement, and the annotation process was

assessed and validated using Cohen's kappa ($\kappa = 0.88$; $p \sim 0.0$; Cohen, 1960). All codes were mutually exclusive.

Third, we divided 12,000 comments into training and testing sets in a ratio of 7 : 3; the label distribution is presented in Table 3.

Fourth, we trained and compared the text classifiers using deep learning methods. Because the new language representation model, the bidirectional encoder representations from transformers (BERT; developed by Google in 2018), is conceptually simple, empirically powerful, and obtains new state-of-the-art results on 11 natural language processing tasks without substantial task-specific architectural modifications, we chose BERT and its two related but improved models to train the classifiers (Devlin et al., 2018). These three models are BERT, RoBERTa-wwm-ext (robustly optimized BERT approach-whole word masking-extended data), and RBT3 (three-layer RoBERTa-wwm-ext; Cui et al., 2019; Cui et al., 2020; Liu et al., 2019). We chose accuracy and macro-F1 score to measure the performance of the classifiers (Van Asch, 2013; Zhang et al., 2015). During the fine-tuning process, we compared 72 sets of hyperparameters based on the three models to obtain the best-performing stance classification model. Table 4 presents the performances of these classifiers under different combinations of hyperparameters. For brevity, we have only shown the hyperparameter results with the best performance for each model. The detailed results of the stance classification are described in Appendix B. Finally, we chose the RoBERTa-wwm-ext model with hyperparameters (140, 16, 3e-5, 3) to predict the SDQC stance for large comment datasets, as it performed satisfactorily with an accuracy of 80.89% and a macro-F1 of 68.06%. The results of our stance classifier are presented in Table 5.

4.1.4. Sample selection

To measure the rebuttal acceptance of fake news refuting postings, we randomly selected a subset of users who commented on refuting postings, including 3865 unique users from a total of four types of stances, and their 564,910 background postings, including 4277 comments on 343 fake news refuting postings (they may have made different comments on different refuting postings from different fake news fields).

First, according to the proportion of users in the four stances in Table 5, among the users who commented on 1721 refuting postings, we randomly selected 147 unique users with a supporting stance, 1276 unique users with a denying stance, 1335 unique users with a querying stance, and 1107 unique users with a commenting stance. All of these users participated in the 4277 comments responding to the 343 fake news refuting postings.

Second, to avoid contact with the COVID-19 fake news refuting information in the users' historical experience, we chose all 3865 users' background postings from the 3 months before the outbreak of COVID-19 in early December 2019 (2019.09.01 00:00:00–2019.11.30 23:59:59). We then collected 564,910 background postings in total through the users' UIDs (a unique identification code for each user on Sina Weibo).

As we sampled users who commented on refuting postings, we are approximating what the users may have been exposed to regarding rebuttal postings. Users' commenting behavior represents their behavioral decisions after they are exposed to postings and process those postings (Liu et al., 2012). Fig. 2 shows the distribution of the number of background postings of users in each stance. The process of data collection and preprocessing and the flowchart of the sample selection is shown in Fig. 4.

Table 3

Label distribution of the training and testing sets.

	Stance			
	Support	Deny	Query	Comment
Training set	227	2305	1126	4742
Testing set	98	956	472	2074

Table 4

Stance detection performance of the classifiers on user comments.

Model	Hyperparameters	Accuracy	Macro-precision	Macro-recall	Macro-F1
BERT	(70, 16, 2e-5, 3)	80.33%	67.71%	65.94%	66.51%
RBT3	(140, 32, 5e-5, 3)	80.17%	68.78%	63.08%	64.53%
RoBERTa-wwm-ext	(140, 16, 3e-5, 3)	80.89%	68.76%	67.88%	68.06%

Note: Hyperparameters (x, y, z, w), where x: max_seq_length (70, 140), y: train_batch_size (16, 32), z: learning_rate (2e-5, 3e-5, 5e-5), and w: num_train_epochs (2, 3). The values in parentheses represent the hyperparameters derived from the fine-tuning process.

Table 5

Comparison of the number of users and the number of comments for the SDQC stances.

Stance	Unique				Mixed	Total
	Support	Deny	Query	Comment		
Total No. of users, n	781	19,111	6832	34,694	6370	67,788
Total No. of comments, n	800	20,546	7359	45,337	26,306	100,348

Note: Unique stance means that the user has only expressed one unique stance in their comments, no matter how many times they commented. A mixed stance means that the user commented at least twice and expressed more than one type of stance.

4.2. Operationalization of variables

4.2.1. Independent variables

The independent variables are the rebuttal source's authority (*RSAutho*), the rebuttal source's influence (*RSInflu*), the information readability of the rebuttal information (*RIRead*), and the argument quality of the rebuttal information (*RAQuali*). These were extracted from the text information through text mining and manual scoring.

(1) Measuring the authority of a rebuttal source

To identify the authority of a rebuttal source (*RSAutho*), referring to Chen et al. (2021), we scored the authority of a rebuttal source according to their homepage information and verification status on Sina Weibo as well as other information online (e.g., <http://baike.baidu.com/>), ranging from 1 to 5, where 1 indicated that the rebuttal source is "completely not an authority," and 5 indicates that it is "completely an authority." At the same time, to score the degree of authority more accurately, we used a fuzzy expression that shields the generally adopted Likert scale from rank restriction, allowing experts to freely express their attitudes on a number axis from 1 to 5 based on scoring criteria, marking the point that best indicates the degree of authority and marking the number represented above this point (Pennycook et al., 2020). The size of the number indicates the expert's perceived value of authority, with larger values indicating higher authority of the account. To ensure that the manual scoring process met reliability standards, we also used a three-person expert panel of social media researchers to score the authority of the account (two Ph.D. students and one expert from Zhiwei Data Sharing Platform who are experienced in fake news research amid public health crises). First, the three researchers iteratively discussed and confirmed the preliminary authoritative scoring criteria by drawing on their practical research experience, sample data characteristics, and reading of research materials. Second, the two researchers independently scored the authority of the accounts based on the preliminary scoring criteria formed above, repeatedly revising and discussing the scoring details during the scoring process, which resulted in the final version of the authority scoring criteria; ultimately, they obtained a consistent scoring result for the authority of the accounts.

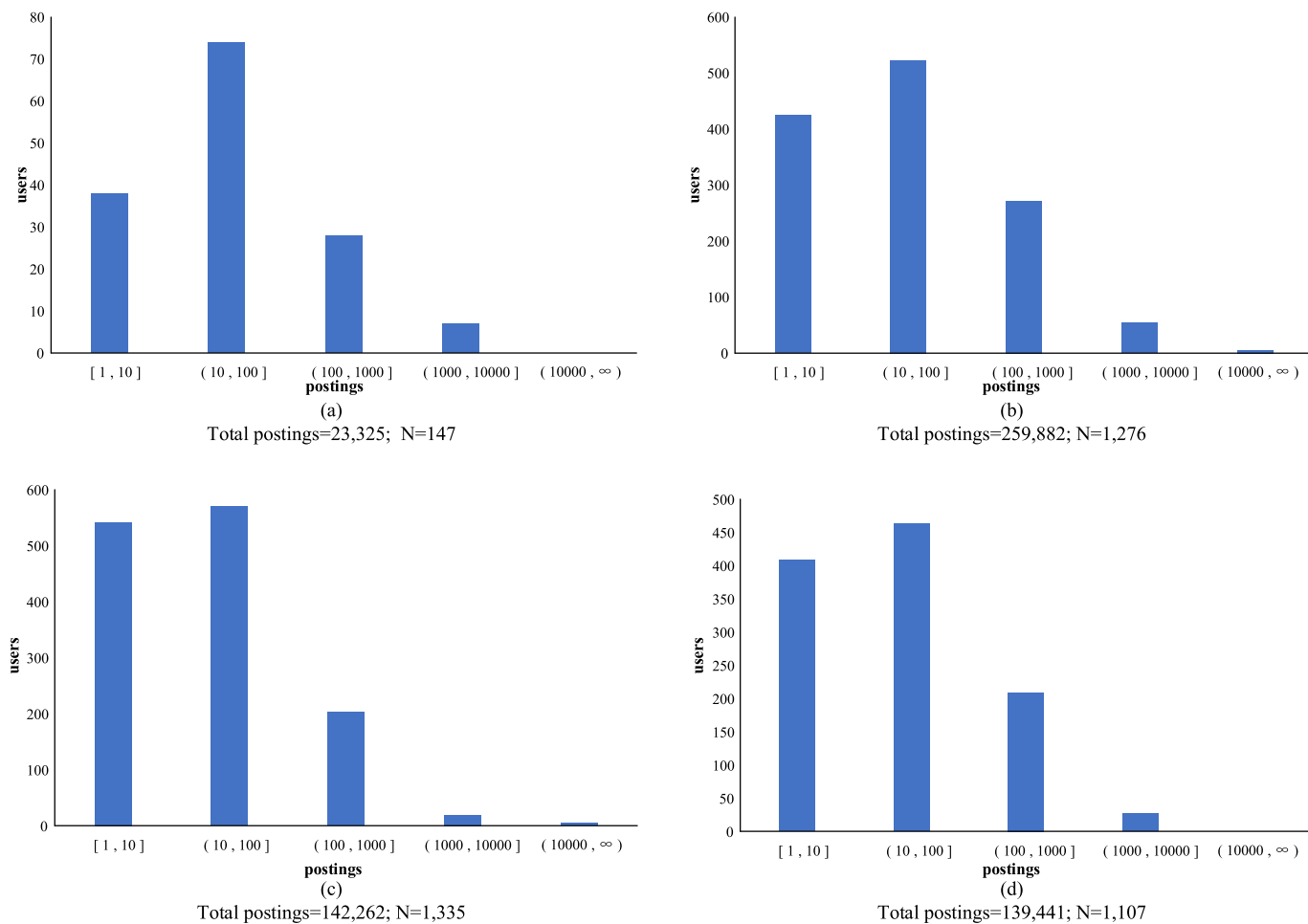


Fig. 2. Distribution of background postings of users with (a) supporting, (b) denying, (c) querying, (d) commenting stances.

Third, the third researcher scored independently based on the final version of the scoring criteria and obtained their scoring results. Fourth, we calculated the consistency of the two scoring results, and the inter-rater agreement was determined to be 0.91 by calculating Cohen’s kappa (κ) ($p \sim 0.0$), which indicated that the scoring process had been validated. The calculation formula is $RSAutho = \sum_{i=1}^n RSAutho_i$. Table 6 shows the final scoring criteria and examples of the results of our scoring of the source authority.

(2) Measuring the influence of a rebuttal source

Users who have large numbers of followers on social media are identified as influential (L. Zhang, Peng, et al., 2014). On Sina Weibo, the number of followers can be used to measure the account’s social media influence (Z. Wang, Liu, et al., 2020). Therefore, the influence of the rebuttal source ($RSInflu$) was measured using the number of followers of the rebuttal source accounts. The calculation formula is $RSInflu = \sum_{i=1}^n RSInflu_i$.

(3) Measuring the information readability of a rebuttal posting

In this study, we used information readability ($RIRead$) to indicate the ease with which a commenting user can understand a rebuttal text (Korfiatis et al., 2012). Readability can be defined as the extent to which a written text is easy to understand, and its main determinant is the complexity of wording (Senter & Smith, 1967). Thus, for all

COVID-19-related fake news rebuttals, the higher the readability of the rebuttal text, the less difficult it is for users to understand the message when the main idea of the message expressed in the rebuttal text is the same (Korfiatis et al., 2012).

In more formal terms, a readability test is a formula produced by applying linear regression to subjects regarding the reading ease of different segments of text that they had been asked to comprehend using specific instruments (Korfiatis et al., 2012). Thus, we measured the information readability of a rebuttal posting ($RIRead$) using the readability index (RI). The calculation of the readability index has been established in English studies (Korfiatis et al., 2012). Among the existing Chinese readability formulas, Jing Xiyu’s proposed readability formula is the most authoritative (see Equation (1)), while the common characters used in Equation (1) defines come from a list of 495 commonly used traditional characters in Taiwan’s national primary school, which is relatively old and not applicable to the classification of common Chinese characters in online social media (Qing & Qin, 2020). Therefore, we chose to cite the 1988 edition of *Modern Chinese Commonly Used Characters* (3500 characters) as a dictionary of commonly used Chinese characters.

$$RI = 17.5255 + 0.0024 \times TW + 0.04415 \times ASL - 18.3344 \times PCW \quad (1)$$

where for any refuting posting containing at least one sentence, TW is the total number of characters (Chinese characters, English letters, and numbers except for spaces and punctuation marks) in a refuting posting, and ASL indicates the average length of a sentence in a refuting posting—that is, the average number of characters in a sentence.

