



OPEN ACCESS

EDITED BY
Wadii Boulila,
Manouba University, Tunisia

REVIEWED BY
Peishan Ning,
Central South University, China
Anjan Pal,
University of York, United Kingdom
Alessandro Rovetta,
R&C Research, Italy

*CORRESPONDENCE
Cun Fu
fucun@cqu.edu.cn

SPECIALTY SECTION
This article was submitted to
Digital Public Health,
a section of the journal
Frontiers in Public Health

RECEIVED 05 May 2022
ACCEPTED 18 July 2022
PUBLISHED 10 August 2022

CITATION
Zhao J, Fu C and Kang X (2022)
Content characteristics predict the
putative authenticity of COVID-19
rumors.
Front. Public Health 10:920103.
doi: 10.3389/fpubh.2022.920103

COPYRIGHT
© 2022 Zhao, Fu and Kang. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Content characteristics predict the putative authenticity of COVID-19 rumors

Jingyi Zhao¹, Cun Fu^{2*} and Xin Kang^{2,3}

¹College of International Studies, Southwest University, Chongqing, China, ²School of Foreign Languages and Cultures, Chongqing University, Chongqing, China, ³Research Center for Language, Cognition and Language Application, Chongqing University, Chongqing, China

Rumors regarding COVID-19 have been prevalent on the Internet and affect the control of the COVID-19 pandemic. Using 1,296 COVID-19 rumors collected from an online platform (piyao.org.cn) in China, we found measurable differences in the content characteristics between true and false rumors. We revealed that the length of a rumor's headline is negatively related to the probability of a rumor being true [odds ratio (OR) = 0.37, 95% CI (0.30, 0.44)]. In contrast, the length of a rumor's statement is positively related to this probability [OR = 1.11, 95% CI (1.09, 1.13)]. In addition, we found that a rumor is more likely to be true if it contains concrete places [OR = 20.83, 95% CI (9.60, 48.98)] and it specifies the date or time of events [OR = 22.31, 95% CI (9.63, 57.92)]. The rumor is also likely to be true when it does not evoke positive or negative emotions [OR = 0.15, 95% CI (0.08, 0.29)] and does not include a call for action [OR = 0.06, 95% CI (0.02, 0.12)]. By contrast, the presence of source cues [OR = 0.64, 95% CI (0.31, 1.28)] and visuals [OR = 1.41, 95% CI (0.53, 3.73)] is related to this probability with limited significance. Our findings provide some clues for identifying COVID-19 rumors using their content characteristics.

KEYWORDS

logistic regression model, content characteristics, authenticity, COVID-19 rumors, linguistic

Introduction

Rumors are unverified information that circulates online and offline (1–3). Rumors regarding COVID-19 can be unverified facts, misunderstandings of facts, a pursuit of factual information, a question on current policies, or deliberate deception (2, 4–7). A recent poll in the United Kingdom showed that 46% of citizens came across rumors about COVID-19 (8). Similarly, the Pew Charitable Trusts in the United States indicated that 48% of the population had been exposed to such rumors (9).

Coronavirus disease 2019 rumors in China can be grouped into five categories: the nature of the virus, pandemic areas and confirmed cases, COVID-19 policies, authorities and organizations (e.g., WHO), and medical supplies (10). Rumors undermine the government's efforts to control the pandemic because many social media users cannot discern their authenticity (3, 11). For example, a recent British study indicated that only 4% of participants could tell fake daily news from real ones (12).

Our study explores whether the content characteristics of COVID-19 rumors in China predict the likelihood of their veracity. We have examined information length (headline and statement), information specification (place, time, source cue, and visual), and information effectiveness (emotion and call for action) as predictors of the veracity of rumors.

Literature review

Rumors spread along with the COVID-19 pandemic and hinder the public from making informed decisions. To debunk COVID-19 rumors, many checklists, guidelines, fact-checking tools, digital training programs for health literacy, and long-term strategic plans have been implemented (11). However, these measures were not as effective as expected because they failed to help the public discern false rumors using “rules of thumb” (13–15). Moreover, although previous studies advocated improving social media users’ health literacy, they did not examine the content characteristics of health rumors (16–19).