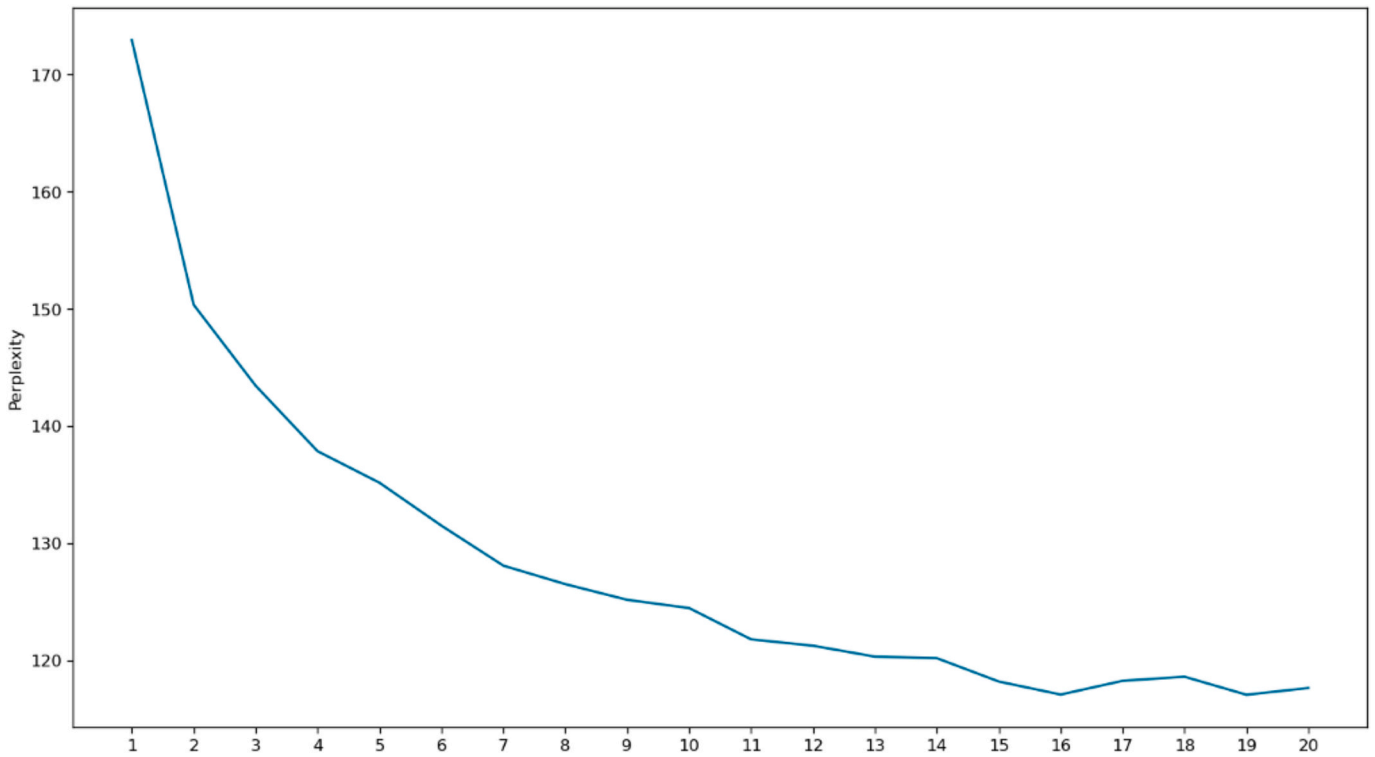


Fig. 3. Perplexity on the dataset depending on the number of topics.

$$ASL = \begin{cases} TW, & Num_other = 1 \\ TW/Num_other, & else \end{cases} \quad (2)$$

where Num_other indicates the number of sentences contained in a refuting posting, and PCW is the percentage of common Chinese characters in a refuting posting.

$$PCW = Count_norm/Num_chi \quad (3)$$

where $Count_norm$ indicates the number of common Chinese characters that appear in a refuting posting, and Num_chi indicates the total number of Chinese characters in a refuting posting.

In Equation (1), a smaller RI value indicates a smaller TW and ASL and a larger PCW , suggesting that the rebuttal text is easier to read; conversely, a larger RI value indicates a larger TW and ASL and a smaller PCW , suggesting that the rebuttal text is less readable (Qing & Qin, 2020). Therefore, to facilitate the analysis of the relationship between information readability and rebuttal acceptance in the regression model, we take the opposite value of the RI (see Equation (1)) to express $RIRead$; a higher score indicates the ease of readability (i.e., the higher the $RIRead$, the easier it is to read the rebuttal text). This was calculated using the following formula:

$$RIRead = -RI = -(17.5255 + 0.0024 \times TW + 0.04415 \times ASL - 18.3344 \times PCW) \quad (4)$$

Similarly, the calculation formula is $RIRead = \sum_{i=1}^n RIRead_i$.

(4) Measuring the argument quality of a rebuttal posting

Similar to the scoring method of $RSAutho$, referring to the previous studies in Section 3.3, which consider the quality of argument in the postings of fake news rebuttals (Bordia et al., 2005; Zubiaga et al., 2018), we also scored for the quality of the argument regarding rebuttal information ($RAQuali$). The score ranged from 0 to 2, showing the strength of the argument contained in the rebuttal information, where 0 indicates that the rebuttal information is “a completely weak argument” and 2 indicates that it is “a completely strong argument.” The higher the score, the higher the argument quality (i.e., the stronger the arguments) provided in the respective rebuttal. The calculation formula is $RAQuali = \sum_{i=1}^n RAQuali_i$. Furthermore, we also conducted interrater agreement reliability tests to ensure intersubjective verifiability (Cohen’s kappa, $\kappa = 0.75$; $p \sim 0.0$). Table 7 shows an example of the results of our scoring for argument quality.

4.2.2. Dependent variable

The dependent variable in this research is rebuttal acceptance ($RAccept$), which indicates the user’s stance reaction when commenting on the refuting posting after it is posted. Each classifier in Section 4.1.3 can output probability values of the four SDQC stances corresponding to each comment text before outputting the stance classification result for

each comment text; for each comment, the sum of the four probability values is 1. Therefore, we used the results from the optimal stance classifier in Section 4.1.3 (value ranges [0,1]) to output the probability value of the fake news denying stance corresponding to each comment text to indicate the degree of acceptance of the fake news rebuttal. The higher the value, the higher the acceptance of the rebuttals by the users.

The calculation formula is $RAccept = \sum_{i=1}^n RAccept_i$.

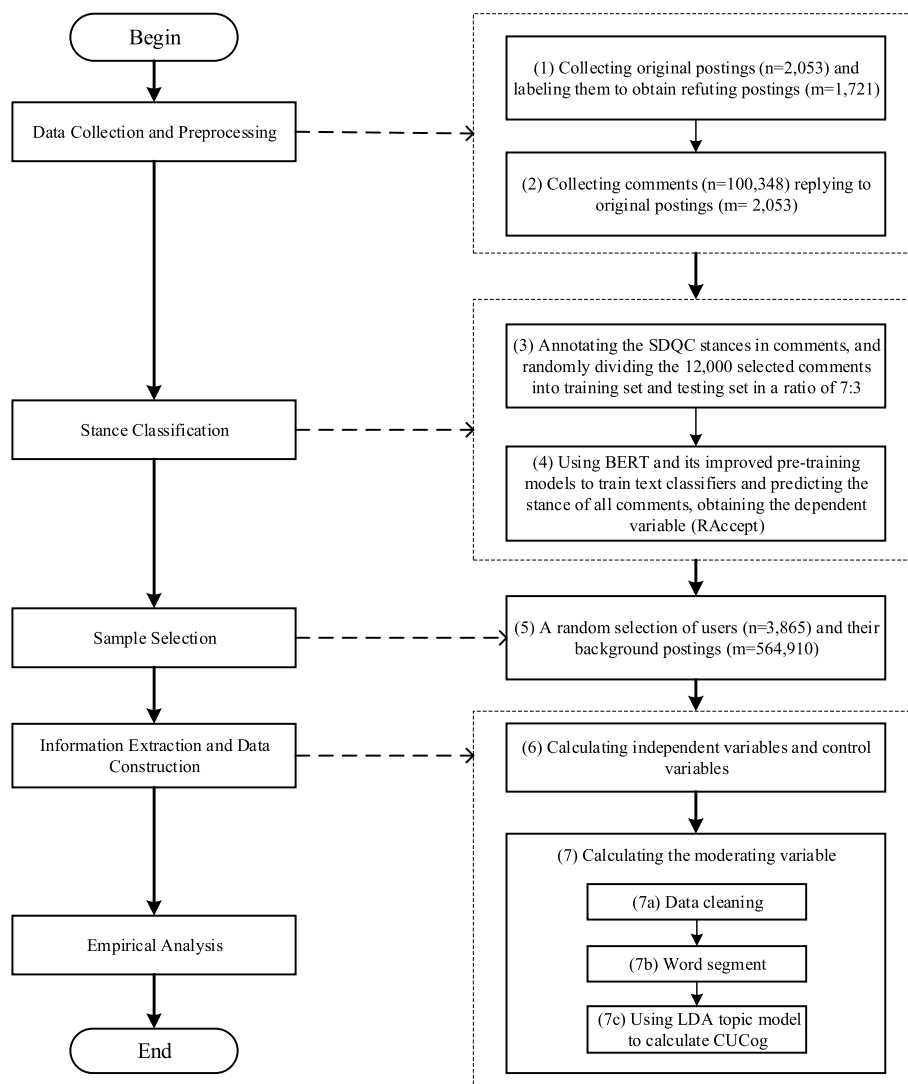


Fig. 4. The process of the research procedures.

4.2.3. Moderator variable

The moderator variable in this research is the commenting users' cognitive ability (*CUCog*).

Petty and Cacioppo (1986) posit that in circumstances in which individuals are motivated and can process information, they will engage in issue-relevant thinking. When this occurs, individuals are likely to attempt to access relevant associations, images, and experiences from memory and analyze and elaborate on the information presented in conjunction with such associations provided by their memory. For example, individuals may also extract inferences concerning the merits of the arguments in support of a recommendation centered on their examination and develop an overall evaluation of or attitude toward the information (Petty & Cacioppo, 1986). As the individual integrates the arguments presented with their previous knowledge, the likelihood of elaboration is high, and there will be considerable support for the allotment of cognitive effort (Petty & Cacioppo, 1986; Robert et al., 2005).

Language is inherently cognitive (Beckage & Colunga, 2016). Cognitive ability is structured around a regnant general factor (general intelligence) and supported by a number of specific factors (e.g., mathematical, spatial, and verbal abilities; Lubinski, 2009). Vocabulary tests share high variance in terms of general intelligence and have frequently been used to assess individuals' cognitive capability (Ahmed, 2021; Brandt & Crawford, 2016). Wechsler (1958) supports the use of a

vocabulary test to measure cognitive ability and argues that "contrary to lay opinion, the size of a man's vocabulary is not only an index of his schooling but also an excellent measure of his general cognitive ability. Its excellence as a test of cognitive ability may stem from the fact that the number of words a man knows is at once a measure of his learning ability, his fund of verbal information, and the general range of his ideas."

As mentioned above, an individual's cognitive ability to process a message depends on many factors, such as prior knowledge of the topic (Metzger, 2007). Context-specific knowledge gained through experience provides a cognitive foundation that enables one to engage in cognitively rigorous elaboration (Petty & Cacioppo, 1986, pp. 1–24). Therefore, we investigate the text similarity between the individuals' past experiences and knowledge and the rebuttal postings that they have commented on to analyze the level of cognitive ability of individuals in relation to fake news rebuttals. The more information (words, topics) related to the rebuttal message in the domain that is contained in the user's historical memory, the more the user can understand the meaning embedded in the rebuttal text, and the higher is their degree of cognitive abilities regarding fake news rebuttals in the domain. To accomplish this, based on previous research (Vosoughi et al., 2018), using an LDA topic model, which is a probabilistic statistical model used to discover the underlying abstract topics in a series of documents or text data (Blei et al., 2003), we calculated the information distance between the

Table 6
Criteria and examples for the process of scoring the authority of a rebuttal source (translated into English from Chinese).

Score	Description	Example
[1,1.5]	Ordinary users who have no authentication information, and the content of their homepage is primarily personal opinions or sharing daily life.	/
(1.5,2]	Personal accounts with authority in certain specific industries, such as social peers in the public health sphere.	/
(2,3]	Non-government-related official media accounts (i.e., official media accounts not affiliated with any government): for example, authoritative media in specific industry sectors (e.g., financial newspapers); portal sites (e.g., Tencent, Sina)	<ul style="list-style-type: none"> ● Sina Finance ● Hongxing News ● Chengdu Business Daily
(3,4]	Official accounts of local governments, emergency services, and political groups and unions: for example, police agencies, local government official accounts (e.g., provincial, municipal governments), media accounts belonging to local governments, and the accounts of their subordinate agencies (e.g., official newspaper [local government-operated]), political groups such as the communist youth league, party groups, and unions.	<ul style="list-style-type: none"> ● Beijing Daily ● Ping An Kaifeng ● Liaoning Communist Youth League
(4,5]	The official media accounts are led by the Central Committee of the Communist Party of China (CCCPC) and the accounts of its subordinate agencies.	<ul style="list-style-type: none"> ● People's Daily ● Xinhua Viewpoint ● Workers' Daily

Note: Personal users are involved in non-public data; thus, no examples of specific user nicknames will be provided.

Table 7
Criteria and examples for scoring the argument quality of a rebuttal (translated into English from Chinese).

Score	Example	Description
[0,0.5]	<ul style="list-style-type: none"> ● Yansong Bai <i>did not</i> talk with Nanshan Zhong. ● 80 Chinese citizens were mistreated in isolation in Russia is <i>not true</i>. 	A flat counterstatement of fake news without providing any valid arguments, i.e., the message's argument reveals weak evidence, slipshod reasoning, and the like, indicating that the strength of the argument is weak.
(0.5,1]	<ul style="list-style-type: none"> ● Refute the rumor! <i>Yansong Bai exclusive reply</i>: tonight I did not dialog Nanshan Zhong. ● <i>Zhengli Shi (A researcher at the Wuhan Institute of Virology)</i>: can guarantee that Wuhan Institute of Virology currently has no infections. 	Refuting fake news with cited evidence and reasons that are not very good, adequate, or powerful, such as using a relevant person or an organization's response, indicating that the strength of the argument is medium.
(1,2]	<ul style="list-style-type: none"> ● <i>Wuhan Institute of Virology</i> can guarantee that it currently has zero infections; <i>Yanling Huang</i> herself responded that she has not returned to Wuhan after graduation and has not been infected. ● <i>Xiaogan</i> (a city in Hubei Province) police debunked the rumor that the owner did not move the car due to death from COVID-19; at the same time, <i>the car owner's company</i> solemnly declared: the car owner is healthy, and the rumors on the online social media are not true. 	Refuting fake news with a combination of evidence and reasons, such as using both a person's and an organization's response; the message contains powerful arguments, good evidence, sound reasoning, and so on, indicating that the strength of the argument is strong.

Note: a. For brevity, we only show the text that can indicate the strength of the rebuttal argument and do not list the full rebuttal postings. b. Keywords concerning the quality of the argument are italicized.

background postings from the comment users and the refuting postings from the rebuttal source.

First, we organized 564,910 background postings corresponding to all 3865 users as well as 343 refuting postings that were commented on by these users into a collection of 4208 (3865 + 343 = 4208) text datasets and used an LDA topic model to train the sample data. In this study, fake news refuting postings and users' background postings were organized into a dataset with IDs corresponding to the texts. Specifically, each rebuttal posting was organized into a unique MID and its corresponding fake news rebuttal text, totaling 343, and each user's background posting data was organized into a unique UID and its corresponding background posting text, totaling 3865.

Second, according to the external reference of the stop words dictionary, we used Python's regular expressions to clean the sample data, including removing the relevant stop words (based on the list of stop words produced by the Harbin Institute of Technology; Zhao et al., 2021), URLs, and punctuation. And we created our user dictionary, including some terms related to COVID-19, and used the Jieba (Jieba Chinese Word Segmentation, China, <http://github.com/fxsjy/jieba/>) in Python (Python 3.6.2, <http://www.python.org/>) to segment the sample data, which is one of the most widely used word segmentation tools for Chinese.

Third, for the clustering analysis of LDA topics, the bag-of-words method was used to form the feature vector corresponding to each text. To reduce the dimensionality of the feature vectors and improve the scientific validity of the experiment, the words that appear in the top 5% and the words that appear only three times were removed, and the top 200 words were selected as features to form a feature vector for each text, resulting in a feature vector matrix of 200 × 4208.

Fourth, a perplexity measurement is a predictive measure of the probabilistic model wherein a low perplexity indicates how good the probability distribution in the sample is (Brown et al., 1992; Principe et al., 2021). Thus, we used a perplexity analysis to estimate the performance of topic clustering. We used the LDA method of the sklearn package in Python to cluster the LDA topics of the above feature matrix, with the number of clusters gradually increasing from 1 to 20 to find the appropriate number of topics, and generated perplexity indexes for each clustering result using the lda_perplexity method. Fig. 3 demonstrates the perplexity of LDA models with varying numbers of topics; the perplexity continues to decrease up to 16 topics. However, perplexity is not the only criterion for deciding on the final number of topics; choosing an appropriate number of topics also depends on domain knowledge and human evaluation of whether the words associated with each topic make structural sense (De Clercq, 2019). Based on these considerations, this study selected 16 as the appropriate number of topics (despite the same value of perplexity for 19 topics); this value had a reasonable value for perplexity while retaining topic interpretability. Finally, we generated the probability distributions of the 16 topics for each text.

Fifth, for each user, we compared the distribution of the 16 topics in the rebuttal posting that they commented on as well as that of the 16 topics in the background postings to which they were exposed during the 3 months (2019.09.01 00:00:00–2019.11.30 23:59:59) prior to the outbreak of the COVID-19 pandemic. Vosoughi et al. (2018) used Bhattacharyya distance (BD; Kailath, 1967) and KL divergence (KL; Cover, 1999; Kullback & Leibler, 1951) in their study to compare the information distance between a rumor and users' background tweets based on LDA topic distribution (Vosoughi et al., 2018), which are two suitable similarity measures for topics (Andrei & Arandjelović, 2016). BD measures the dissimilarity between two probability distributions (Kailath, 1967). Similarly, KL measures how one probability distribution diverges from a second expected probability distribution (Kullback & Leibler, 1951). In our case, the two probability distributions correspond to the topic distributions of the rebuttal posting and the background postings. The higher the BD or KL, the greater the distance between the topic distributions in the two documents (one a rebuttal posting that a user commented on and the other a user's background postings;

Vosoughi et al., 2018). To calculate the level of cognitive ability of the individual in relation to fake news rebuttals, we used the opposite of BD ($-BD$) or the opposite of KL ($-KL$) for this comparison to measure $CUCog$, which indicates the similarity between the two aforementioned documents. Similarly, the higher the $-BD$ or $-KL$ is, the less the distance between and the more similar the topic distributions in a rebuttal posting that the user commented on and that user's background postings.

In particular, in the main effect, we used $-BD$ to measure $CUCog$ (see Equation (5)); while in the robustness test, we used $-KL$ to measure $CUCog$ (see Equation (6)). In addition, KL divergence is not symmetric; referring to Vosoughi et al. (2018), as shown in Equation (6), we used a symmetric metric to measure this divergence. A higher $CUCog$ represents individuals with higher cognitive abilities regarding fake news rebuttals. The calculation formula is $CUCog = \sum_{i=1}^n CUCog_i$.

$$CUCog = -BD(\Gamma_{rp}, \Gamma_{bp}) = \ln \left(\sum_{x \in X} \sqrt{\Gamma_{rp}(x)\Gamma_{bp}(x)} \right) \quad (5)$$

$$CUCog = -KL(\Gamma_{rp}, \Gamma_{bp}) = -\frac{1}{2} \left(\sum_{x \in X} \Gamma_{rp}(x) \log \left(\frac{\Gamma_{rp}(x)}{\Gamma_{bp}(x)} \right) + \sum_{x \in X} \Gamma_{bp}(x) \log \left(\frac{\Gamma_{bp}(x)}{\Gamma_{rp}(x)} \right) \right) \quad (6)$$

where Γ_{rp} and Γ_{bp} correspond to the topic probability distributions of a fake news rebuttal posting and the background postings of the user who commented on this rebuttal posting, respectively; X corresponds to the set of 16 topics; $\Gamma_{rp}(x)$ and $\Gamma_{bp}(x)$ correspond to individual topics in Γ_{rp}

and Γ_{bp} . $\sum_{x \in X} \Gamma_{rp}(x) \log \left(\frac{\Gamma_{rp}(x)}{\Gamma_{bp}(x)} \right)$ measures the amount of information lost in the approximation of the probability distribution $\Gamma_{rp}(x)$ with $\Gamma_{bp}(x)$. Lower KL or BD indicates lower divergence between the two documents; therefore, the higher $-KL$ or $-BD$ is, the more similar Γ_{rp} and Γ_{bp} are.

4.2.4. Control variables

To eliminate the interference of other factors in the results, control variables, including the gender of commenting users ($CUGen$), the registration time of commenting users ($CURegTime$), the social relation between the rebuttal source and commenting users ($SocialRela$), the category of fake news ($Category$), the frequency of the involvement of commenting users ($Involve$), the number of postings of commenting users ($CUTotalPos$), and the number of followers of commenting users ($CUTotalFolw$), were examined. These variables are dynamic and constantly change; therefore, we used the values at the moment when the users posted the relevant comments.

The definitions of the main variables in this study are listed in Table 8. This table includes the dependent variable, independent variables, moderate variable, and control variables.

4.3. Research Procedures

Regression analysis would be more appropriate than the structural equation model to investigate the specific mechanisms of action in the model (Gefen et al., 2000). Particularly is the test for the moderating effect, regression is powerful when dealing with various forms of correlated errors and model testing and has been among the popular statistical techniques for testing theory in several academic disciplines (Hair, 2009). In this study, we investigated the relationships among the central route (rebuttal's information readability and argument quality), peripheral route (rebuttal's source credibility, including authority and influence), and rebuttal acceptance, as well as the moderating effect of receiver's cognitive ability on these relationships. The topic of this study is the users' comments on the postings of fake news rebuttals to reveal the underlying mechanisms of users' rebuttal acceptance. Therefore, in this study, it is feasible to use regression analysis as an empirical

Table 8
Description of variables.

Variable	Description
Dependent Variable	
$RAccept$	$= \sum_{i=1}^n RAcept_i$ (The commenting user's acceptance of a fake news rebuttal posting)
Independent Variables	
$RSAutho$	$= \sum_{i=1}^n RSAutho_i$ (The authority of a rebuttal source)
$RSInflu$	$= \sum_{i=1}^n RSInflu_i$ (The influence of a rebuttal source)
$RIRead$	$= \sum_{i=1}^n RIRead_i$ (The information readability of a rebuttal posting)
$RAQuali$	$= \sum_{i=1}^n RAQuali_i$ (The argument quality of a rebuttal posting)
Moderator Variable	
$CUCog$	$= \sum_{i=1}^n CUCog_i$ (The level of the commenting user's cognitive ability regarding a rebuttal posting)
Control Variables	
$CUGen$	The gender of commenting users; 0 for Female; 1 for Male
$CURegTime$	The registration time of commenting users: the time at which the user commented on the fake news rebuttal minus the length of time he/she registered for Sina Weibo (in seconds)
$Category$	The type of fake news (see Table 2); 1 for authority; 2 for society; 3 for politics
$Involve$	The frequency of the involvement of commenting users
$SocialRela$	Whether there is a following or followed relation between the rebuttal source and the commenting user; 1 for following or followed relation; 0 for no relation
$CUTotalPos$	The number of postings of commenting users
$CUTotalFolw$	The number of followers of commenting users

method. The process of this study consisted of two parts.

- (1) The first part was a text mining analysis based on deep learning algorithms and the LDA topic model. The main purpose of text mining was to extract independent, control, moderating, and dependent variables from unstructured text information from rebuttal sources and commenting users. For text mining, Python version 3.6.2 was used as the main analysis platform. The process of text mining included text data acquisition and preprocessing (see (1)–(2) and (5) in Fig. 4), manual tagging of a corpus, classifier construction based on a deep learning algorithm, classification evaluation, obtaining the dependent variable (see (3)–(4) in Fig. 4), and extracting information to obtain the independent, control, and moderating variables of this study (see (6)–(7) in Fig. 4).
- (2) The second part was an empirical analysis based on econometrics (see Fig. 4). The main purpose of the empirical analysis was to explore the impact of source credibility, information readability, and argument quality on rebuttal acceptance as well as the moderating effect of cognitive ability on the relationships between them. A regression model was used to analyze the data. First, we specified a main model (Model 1), which includes four variables. This mainly tests peripheral route and central route determinants refuting fake news, namely, the influence of the rebuttal's source credibility (authority and influence) on rebuttal acceptance as well as the influence of the rebuttal's information readability, and argument quality on rebuttal acceptance. Second, we added moderating variables to the model (Model 2). In this way, we were able to test our hypotheses, both for the peripheral route and the central route of information processing, including the effects of the individual's cognitive ability on rebuttal acceptance. This enabled us to test our hypotheses, including the moderating effect of an individual's cognitive ability on the relationship between the rebuttal's source credibility and rebuttal acceptance as well as the moderating effects of

an individual’s cognitive ability on the relationship between the rebuttal’s information readability, and argument quality and rebuttal acceptance.

The equations are presented as follows.

(1) Model 1.

$$RAccept_i = \beta_0 + \beta_1 * RSAutho_i + \beta_2 * RSInflu_i + \beta_3 * RIRead_i + \beta_4 * RAQuali_i + \beta_5 * Control_i + \varepsilon_i + \mu_i$$

(2) Model 2.

$$RAccept_i = \beta_0 + \beta_1 * RSAutho_i + \beta_2 * RSInflu_i + \beta_3 * RIRead_i + \beta_4 * RAQuali_i + \beta_5 * CUCog_i * RSAutho_i + \beta_6 * CUCog_i * RSInflu_i + \beta_7 * CUCog_i * RIRead_i + \beta_8 * CUCog_i * RAQuali_i + \beta_9 * Control_i + \varepsilon_i + \mu_i$$

where the variable definitions in the equations are as described in the previous section.

5. Data analysis and results

5.1. Data analysis

Given that the variables varied in their initial magnitudes, we standardized the independent, control, and dependent variables for the main analysis. STATA (version 16.0) was used to analyze the data. Table 9 presents the variable statistics.

We used a regression model to test the rebuttal’s source credibility, information readability, and argument quality in relation to rebuttal acceptance as well as the moderating effects caused by an individual’s cognitive ability. The model was estimated in three stages: the first stage only included the control variables (Model I), the second stage included the independent variables (Model II), and the third stage tested the full model with the interactions (Model III). All variables were standardized in the regressions. Table 10 summarizes the regression results.

First, the results indicated that the source authority had a negative effect on rebuttal acceptance ($\beta = -0.039, p < .01$ in Model II), while source influence had a positive effect on rebuttal acceptance ($\beta = 0.041, p < .01$ in Model II). Moreover, the effect of source influence outweighed that of source authority regarding rebuttal acceptance. Therefore, H1 and H2 are supported.

Second, the results indicated that both information readability ($\beta = 0.047, p < .01$ in Model II) and argument quality ($\beta = 0.089, p < .01$ in Model II) had a positive effect on rebuttal acceptance. Moreover, the effect of argument quality outweighed that of information readability regarding rebuttal acceptance. Therefore, H3 and H4 are supported.

Third, the moderating effect of an individual’s cognitive ability on the effect of source authority ($\beta = 0.022, p < .05$ in Model III) on rebuttal acceptance was positive, thus supporting H5a. The moderating effect of an individual’s cognitive ability on the effect of source influence

Table 9
Variable statistics.

Variable	Mean	Std. Dev.	Min	Max	VIF
(1) RAccept	0.317	0.444	0	0.998	/
(2) RSAutho	3.771	0.926	1	5	2.81
(3) RSInflu	41,791,016	36,173,742	104	119,089,319	2.73
(4) RIRead	-0.709	0.772	-3.874	11.570	1.34
(5) RAQuali	1.439	0.553	0	2	1.44
(6) CUCog	-1.082	0.417	-2.589	0	1.13
(7) CUGen	0.622	0.485	0	1	1.03
(8) CURegTime	205,105,842	88,789,950	6,780,371	338,833,382	1.04
(9) Category	1.443	0.672	1	3	1.43 ^a 1.33 ^b
(10) Involve	1.340	1.183	1	13	1.02

Note: a. the number of observations in the final regression model was 4277. b. Category = 2: 1.43^a; Category = 3: 1.33^b.

Table 10
Regression results.

	Model I	Model II	Model III
CUGen	-0.007 (0.013)	-0.006 (0.013)	-0.005 (0.013)
CURegTime	-0.013** (0.006)	-0.010 (0.006)	-0.009 (0.006)
Category	0.383*** (0.015)	0.462*** (0.017)	0.438*** (0.017)
	0.002 (0.021)	-0.002 (0.023)	-0.014 (0.023)
Involve	-0.034*** (0.006)	-0.030*** (0.006)	-0.031*** (0.006)
RSAutho		-0.039*** (0.010)	-0.032*** (0.010)
RSInflu		0.041*** (0.010)	0.036*** (0.010)
RIRead		0.047*** (0.007)	0.045*** (0.007)
RAQuali		0.089*** (0.007)	0.082*** (0.007)
CUCog* RSAutho			0.022** (0.010)
CUCog* RSInflu			-0.021** (0.010)
CUCog* RIRead			0.025*** (0.006)
CUCog* RAQuali			-0.022*** (0.007)
Constant	0.230*** (0.011)	0.211*** (0.011)	0.220*** (0.011)
R-squared	0.143	0.184	0.193
N	4277	4277	4277
F	142.77	106.72	72.78

Note: Significance levels: *** $p < .01$, ** $p < .05$, * $p < .1$.

Table 11
Summary of hypotheses testing.

Hypotheses	Results
H1: The authority of a rebuttal source has a negative effect on the acceptance of fake news rebuttals.	Supported
H2: The influence of a rebuttal source has a positive effect on the acceptance of fake news rebuttals.	Supported
H3: The information readability of a rebuttal has a positive effect on the acceptance of fake news rebuttals.	Supported
H4: The argument quality of a rebuttal has a positive effect on the acceptance of fake news rebuttals.	Supported
H5a: The effect of source authority on the acceptance of fake news rebuttals will be negatively moderated by individuals’ cognitive ability regarding fake news rebuttals such that this effect will be weakened by individuals with higher cognitive abilities.	Supported
H5b: The effect of source influence on the acceptance of fake news rebuttals will be negatively moderated by individuals’ cognitive ability regarding fake news rebuttals such that this effect will be weakened by individuals with higher cognitive abilities.	Supported
H5c: The effect of information readability on the acceptance of fake news rebuttals will be positively moderated by individuals’ cognitive ability regarding fake news rebuttals such that this effect will be strengthened by individuals with higher cognitive abilities.	Supported
H5d: The effect of argument quality on the acceptance of fake news rebuttals will be positively moderated by individuals’ cognitive ability regarding fake news rebuttals such that this effect will be strengthened by individuals with higher cognitive abilities.	Not Supported

($\beta = -0.021, p < .05$ in Model III) on rebuttal acceptance was negative; thus, H5b is supported.

Fourth, the moderating effect of an individual's cognitive ability on the effect of information readability ($\beta = 0.025, p < .01$ in Model III) on rebuttal acceptance was positive; thus, H5c is supported. Conversely, the moderating effect of an individual's cognitive ability on the effect of argument quality ($\beta = -0.022, p < .01$ in Model III) on rebuttal acceptance was negative; thus, H5d is not supported.

Table 11 summarizes the results of the hypotheses testing.

Furthermore, to express the interaction effects more clearly, an interaction diagram was constructed according to the procedures of Aiken et al. (1991). The results are shown in Figs. 5–8.

Fig. 5 shows that the effects of source authority on rebuttal acceptance (*RAccept*) were found to be negative under both high and low cognitively abled individuals. However, for high cognitive ability (dotted line), rebuttal acceptance decreases more slowly than for low cognitive abilities (solid line), indicating that high cognitive ability decreases the negative effect of source authority on rebuttal acceptance. Fig. 6 shows that the effects of source influence on rebuttal acceptance were found to be positive under both high and low cognitive ability. However, for high cognitive ability (dotted line), rebuttal acceptance increases more slowly than for low cognitive ability (solid line), indicating that high cognitive ability weakens the positive effect of source influence on rebuttal acceptance. Fig. 7 also shows that the effects of information readability on rebuttal acceptance were found to be positive under high and low cognitive abilities, and the effects of the interaction of information readability on rebuttal acceptance were stronger under high cognitive ability (dotted line). As shown in Fig. 8, for low cognitive ability (solid line), rebuttal acceptance increases more rapidly than for high cognitive ability (dotted line), indicating that high cognitive ability weakens the positive effect of argument quality on rebuttal acceptance.

5.2. Robustness checks

Furthermore, to test the robustness of our main findings, we used two different methods. First, we tested the robustness of the main effect model by changing the control variables. Specifically, we provided a version that adds other control variables, including *SocialRelaCUTotalPos*, and *CUTotalFollow*, which helped control for unobservable factors related to each commenting user. Table 12 (Models II and III) shows that the results of the robustness tests are consistent with the regression results (Models II and III in Table 10).

Second, we considered using the calculation method in Equation (6) to measure the individual's cognitive ability and to reexamine the main effect. The results of the robustness tests are presented in Table 12.

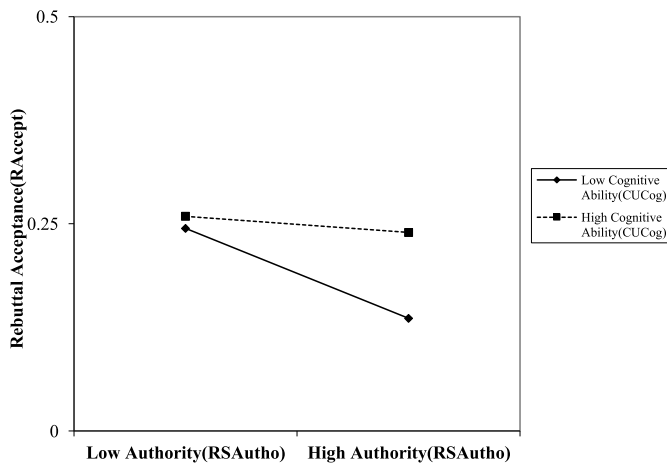


Fig. 5. The moderating effects of cognitive ability on rebuttal acceptance (*RAccept*).

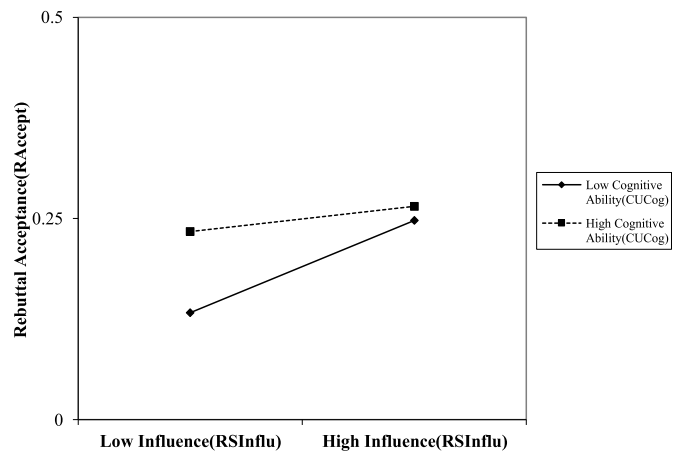


Fig. 6. The moderating effects of cognitive ability on rebuttal acceptance (*RAccept*).

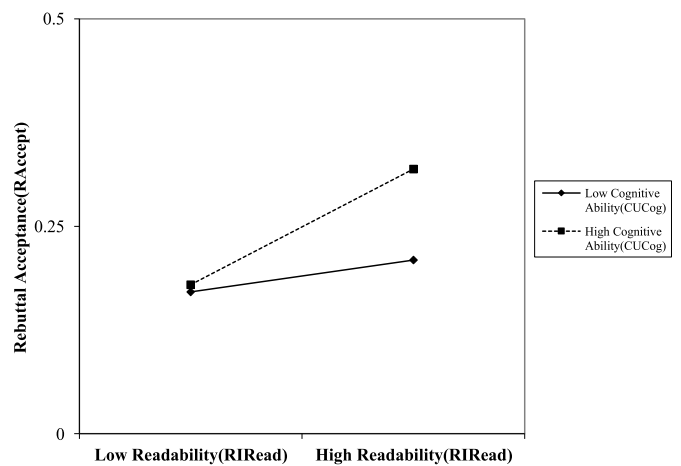


Fig. 7. The moderating effects of cognitive ability on rebuttal acceptance (*RAccept*).

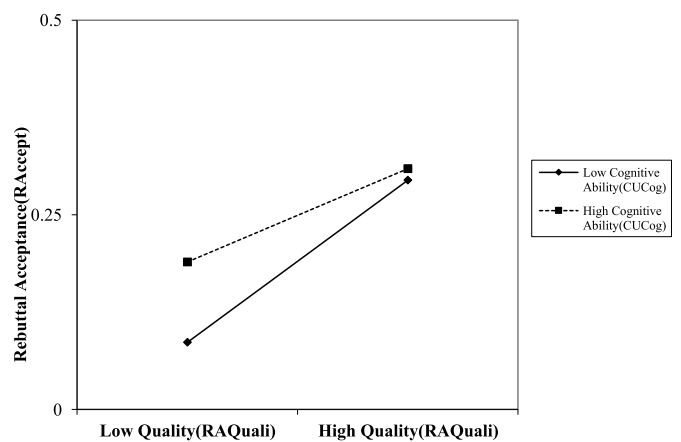


Fig. 8. The moderating effects of cognitive ability on rebuttal acceptance (*RAccept*).

Compared to Table 10, the results of the robustness tests (Models IV and V in Table 12) are consistent with the regression results (Model III in Table 10 and Model III in Table 12, respectively). The range of VIF values for all models was below the acceptable limits (Model III: 1.02–2.82; Model IV: 1.02–2.77; Model V: 1.02–2.78). Therefore, we

Table 12
Results of robustness checks.

	Model I	Model II	Model III	Model IV	Model V
CUGen	-0.010 (0.013)	-0.008 (0.013)	-0.007 (0.013)	-0.005 (0.013)	-0.006 (0.013)
CURegTime	-0.013** (0.006)	-0.011* (0.006)	-0.010 (0.006)	-0.010 (0.006)	-0.011* (0.006)
Category	0.383*** (0.015)	0.460*** (0.017)	0.437*** (0.017)	0.448*** (0.017)	0.446*** (0.017)
	-0.002 (0.021)	-0.003 (0.023)	-0.015 (0.023)	-0.002 (0.023)	-0.003 (0.023)
Involve	-0.034*** (0.006)	-0.031*** (0.006)	-0.031*** (0.006)	-0.030*** (0.006)	-0.031*** (0.006)
SocialRela	0.066*** (0.014)	0.032** (0.014)	0.027* (0.014)		0.029** (0.014)
CUTotalPos	0.009 (0.006)	0.011* (0.006)	0.010* (0.006)		0.011* (0.006)
CUTotalFollow	0.005 (0.006)	0.005 (0.006)	0.004 (0.006)		0.005 (0.006)
RSAutho		-0.039*** (0.010)	-0.032*** (0.010)	-0.034*** (0.010)	-0.034*** (0.010)
RSInflu		0.038*** (0.010)	0.035*** (0.010)	0.037*** (0.010)	0.035*** (0.010)
RIRead		0.047*** (0.007)	0.044*** (0.007)	0.049*** (0.007)	0.048*** (0.007)
RAQuali		0.087*** (0.007)	0.081*** (0.007)	0.084*** (0.007)	0.083*** (0.007)
CUCog* RSAutho			0.021** (0.010)	0.023** (0.010)	0.022** (0.010)
CUCog*RSInflu			-0.021** (0.010)	-0.020** (0.010)	-0.020** (0.010)
CUCog*RIRead			0.024*** (0.006)	0.026*** (0.007)	0.025*** (0.007)
CUCog*RAQuali			-0.021*** (0.007)	-0.022*** (0.007)	-0.021*** (0.007)
Constant	0.212*** (0.012)	0.203*** (0.012)	0.213*** (0.012)	0.216*** (0.011)	0.208*** (0.012)
R-squared	0.148	0.186	0.194	0.191	0.193
N	4277	4277	4277	4277	4277
F	92.90	80.97	60.44	72.06	59.95

Note: Significance levels: *** $p < .01$, ** $p < .05$, * $p < .1$.

concluded that most of our findings are robust.

6. Discussion and conclusion

6.1. Key findings

This study offers three key insights. First, we found that in the peripheral route of ELM, the credibility of a rebuttal source influenced rebuttal acceptance. These results accord with those of Eagly and Chaiken (1993), which indicated that source credibility was an important determinant of refutation. That is, our results show that source authority can weaken rebuttal acceptance, which is not in agreement with the ELM, while source influence can enhance rebuttal acceptance, concordant with the ELM. Concomitantly, source influence had a stronger positive effect of refuting fake news than the negative effect of source authority, which confirms that the greater the perceived influence of the source, the more persuasive the source will be (Susarla et al., 2012). Since our source influence was measured by the number of followers of a refuter on social media, our results indicate that the relationships on social media can enhance the rebuttal acceptance of sources with high influence. This is similar to the results of Rui et al. (2013), wherein the positive effect of influencing the acceptance of users with more followers was significantly larger than that of users with fewer followers. However, another study suggests that users show rebuttal distrust or resistance to refuters with higher authority. This confirms that because of psychological reactance (Brehm & Brehm, 2013), people generally do not like to be told what to think and how to act; they may reject particularly authoritative fake news rebuttals, especially if they concern complex issues such as public health crises (Lewandowsky et al., 2012). In such real-world scenarios, people will refer more to fake news that is in line with their attitudes and will be relatively immune to corrections, which may even backfire and strengthen their initially-held beliefs (Nyhan & Reifer, 2010). The people may not only reject the message at hand but also become predisposed to reject any future messages from its source (Brehm & Brehm, 2013). Some researchers have indicated that during the COVID-19 outbreak, citizens can be persuaded by authorities' denials and reduce their belief in a piece of news that has been declared "fake" (Wang & Huang, 2021). However, when the denied news turns out to be real, citizens will reduce their belief not only in the denial at hand but also in a similar denial in the future and will reduce their satisfaction with the government (Wang & Huang, 2021). An example of this is the case of Dr. Wenliang Li, who warned of a SARS-like disease in the early stages of the

COVID-19 outbreak, along with several other doctors in China, and was reprimanded by authorities for rumor-mongering and then, several months later, amid the global pandemic, was officially recognized as a martyr (Wang & Huang, 2021). Particularly during the COVID-19 pandemic, the public has expressed a high degree of trust and expectations in the authority's media outlets and organizations (Falcone & Sapienza, 2020). This also reflects the urgent need for information regarding the latest developments in the treatment of COVID-19. Taking Event 3 (a postgraduate from the Wuhan Institute of Virology was "Patient Zero") in our dataset as an example, owing to the public dissatisfaction with the delayed treatment during the early stages of the COVID-19 epidemic on the part of the Wuhan government as well as the accumulation of distrust in them since the situation with Dr. Wenliang Li, people have doubted the ability of the authoritative organizations to manage the epidemic, thereby weakening the acceptance of receivers to fake news rebuttals released by high-authority sources. Therefore, irrespective of the measures taken, the public will have an attitude of distrust and unacceptance, creating a reverse authority effect (Jung & Kellaris, 2006), resulting in the source authority having a negative effect on rebuttal acceptance; consequently, the authority's organizations and media outlets fall into the Tacitus trap owing to the loss of credibility (Horn, 2011).

Second, we found that in the central route of ELM, argument quality and information readability could enhance rebuttal acceptance, which is concordant with the ELM (Bordia et al., 2005; Zhou et al., 2021). Our results indicate that when posting a rebuttal, the audience needs not only information readability but also argument quality (Petty & Cacioppo, 1986, pp. 1–24). Moreover, the effect of argument quality on rebuttal acceptance was more significant than that of information readability, which confirms that in the ELM, when argument quality is strong, the message contains facts that are justified and compelling as well as generally more persuasive (Angst & Agarwal, 2009; Petty & Cacioppo, 1986, pp. 1–24). Our results show that rebuttals with strong arguments (e.g., an argument embedded with reasonable, detailed, and adequate evidence) can provide an element of vividness to messages (Iyer & Debevec, 1991). Further, they may be perceived as more lucid and persuasive than a flat denial (Paek & Hove, 2019), especially in fake news rebuttals. Thus, using convincing and persuasive arguments can significantly improve rebuttal acceptance (Pal et al., 2020). Our results also show that rebuttals tend to be informative and factual in nature; thus, their information readability (i.e., assurance of ease of readability) is significantly associated with rebuttal acceptance, which is similar to the research of Korfiatis et al. (2012); that is, a text that is easy to read is

more effective in persuasion. In addition, we also found that although both routes are important as they relate to fake news rebuttals, the central route is more influential (Bhattacharjee & Sanford, 2006), that is, argument quality and information readability are more salient than source credibility, which confirms that the gist of a message is often more memorable than its source (Eagly & Chaiken, 1993; Hachaturyan et al., 2021).

Third, we also found that individuals' cognitive ability played a role in moderating the relationship between source credibility and rebuttal acceptance as well as the relationship between information readability, argument quality, and rebuttal acceptance. Our results show that individuals with low cognitive abilities may rely more on source influence and authority in source credibility. According to the ELM, these results can be interpreted as follows: When individuals are weak in processing the information they are exposed to, they will adopt a simple processing mode of information—that is, acceptance of the persuasion of peripheral cues (Bordia et al., 2005). Interestingly, despite the negative effect of authority on rebuttal acceptance, users with low cognitive abilities are still more inclined to rely on peripheral routes for decision-making; therefore, in Fig. 5, rebuttal acceptance declines faster than high cognitive abilities. Under certain circumstances, the effectiveness of authority may not be as persuasive as the authority principle suggests, and a reverse authority effect is likely to be amplified or attenuated under different cultural circumstances and receivers' characteristics (Jung & Kellaris, 2006). In addition, previous findings highlight that the ability to distinguish the false nature of fake news from real information lies in the differences in the cognitive ability of users (Ahmed, 2021). Individuals with lower (versus higher) levels of cognitive abilities were less responsive to refuting messages, and the initial exposure to the fake news had a persevering influence on their attitudes (Roets, 2017). For those with lower levels of cognitive abilities, even after the explicit disconfirmation of the false information, adjusted attitudes remained biased (Roets, 2017). Thus, in the particular research context mentioned in the first finding above, source authority can weaken rebuttal acceptance; when faced with fake news rebuttals from a higher authority source, the adjusted attitudes of individuals with lower levels of cognitive abilities were more negative compared to individuals with higher levels of cognitive abilities (Roets, 2017), and thus, individuals with lower cognitive abilities reinforce such a negative effect of authority on rebuttal acceptance.

At the same time, for individuals with high cognitive abilities, the ELM's traditional research also indicates that they will rely more on central cues (Petty & Cacioppo, 1986, pp. 1–24). However, our results show that individuals with high cognitive abilities simply had a higher demand for information readability in the central route, indicating that for them, there is a greater likelihood of information readability to improve their acceptance of fake news. Meanwhile, we can also see from the formula of *RRead* (Equation (4)) that the smaller the length of the sentence (*ASL*), the smaller the total number of characters (*TW*), and the larger the percentage of common Chinese characters (*PCW*) in a refuting posting, the easier it is to read the rebuttals. That is, the more concise the rebuttal and the higher the proportion of commonly used words, the more likely it is that individuals with high cognitive abilities will accept the rebuttal. Moreover, although individuals with low cognitive abilities have a higher demand for peripheral cues, their demand for the argument quality in the central route is equally high. Our results show that the change in the increase in rebuttal acceptance by argument quality was not as pronounced for individuals with high cognitive abilities as for individuals with low cognitive abilities; that is, providing strong argument plays an important role in increasing rebuttal acceptance by individuals with low cognitive abilities. This interesting and surprising result is similar to that of Yada and Head (2019); in the same health-related research, attitude formation of those with low levels of user expertise was most influenced by central rather than peripheral routes of persuasion, although this conflicts with extant ELM research in non-healthcare or public health contexts. Our results confirm that, for

those with lower levels of cognitive abilities, strong arguments are indeed an effective messaging tactic to improve the acceptance of fake news rebuttals, especially in the public health field (Yada & Head, 2019); While the strong-argument condition may have provided no new information for individuals with high cognitive abilities (Rosen, 2000). We believe that individuals with high cognitive abilities have low demand for argument quality because of two reasons. One explanation is that individuals with high cognitive abilities are relatively more knowledgeable and familiar with the field of fake news rebuttals, and they are more likely to accept familiar information as true (Swire et al., 2017). Therefore, simply enhancing the quality of argument did not have a significant impact on increasing the rebuttal acceptance of individuals with higher cognitive abilities. Another explanation is that for individuals with high cognitive abilities, information overloading is a hindrance (Laato et al., 2020). Therefore, a strong and high-quality argument is not necessary for individuals with high cognitive abilities. In contrast, individuals with high cognitive abilities may prefer to disseminate information that is simple, easy to digest, and easy to read on social media (Metzger & Flanagin, 2013).

6.2. Theoretical contributions

This study has implications for both theory and practice. On the theoretical front, the study provides two theoretical insights. First, to the best of our knowledge, by applying to ELM, this study is one of the first to explore the underlying mechanism of rebuttal acceptance during the COVID-19 pandemic from a large volume of text information based on text mining (e.g., deep learning algorithms, LDA topic model) and econometrics analysis technology on Chinese social media platforms. By seeing the growing concern of online fake news, an emerging body of research has begun to focus on how to effectively debunk fake news (Li et al., 2021). As a way to extend such research, this paper examines how individuals with different cognitive abilities accept fake news rebuttals on social media in China. Although fake news rebuttals have been thoroughly researched in recent decades, most prior research has focused on traditional contexts and quantification methods (Pal et al., 2020). Our findings have enriched the literature on the mechanism of online rebuttal management and intervention from a text mining perspective on social media.

Second, as a highly influential, valuable, and popular framework (Kitchen et al., 2014; Morris et al., 2005), the majority of applied research on ELM has been conducted in the fields of social psychology and marketing (e.g., advertising and consumer behavior) to describe how individuals process information (Petty & Cacioppo, 1986, pp. 1–24). However, in the context of debunking fake news using rebuttals, this study extends the applicability of the ELM by analyzing the two routes of information processing in the fake news rebuttal field on social media: peripheral cues (influence and authority) and central cues (information readability and argument quality). It develops a research model by which to examine the underlying mechanism of how the peripheral and central clues of rebuttals are moderated by individuals' differing cognitive abilities to determine rebuttal acceptance. On the one hand, this study extracts these important features from the acceptance of fake news rebuttals based on ELM. Although there might be additional influential characteristics that we have not yet identified, the ELM model provides us with a well-grounded foundation for the identification of possible characteristics related to fake news rebuttals. On the other hand, the ELM has not focused much on deciphering how features are measured, such as how stronger and weaker arguments are measured (Srivastava & Kalro, 2019). In the context of fake news rebuttals during the COVID-19 pandemic in China, this study extends the literature on ELM and online fake news by measuring factors (source authority, source influence, information readability, argument quality, and cognitive ability) from a text-mining perspective. Specifically, this study proposes a measurement of argument quality from the perspective of fake news rebuttals, combines the Chinese readability formula to

measure the readability of rebuttal postings, proposes to assess source credibility from both authority and influence perspectives, proposes a measurement to score the authority of the rebuttal source, measures the source influence based on the number of followers, and uses the LDA topic model to measure individuals' cognitive abilities related to fake news rebuttals. By supporting the peripheral route, the findings highlight that source credibility plays a crucial role in rebuttal acceptance. Source authority can weaken rebuttal acceptance, whereas source influence can enhance it. By supporting the central route, the results demonstrate a greater positive role of information readability and argument quality on rebuttal acceptance, and argument quality has a greater impact. By supporting the moderating role of individual cognitive ability, we found that individuals with low cognitive abilities rely more on peripheral cues and argument quality in central cues to accept rebuttals, while individuals with high cognitive abilities rely more on information readability in central cues to accept rebuttals. Thus, the study offers greater insight into how the use of rebuttals can be harnessed as an effective way to influence individuals' likelihood of accepting rebuttals. It also identifies the pathways through which individuals with different cognitive abilities accept fake news rebuttals, which has enriched the application of the ELM theory in the online fake news rebuttal field, particularly during the COVID-19 pandemic.

6.3. Practical implications

On the practical front, this study also provides two implications. First, it offers insight for news outlet professionals, social platform managers, and Chinese government regulators regarding the use of rebuttals in combatting online fake news during the COVID-19 pandemic. This paper has highlighted the information readability, argument quality, and source credibility features of rebuttals and has shown how rebuttals can be crafted effectively to serve the purpose of refuting fake news on social media. The findings suggest that the information readability and argument quality of rebuttal and the influence of source credibility are key features that should be leveraged to develop rebuttals, particularly argument quality, which has a greater impact. For example, arguments provided in an online rebuttal text must be credible; account for the important causal qualities; and provide persuasive, reasonable, and concrete evidence in support, such as by explaining why fake news was considered correct in the first place (Lewandowsky et al., 2012). In addition, source authority has a negative effect on rebuttal acceptance, while source influence could enhance rebuttal acceptance. Thus, to effectively combat fake news, on the one hand, we suggest encouraging individuals with a higher number of followers to share high-quality arguments and easy-to-read refuting messages, leveraging their influence on social media to help people receive rebuttals. On the other hand, those attempting to refute fake news should improve their authority to further develop refuting information. For instance, state media outlets and government-affiliated accounts with high authority should consider the ecological construction of public opinion and improve their credibility to win the public's approval. They should leverage social platforms to combat fake news, attend to the needs and intentions of the public, pay attention to public feedback, appropriately encourage citizen engagement to build trusted relationships beyond simple information exchange on social media platforms, enhance understanding and interactive communication, and formulate effective refuting strategies to eliminate the existence of fake news as early as possible (Chen et al., 2020).

Second, while most practitioners are facing growing concern regarding fake news within the realm of social media, this paper shows the extent to which individuals with different cognitive abilities accept fake news rebuttals in terms of source credibility, information readability, and argument quality. These results demonstrate that source credibility, information readability, and argument quality have varying effects on rebuttal acceptance depending on the differing cognitive ability of various individuals. Individuals with high cognitive abilities

were more likely to appreciate information readability in rebuttals, which showed their acceptance of rebuttals by central route processing. Furthermore, individuals with low cognitive abilities were more likely to appreciate source influence, source authority, and argument quality in rebuttals, which showed their acceptance of rebuttals by both peripheral and central route processing. Thus, rebuttals must be customized to their intended recipients as individuals with different cognitive abilities can perceive the same things via dramatically different routes. Hence, we suggest that practitioners must understand how factors in peripheral and central routes can enhance the overall effectiveness of rebuttals by influencing individuals with different cognitive abilities to respond to fake news refutation. For example, if an individual's level of cognitive ability is clear, for those with low cognitive abilities, rebuttals can be designed by focusing on improving peripheral clues and argument quality. In addition, more efforts are needed in the form of digital media literacy and cognitive development programs to educate individuals with lower cognitive abilities (Ahmed, 2021) and improve their acceptance of fake news rebuttals. For individuals with high cognitive abilities, only information readability needs to be emphasized in rebuttals by focusing on improving the skills and motivation of the receivers to make a reading effort, such as encouraging the individual to acquire news literacy that involves more thorough knowledge of the processes by which news is produced, distributed, and consumed and the skills to control some of these processes (Vraga et al., 2021). In this way, an accurate, efficient, and cost- and resource-efficient rebuttal effect can be created for individuals with different cognitive abilities. Given the lack of control and management of individuals with different cognitive abilities in the online setting, practitioners should develop rebuttals by enhancing both peripheral routes (e.g., source influence and authority) and central routes (e.g., information readability and argument quality) to maximize their impact. These results provide guidance and reference for practitioners to provide targeted rebuttals for individuals with different cognitive abilities. This paper indicates some confidence in the use of rebuttals, particularly when COVID-19-related fake news disrupts the social order and threatens the reputation of state media and government organizations.

6.4. Limitations and future research

Drawing on the ELM of persuasive theory, this study primarily examines how individuals with different cognitive abilities accept fake news rebuttals through two aspects: source credibility (authority and influence) in the peripheral route and information readability and argument quality in the central route. The study aimed to explore an effective way to refute fake news to further eliminate its adverse effects during the COVID-19 pandemic on social media in China. In addition to its theoretical contributions and practical implications, this study has several limitations.

First, five fake news events related to the COVID-19 pandemic were chosen for investigation in the Chinese social media environment; the generalization of the results to other countries and different cultures is constrained. Thus, future studies should be focused on research in other countries and contexts to test whether these results can extend to those countries in a more general context, such as to investigate the varying levels of the effect of authority across countries and cultural backgrounds and to examine factors that might exacerbate or dissipate the effect of authority by analyzing the potential moderators of psychological reactance. Meanwhile, future research could consider expanding the study of fake news rebuttals to those from other domains, such as health and politics. Furthermore, in future relevant studies, more generalizable and interesting results might be achieved using adequate sample size and multiple platforms and making comparisons between platforms in different cultural environments, such as Sina Weibo and Twitter.

Second, although we attempted to use the optimal classifier for text classification in this study, some errors cannot be avoided in the text analysis process. Future research could improve the accuracy of the

classifier by improving the algorithm to classify users who comment on fake news rebuttals in a more diverse and fine-grained manner.

Third, this research only considers cognitive ability as a moderator and simplifies the peripheral and central routes of ELM by using source credibility (influence and authority), information readability, and argument quality. Early persuasion theory contends that source credibility may facilitate the attention and comprehension of a message's arguments (Hamilton et al., 1993); judgments regarding the source may not occur independently of judgments regarding the quality of the message (Stephenson et al., 2001), a relationship that was not predicted by the ELM in our study. Future work should consider extending the ELM, including examining more central and peripheral cues, and considering whether the presence of source credibility could increase the favorability of an argument. In addition, some research (e.g., online review research on e-commerce platforms) has indicated that text readability tends to influence readers' motivation and ability to process arguments and amplify the impact of an argument by playing a moderating role in the relationship between online review factors and review helpfulness (Li et al., 2017). Since this research considers the moderating role of only an internal ability factor (cognitive ability) and not that of an external ability factor (readability), it may limit the study's ability to process the argument thoroughly (Chebat et al., 2003). In other words, the potential moderating role of readability in the relationship between argument quality and rebuttal acceptance was not predicted by the ELM in our study. Future work should consider examining the moderating effect of readability on central routes to gain more insights on how factors influence rebuttal acceptance. Similarly, future research should further consider how the individual, with their high/low motivation and/or high/low ability, selects central and peripheral cues to accept fake news rebuttals and whether there are some relations between the two routes.

Fourth, although we have tried our best to improve the accuracy with which we extract variables from the abundance of text data available from social media, most of the effects in our results are significant but small and examine correlational rather than causal effects. We think that there remains some deviation in the accuracy of variable extraction; for example, this study chose to operationalize cognitive ability via verbal intelligence, a dimension closely associated with other forms of intelligence (Colom et al., 2005). In the future, we hope to find a more accurate method with which to extract more variables from online text data to improve the explanatory power of the model as well as to consider causality studies. This will make our results more convincing. In addition, contrary to our research hypothesis, cognitive ability plays a negative moderating role between argument quality and rebuttal acceptance. Although we have given two relatively reasonable explanations in this paper, scholars can further test the impact and find the real mechanism behind it.

Finally, based on the ELM, this study only considered the moderating effect of cognitive ability and did not consider the emotional connections generated during user interaction. Thus, future research should investigate the factors influencing fake news rebuttal acceptance by focusing on uncovering the mediating and moderating effects of emotion and cognitive ability, respectively, which is another avenue for future research.

Funding

This work was supported by the National Natural Science Foundation of China [grant number 72074060].

CRediT author statement

Xin Wang (X.W.): Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing - original draft, Writing - review & editing. Fan Chao (F.C.): Methodology, Formal analysis. Guang Yu (G.Y.): Conceptualization, Writing - review & editing, Supervision, Funding

acquisition. Kaihang Zhang (K.Z.): Data curation, Software.

Declaration of competing interest

The author(s) declared no potential conflicts of interest concerning the research, authorship, and/or publication of this article.

Acknowledgments

We would like to acknowledge the editor's contribution and show appreciation to the reviewers for their helpful comments and recommendations.

We would like to acknowledge Zhiwei Data Sharing Platform (<http://university.zhiweidata.com/>) for our data support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.chb.2021.107174>.

References

- Agnihotri, A., & Bhattacharya, S. (2016). Online review helpfulness: Role of qualitative factors. *Psychology and Marketing*, 33(11), 1006–1017.
- Ahmed, S. (2021). Who inadvertently shares deepfakes? Analyzing the role of political interest, cognitive ability, and social network size. *Telematics and Informatics*, 57, 101508.
- Aiken, L. S., West, S. G., & Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions*. sage.
- Allport, G. W., & Postman, L. (1946). An analysis of rumor. *Public Opinion Quarterly*, 10(4), 501–517.
- Andrei, V., & Arandjelović, O. (2016). Complex temporal topic evolution modelling using the Kullback-Leibler divergence and the Bhattacharyya distance. *EURASIP Journal on Bioinformatics and Systems Biology*, 1–11, 2016(1).
- Angst, C. M., & Agarwal, R. (2009). Adoption of electronic health records in the presence of privacy concerns: The elaboration likelihood model and individual persuasion. *MIS Quarterly*, 339–370.
- Barrón-Cedeno, A., Jaradat, I., Da San Martino, G., & Nakov, P. (2019). Propopy: Organizing the news based on their propagandistic content. *Information Processing & Management*, 56(5), 1849–1864.
- Bastick, Z. (2021). Would you notice if fake news changed your behavior? An experiment on the unconscious effects of disinformation. *Computers in Human Behavior*, 116, 106633.
- Beckage, N. M., & Colunga, E. (2016). Language networks as models of cognition: Understanding cognition through language. In *Towards a theoretical framework for analyzing complex linguistic networks* (pp. 3–28).
- Bernstein, D., Penner, L., Clarke-Stewart, A., & Roy, E. (2012). *Psychology* (9th, International ed.). Boston, MA: Cengage Learning.
- Bhattacharjee, A., & Sanford, C. (2006). Influence processes for information technology acceptance: An elaboration likelihood model. *MIS Quarterly*, 805–825.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research*, 3, 993–1022.
- Bordia, P., DiFonzo, N., Haines, R., & Chaseling, E. (2005). Rumors denials as persuasive messages: Effects of personal relevance, source, and message characteristics. *Journal of Applied Social Psychology*, 35(6), 1301–1331.
- Bordia, P., DiFonzo, N., & Schulz, C. A. (2000). Source characteristics in denying rumors of organizational closure: Honesty is the best policy. *Journal of Applied Social Psychology*, 30(11), 2309–2321.
- Brandt, M. J., & Crawford, J. T. (2016). Answering unresolved questions about the relationship between cognitive ability and prejudice. *Social Psychological and Personality Science*, 7(8), 884–892.
- Brehm, S. S., & Brehm, J. W. (2013). *Psychological reactance: A theory of freedom and control*. Academic Press.
- Brown, P. F., Della Pietra, S. A., Della Pietra, V. J., Lai, J. C., & Mercer, R. L. (1992). An estimate of an upper bound for the entropy of English. *Computational Linguistics*, 18(1), 31–40.
- Cao, Q., Duan, W., & Gan, Q. (2011). Exploring determinants of voting for the "helpfulness" of online user reviews: A text mining approach. *Decision Support Systems*, 50(2), 511–521.
- Chang, Y. T., Yu, H., & Lu, H. P. (2015). Persuasive messages, popularity cohesion, and message diffusion in social media marketing. *Journal of Business Research*, 68(4), 777–782.
- Chebat, J. C., Gelinias-Chebat, C., Hombourger, S., & Woodside, A. G. (2003). Testing consumers' motivation and linguistic ability as moderators of advertising readability. *Psychology and Marketing*, 20(7), 599–624.
- Chen, B., Chen, X., Pan, J., Liu, K., Xie, B., Wang, W., ... Jiang, J. (2021). Dissemination and refutation of rumors during the COVID-19 outbreak in China: Infodemiology study. *Journal of Medical Internet Research*, 23(2), Article e22427.

- Chen, C. D., & Ku, E. C. (2013). Bridging indistinct relationships and online loyalty: Evidence from online interest-based communities. *Online Information Review*, 37(5), 731–751.
- Chen, Q., Min, C., Zhang, W., Wang, G., Ma, X., & Evans, R. (2020). Unpacking the black box: How to promote citizen engagement through government social media during the COVID-19 crisis. *Computers in Human Behavior*, 110, 106380.
- Cheung, C. M., Lee, M. K., & Rabjohn, N. (2008). *The impact of electronic word-of-mouth: The adoption of online opinions in online customer communities* (Internet research).
- Cheung, M. Y., Luo, C., Sia, C. L., & Chen, H. (2009). Credibility of electronic word-of-mouth: Informational and normative determinants of on-line consumer recommendations. *International Journal of Electronic Commerce*, 13(4), 9–38.
- Choudhary, A., & Arora, A. (2021). Linguistic feature based learning model for fake news detection and classification. *Expert Systems with Applications*, 169, 114171.
- Chua, A. Y. K., & Banerjee, S. (2018). Intentions to trust and share online health rumors: An experiment with medical professionals. *Computers in Human Behavior*, 87, 1–9.
- Cialdini, R. B. (2001). *Influence: Science and practice*. Book, 3rd.
- Ciampaglia, G. L., Shiralkar, P., Rocha, L. M., Bollen, J., Menczer, F., & Flammini, A. (2015). Computational fact checking from knowledge networks. *PLoS One*, 10(6), Article e0128193.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46.
- Colom, R., Abad, F. J., Rebollo, I., & Shih, P. C. (2005). Memory span and general intelligence: A latent-variable approach. *Intelligence*, 33(6), 623–642.
- Coppola, V., & Girandola, F. (2018). When increasing the strength of the argument becomes counterproductive: The role of argumentative markers in the arousal of the psychological reactance. *Journal of Language and Social Psychology*, 37(4), 475–496.
- Cover, T. M. (1999). *Elements of information theory*. John Wiley & Sons.
- Cui, Y., Che, W., Liu, T., Qin, B., Wang, S., & Hu, G. (2020). Revisiting pre-trained models for chinese natural language processing. *arXiv preprint*. arXiv:2004.13922.
- Cui, Y., Che, W., Liu, T., Qin, B., Yang, Z., Wang, S., & Hu, G. (2019). Pre-training with whole word masking for chinese bert. *arXiv preprint*. arXiv:1906.08101.
- Daraz, L., Morrow, A. S., Ponce, O. J., Farah, W., Katabi, A., Majzoub, A., & Larry, P. (2018). Readability of online health information: A meta-narrative systematic review. *American Journal of Medical Quality*, 33(5), 487–492.
- De Clercq, D., Wen, Z., & Song, Q. (2019). Innovation hotspots in food waste treatment, biogas, and anaerobic digestion technology: A natural language processing approach. *The Science of the Total Environment*, 673, 402–413.
- Deng, S., Fu, S., Liu, Y., & Li, H. (2021). Modelling users' trust in online health rumours: An experiment-based study in China. *INFORMATION RESEARCH-AN INTERNATIONAL ELECTRONIC JOURNAL*, 26(1), 890.
- Derczynski, L., Bontcheva, K., Liakata, M., Procter, R., Hoi, G. W. S., & Zubiaga, A. (2017). *SemEval-2017 task 8: RumourEval: Determining rumour veracity and support for rumours*. arXiv preprint arXiv:1704.05972.
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint*. arXiv:1810.04805.
- DiFonzo, N., & Bordia, P. (2000). How top PR professionals handle hearsay: Corporate rumors, their effects, and strategies to manage them. *Public Relations Review*, 26(2), 173–190.
- DiFonzo, N., & Bordia, P. (2007). *Rumor psychology: Social and organizational approaches*. American Psychological Association.
- DiFonzo, N., Bordia, P., & Rosnow, R. L. (1994). Reining in rumors. *Organizational Dynamics*, 23(1), 47–62.
- DuBay, W. H. (2004). *The principles of readability*. Online Submission.
- Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt brace Jovanovich college publishers.
- Erku, D. A., Belachew, S. A., Abrha, S., Sinnollareddy, M., Thomas, J., Steadman, K. J., & Tesfaye, W. H. (2021). When fear and misinformation go viral: Pharmacists' role in deterring medication misinformation during the 'infodemic' surrounding COVID-19. *Research in Social and Administrative Pharmacy*, 17(1), 1954–1963.
- Falcone, R., & Sapienza, A. (2020). How COVID-19 changed the information needs of Italian citizens. *International Journal of Environmental Research and Public Health*, 17(19), 6988.
- Gallotti, R., Valle, F., Castaldo, N., Sacco, P., & De Domenico, M. (2020). Assessing the risks of 'infodemics' in response to COVID-19 epidemics. *Nature Human Behaviour*, 11.
- Gans, R. (2014). The politics of HPV vaccination advocacy: Effects of source expertise on effectiveness of a pro-vaccine message. *Proceedings of the New York State Communication Association*, 3, 2013(2013).
- Gefen, D., Straub, D., & Boudreau, M. C. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the Association for Information Systems*, 4(1), 7.
- Goh, D., & Chi, J. (2017). Central or peripheral? Information elaboration cues on childhood vaccination in an online parenting forum. *Computers in Human Behavior*, 69, 181–188.
- Gonzalez, C., Thomas, R. P., & Vanyukov, P. (2005). The relationships between cognitive ability and dynamic decision making. *Intelligence*, 33(2), 169–186.
- Griffith, E. E., Nolder, C. J., & Petty, R. E. (2018). The elaboration likelihood model: A meta-theory for synthesizing auditor judgment and decision-making research. *Auditing: A Journal of Practice & Theory*, 37(4), 169–186.
- 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.
- Guo, L. (2020). *China's "fake news" problem: Exploring the spread of online rumors in the government-controlled news media*.
- Hachaturyan, V., Adam, M., Favaretti, C., Greuel, M., Gates, J., Bärnighausen, T., & Vandormael, A. (2021). Reactance to social authority in a sugar reduction informational video: Web-based randomized controlled trial of 4013 participants. *Journal of Medical Internet Research*, 23(11), Article e29664.
- Hair, J. F. (2009). *Multivariate data analysis*.
- Hamilton, M. A., Hunter, J. E., & Boister, F. J. (1993). The elaboration likelihood model as a theory of attitude formation: A mathematical analysis. *Communication Theory*, 3(1), 50–65.
- Han, J. T., Chen, Q., Liu, J. G., Luo, X. L., & Fan, W. (2018). The persuasion of borrowers' voluntary information in peer to peer lending: An empirical study based on elaboration likelihood model. *Computers in Human Behavior*, 78, 200–214.
- Horn, E. (2011). Logics of political secrecy. *Theory, Culture & Society*, 28(7–8), 103–122.
- Houdek, P., Koblovsky, P., & Vranka, M. (2021). *The challenge of human psychology to effective management of the COVID-19 pandemic*. Society.
- 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.
- Hui, H., Zhou, C., Lü, X., & Li, J. (2020). Spread mechanism and control strategy of social network rumors under the influence of COVID-19. *Nonlinear Dynamics*, 101(3), 1933–1949.
- Invernizzi, F., Falomir-Pichastor, J. M., Muñtöz-Rojas, D., & Mugny, G. (2003). Social influence in personally relevant contexts: The respect attributed to the source as a factor increasing smokers' intention to quit smoking 1. *Journal of Applied Social Psychology*, 33(9), 1818–1836.
- Iyer, E. S., & Debevec, K. (1991). Origin of rumor and tone of message in rumor quelling strategies. *Psychology and Marketing*, 8(3), 161–175.
- Jang, J. W., Lee, E. J., & Shin, S. Y. (2019). What debunking of misinformation does and doesn't. *Cyberpsychology, Behavior, and Social Networking*, 22(6), 423–427.
- Ji, K., Liu, J., & Xiang, G. (2014). Anti-rumor dynamics and emergence of the timing threshold on complex network. *Physica A: Statistical Mechanics and its Applications*, 411, 87–94.
- Jung, J. M., & Kellaris, J. J. (2006). Responsiveness to authority appeals among young French and American consumers. *Journal of Business Research*, 59(6), 735–744.
- Jung, J. M., Polyorat, K., & Kellaris, J. J. (2009). A cultural paradox in authority-based advertising. *International Marketing Review*, 26(6), 601–632.
- Jung, A. K., Ross, B., & Stieglitz, S. (2020). Caution: Rumors ahead-A case study on the debunking of false information on Twitter. *Big Data & Society*, 7(2), 15.
- Kailath, T. (1967). The divergence and Bhattacharyya distance measures in signal selection. *IEEE Transactions on Circuit Theory*, 15(1), 52–60.
- Kitchen, P. J., Kerr, G., Schultz, D. E., McColl, R., & Pals, H. (2014). The elaboration likelihood model: Review, critique and research agenda. *European Journal of Marketing*, 48(11/12), 2033–2050.
- Korfatis, N., García-Bariocanal, E., & Sánchez-Alonso, S. (2012). Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content. *Electronic Commerce Research and Applications*, 11(3), 205–217.
- Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1), 79–86.
- Laato, S., Islam, A. N., Islam, M. N., & Whelan, E. (2020). What drives unverified information sharing and cyberchondria during the COVID-19 pandemic? *European Journal of Information Systems*, 29(3), 288–305.
- Lazer, D., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., ... Rothschild, D. (2018). The science of fake news. *Science*, 359(6380), 1094–1096.
- Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13(3), 106–131.
- Li, C. Y. (2013). Persuasive messages on information system acceptance: A theoretical extension of elaboration likelihood model and social influence theory. *Computers in Human Behavior*, 29(1), 264–275.
- Li, B., Hou, F., Guan, Z., Chong, Y. L., & Pu, X. (2017). Evaluating online review helpfulness based on elaboration likelihood model: The moderating role of readability. In *Pacific Asia conference on information systems*.
- Lin, X., Spence, P. R., & Lachlan, K. A. (2016). Social media and credibility indicators: The effect of influence cues. *Computers in Human Behavior*, 63, 264–271.
- Liu, Z., Liu, L., & Li, H. (2012). Determinants of information retweeting in microblogging. *Internet Research*, 22(4).
- Liu, Z., & Park, S. (2015). What makes a useful online review? Implication for travel product websites. *Tourism Management*, 47, 140–151.
- Li, Z., Zhang, Q., Du, X., Ma, Y., & Wang, S. (2021). Social media rumor refutation effectiveness: Evaluation, modelling and enhancement. *Information Processing & Management*, 58(1), 102420.
- Lodge, M., & Hamill, R. (1986). A partisan schema for political information processing. *American Political Science Review*, 505–520.
- Lubinski, D. (2009). Exceptional cognitive ability: The phenotype. *Behavior Genetics*, 39(4), 350–358.
- Mak, B., Schmitt, B. H., & Lyytinen, K. (1997). User participation in knowledge update of expert systems. *Information & Management*, 32(2), 55–63.
- van der Meer, T. G., & Jin, Y. (2020). Seeking formula for misinformation treatment in public health crises: The effects of corrective information type and source. *Health Communication*, 35(5), 560–575.
- Metzger, M. J. (2007). Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*, 58(13), 2078–2091.
- Metzger, M. J., & Flanagin, A. J. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics*, 59, 210–220.

- Morris, J. D., Woo, C., & Singh, A. J. (2005). Elaboration likelihood model: A missing intrinsic emotional implication. *Journal of Targeting, Measurement and Analysis for Marketing*, 14(1), 79–98.
- Musuva, P. M., Getao, K. W., & Chepken, C. K. (2019). A new approach to modelling the effects of cognitive processing and threat detection on phishing susceptibility. *Computers in Human Behavior*, 94, 154–175.
- Naem, S. B., Bhatti, R., & Khan, A. (2021). An exploration of how fake news is taking over social media and putting public health at risk. *Health Information and Libraries Journal*, 38(2), 143–149.
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, 2(2), 175–220.
- Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32(2), 303–330.
- Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of Advertising*, 19(3), 39–52.
- Osatuyi, B., & Hughes, J. (2018). A tale of two internet news platforms—real vs. fake: An elaboration likelihood model perspective. In *Proceedings of the 51st Hawaii international conference on system sciences*.
- Ozbay, F. A., & Alatas, B. (2020). Fake news detection within online social media using supervised artificial intelligence algorithms. *Physica A: Statistical Mechanics and Its Applications*, 540, 123174.
- Paek, H. J., & Hove, T. (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), 2653–2667.
- Pal, A., Chua, A. Y., & Goh, D. H. L. (2020). How do users respond to online rumor rebuttals? *Computers in Human Behavior*, 106, 106243.
- Pallak, S. R., Murrioni, E., & Koch, J. (1983). Communicator attractiveness and expertise, emotional versus rational appeals, and persuasion: A heuristic versus systematic processing interpretation. *Social Cognition*, 2(2), 122–141.
- Pavey, L., Churchill, S., & Sparks, P. (2021). Proscriptive injunctions can elicit greater reactance and lower legitimacy perceptions than prescriptive injunctions. *Personality and Social Psychology Bulletin*, Article 01461672211021310.
- Pennycook, G., Bear, A., Collins, E. T., & Rand, D. G. (2020). The implied truth effect: Attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. *Management Science*, 66(11), 4944–4957.
- Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, 188, 39–50.
- Petty, R. E., & Cacioppo, J. T. (1986). *The elaboration likelihood model of persuasion: Communication and persuasion*. Springer.
- Petty, R. E., & Cacioppo, J. T. (2012). *Communication and persuasion: Central and peripheral routes to attitude change*. Springer Science & Business Media.
- Preston, S., Anderson, A., Robertson, D. J., Shephard, M. P., & Huhe, N. (2021). Detecting fake news on Facebook: The role of emotional intelligence. *PLoS One*, 16(3), Article e0246757.
- Principe, V. A., de Souza Vale, R. G., de Castro, J. B. P., Carvano, L. M., Henriques, R. A. P., e Sousa, V. J. d. A., & Nunes, R. d. A. M. (2021). A computational literature review of football performance analysis through probabilistic topic modeling. *Artificial Intelligence Review*, 1–21.
- Qing, K., Songyun, D., & Qin, Q. (2020). The impacts of readability of health information on users' cognitive load and information processing performance: An eye-tracking study. *Data Analysis and Knowledge Discovery*, 1.
- Robert, L. P., & Dennis, A. R. (2005). Paradox of richness: A cognitive model of media choice. *IEEE Transactions on Professional Communications*, 48(1), 10–21.
- Rodríguez, C. P., Carballido, B. V., Redondo-Sama, G., Guo, M., Ramis, M., & Flecha, R. (2020). False news around COVID-19 circulated less on Sina Weibo than on Twitter. How to overcome false information? *International and Multidisciplinary Journal of Social Sciences*, 9(2), 107–128.
- Roets, A. (2017). 'Fake news': Incorrect, but hard to correct. The role of cognitive ability on the impact of false information on social impressions. *Intelligence*, 65, 107–110.
- Rosen, C. S. (2000). Integrating stage and continuum models to explain processing of exercise messages and exercise initiation among sedentary college students. *Health Psychology*, 19(2), 172.
- Rucker, D. D., & Petty, R. E. (2006). Increasing the effectiveness of communications to consumers: Recommendations based on elaboration likelihood and attitude certainty perspectives. *Journal of Public Policy and Marketing*, 25(1), 39–52.
- Rui, H., Liu, Y., & Whinston, A. (2013). Whose and what chatter matters? The effect of tweets on movie sales. *Decision Support Systems*, 55(4), 863–870.
- Senter, R. J., & Smith, E. A. (1967). *Automated readability index*. CINCINNATI UNIV OH.
- Shahab, M. H., Ghazali, E., & Mohtar, M. (2021). The role of elaboration likelihood model in consumer behaviour research and its extension to new technologies: A review and future research agenda. *International Journal of Consumer Studies*, 45(4), 664–689.
- Shrivastava, G., Kumar, P., Ojha, R. P., Srivastava, P. K., Mohan, S., & Srivastava, G. (2020). Defensive modeling of fake news through online social networks. *IEEE Transactions on Computational Social Systems*, 7(5), 1159–1167.
- Singh, V. K., Ghosh, I., & Sonagara, D. (2021). Detecting fake news stories via multimodal analysis. *Journal of the Association for Information Science and Technology*, 72(1), 3–17.
- Song, Y., Kwon, K. H., Lu, Y., Fan, Y., & Li, B. (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, Article 00027642211003153.
- Srivastava, V., & Kalro, A. D. (2019). Enhancing the helpfulness of online consumer reviews: The role of latent (content) factors. *Journal of Interactive Marketing*, 48, 33–50.
- Stephenson, M. T., Benoit, W. L., & Tschida, D. A. (2001). Testing the mediating role of cognitive responses in the elaboration likelihood model. *Communication Studies*, 52(4), 324–337.
- Sui, Y. J., & Zhang, B. (2021). Determinants of the perceived credibility of rebuttals concerning health misinformation. *International Journal of Environmental Research and Public Health*, 18(3), 17.
- Susarla, A., Oh, J. H., & Tan, Y. (2012). Social networks and the diffusion of user-generated content: Evidence from YouTube. *Information Systems Research*, 23(1), 23–41.
- Sussman, S. W., & Siegal, W. S. (2003). Informational influence in organizations: An integrated approach to knowledge adoption. *Information Systems Research*, 14(1), 47–65.
- Swire, B., Ecker, U. K., & Lewandowsky, S. (2017). The role of familiarity in correcting inaccurate information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(12), 1948.
- Szmuda, T., Özdemir, C., Ali, S., Singh, A., Syed, M., & Słoniewski, P. (2020). Readability of online patient education material for the novel coronavirus disease (COVID-19): A cross-sectional health literacy study. *Public Health*, 185, 21–25.
- Tambuscio, M., Ruffo, G., Flammini, A., & Menczer, F. (2015). Fact-checking effect on viral hoaxes: A model of misinformation spread in social networks. In *Paper presented at the Proceedings of the 24th international conference on world wide web*.
- Tanaka, Y., Sakamoto, Y., & Matsuka, T. (2013). Toward a social-technological system that inactivates false rumors through the critical thinking of crowds. In *Paper presented at the Hawaii international conference on system sciences*.
- Tian, X., Yu, G., & He, F. (2016). An analysis of sleep complaints on Sina Weibo. *Computers in Human Behavior*, 62, 230–235.
- Van Asch, V. (2013). *Macro-and micro-averaged evaluation measures* (Vol. 49). Belgium: CLIPS.
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151.
- Vraga, E. K., Tully, M., Maksl, A., Craft, S., & Ashley, S. (2021). Theorizing news literacy behaviors. *Communication Theory*, 31(1), 1–21.
- Wang, R., He, Y., Xu, J., & Zhang, H. (2020). Fake news or bad news? Toward an emotion-driven cognitive dissonance model of misinformation diffusion. *Asian Journal of Communication*, 30(5), 317–342.
- Wang, C., & Huang, H. (2021). When "fake news" becomes real: The consequences of false government denials in an authoritarian country. *Comparative Political Studies*, 54(5), 753–778.
- Wang, Z., Liu, H., Liu, W., & Wang, S. (2020). Understanding the power of opinion leaders' influence on the diffusion process of popular mobile games: Travel Frog on Sina Weibo. *Computers in Human Behavior*, 109, 106354.
- Wang, D., & Qian, Y. (2021). Echo chamber effect in the discussions of rumor rebuttal about COVID-19 in China: Existence and impact. *Journal of Medical Internet Research*, 23(3), e27009.
- Wang, Q. S., Yang, X., & Xi, W. Y. (2018). Effects of group arguments on rumor belief and transmission in online communities: An information cascade and group polarization perspective. *Information & Management*, 55(4), 441–449. <https://doi.org/10.1016/j.im.2017.10.004>
- Wechsler, D. (1958). *The measurement and appraisal of adult intelligence*.
- Wen, S., Jiang, J., Xiang, Y., Yu, S., Zhou, W., & Jia, W. (2014). To shut them up or to clarify: Restraining the spread of rumors in online social networks. *IEEE Transactions on Parallel and Distributed Systems*, 25(12), 3306–3316.
- Westerman, D., Spence, P. R., & Van Der Heide, B. (2014). Social media as information source: Recency of updates and credibility of information. *Journal of Computer-Mediated Communication*, 19(2), 171–183.
- World Health Organisation. (2020a). *Novel Coronavirus (2019-nCoV) Situation Report-13*. Geneva, Switzerland: WHO.
- World Health Organisation. (2020b). *Novel Coronavirus (2019-nCoV) situation Report-86*. Geneva, Switzerland: WHO.
- Worrall, A. P., Connolly, M. J., O'Neill, A., O'Doherty, M., Thornton, K. P., McNally, C., & de Barra, E. (2020). Readability of online COVID-19 health information: A comparison between four English speaking countries. *BMC Public Health*, 20(1), 1–12.
- Xing, Y., Wang, X., Wang, F.-K., Shi, Y., He, W., & Chang, H. (2021). Ranking influential nodes of fake news spreading on mobile social networks. *Journal of Global Information Management*, 29(4), 93–130.
- Yada, N., & Head, M. (2019). Attitudes toward health care virtual communities of practice: Survey among health care workers. *Journal of Medical Internet Research*, 21(12), Article e15176.
- Yu, R. P. (2021). How types of facebook users approach news verification in the mobile media age: Insights from the dual-information-processing model. *Mass Communication & Society*, 24(2), 233–258.
- Zhang, X., & Ghorbani, A. A. (2020). An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management*, 57(2), 102025.
- Zhang, L., Peng, T.-Q., Zhang, Y.-P., Wang, X.-H., & Zhu, J. J. (2014). Content or context: Which matters more in information processing on microblogging sites. *Computers in Human Behavior*, 31, 242–249.
- Zhang, Y., Su, Y., Li, W., & Liu, H. (2018). Modeling rumor propagation and refutation with time effect in online social networks. *International Journal of Modern Physics C*, 29, 1850068, 08.
- Zhang, D., Wang, J., & Zhao, X. (2015). Estimating the uncertainty of average F1 scores. In *Proceedings of the 2015 international conference on the theory of information retrieval* (pp. 317–320). September.
- Zhang, K. Z., Zhao, S. J., Cheung, C. M., & Lee, M. K. (2014). Examining the influence of online reviews on consumers' decision-making: A heuristic-systematic model. *Decision Support Systems*, 67, 78–89.

- Zhao, Y., Da, J., & Yan, J. (2021). Detecting health misinformation in online health communities: Incorporating behavioral features into machine learning based approaches. *Information Processing & Management*, 58(1), 102390.
- Zha, X., Yang, H., Yan, Y., Liu, K., & Huang, C. (2018). Exploring the effect of social media information quality, source credibility and reputation on informational fit-to-task: Moderating role of focused immersion. *Computers in Human Behavior*, 79, 227–237.
- Zhou, L., Zhang, Z., Zhao, L., & Yang, P. (2021). How user personality and information characteristics influence the creative information quality on open innovation platforms: An elaboration likelihood model. *Kybernetes*.
- Zhu, L., Yin, G., & He, W. (2014). Is this opinion leader's review useful? Peripheral cues for online review helpfulness. *Journal of Electronic Commerce Research*, 15(4), 267.
- Zubiaga, A., Kochkina, E., Liakata, M., Procter, R., Lukasiak, M., Bontcheva, K., ... Augenstein, I. (2018). Discourse-aware rumour stance classification in social media using sequential classifiers. *Information Processing & Management*, 54(2), 273–290.