By contrast, researchers have attempted to use the content characteristics of deceptive information to discern truthful and deceptive statements (20, 21). For instance, Fuller et al. (22) demonstrated that word count was a significant indicator of deceptive information. Huang et al. (23) further revealed that deceivers were most likely to offer more information to convince receivers. Similarly, Zhou et al. (24) and Luo et al. (25) found that the false statement was longer because deceivers must provide supporting evidence and details to persuade their receivers to construe the message. In addition, some studies have examined misinformation detection using content-based features such as paralinguistic features (10, 26–29). For example, Qazvinian et al. (27) reported that tweets with more hashtags were more likely to be misinformation than those without them.

Nonetheless, few studies have examined the content characteristics of rumors (29, 30). Zhou and Zhang (31) found that false rumors were more likely to have lower lexical diversity and contained more uncertain words than true rumors. Luca and Zervas (32) showed that fake reviews on Yelp tended to use more extreme words and expressions than real reviews. In addition, Zhang et al. (33) revealed that the presence of numbers was a good indicator for rumors. Chua and Banerjee (34) indicated that information using more exclusive words was more likely to be false. Furthermore, previous studies investigated whether the frequency of question marks, sentiment markers, arbitrary words, and tentative words (e.g., “maybe”) could differentiate false rumors from true ones (10, 13, 35, 36).

According to information manipulation theory (37), deceptive information misleads receivers by covertly violating the principles of quantity, quality, manner, and relevance. Thus, deceivers may create rumors by manipulating these principles. Indeed, some online rumors are generated intentionally by rumor creators who use specific writing styles with unique

content characteristics to avoid being detected as rumors (9, 38). However, studies on the content characteristics of COVID-19 rumors are particularly scarce (29, 39). In addition, since previous studies are diverse in objectives, methodologies, and topics, they are insufficient for establishing a reliable and effective rumor detection system or “rules of thumb” for debunking COVID-19 rumors (35, 40).

Our study aims to address these research gaps. Following the framework proposed by Zhang and Ghorbani (29), we define the content characteristics of rumors as fundamental components of the natural language and categorize them into three types: information length, information specification, and information effectiveness. Information length refers to the number of words in the headline and the statement of rumors. Information specification refers to details of the content, including place and time of events, source cues of the information, and the use of visuals. Information effectiveness refers to emotions that rumors intend to evoke and actions that receivers are expected to take.

Hypotheses

Hypothesis 1 (H1): Information with a longer headline or statement is more likely to be false.

Information manipulation theory states that deceptive information often violates the conversational principle of quantity by altering the amount of the information supplied to receivers (37). Word count is known as a significant predictor of deceptive information (22), but the directionality was not consistent in previous studies. Huang et al. (23) found that deceivers were most likely to offer more information to convince receivers. Similarly, Zhou et al. (24) and Luo et al. (25) found that deceptive information had a longer length than truthful ones since deceivers had to provide supporting evidence and details to persuade their receivers to construe this information as true. They also showed that information length was a good predictor for distinguishing true and false rumors in various topics, including politics, science, and public health. In addition, Zhang et al. (33) found that the length of both the headline and the statement was associated with the authenticity of online health rumors in China. The longer the headline and the statement, the more likely the rumor was false.

Hypothesis 2 (H2): Information with an ambiguous place or time of events and an ambiguous source cue is more likely to be false.

Several studies have indicated that deceptive statements contain less detailed content than truthful ones (20, 24). Deceivers may lack actual experiences and, thus, may be unable to provide detailed information. For instance, Zhang et al. (33) revealed that the presence of numbers was a good indicator of true rumors. Bond and Lee (20) found that deceptive statements were less likely to contain sensory and temporal vocabularies.

Banerjee and Chua (11) found that deceptive online reviews had more diminutive nouns than authentic ones.

Moreover, online rumors often lack concrete source cues. Previous studies have focused on governmental organizations and news agencies such as the Xinhua News of China or the Cable News Network (CNN) of the USA (1, 33, 41). For example, Zhang et al. (33) examined domestic and foreign source cues and found that rumors with either type of source cues were positively related to the probability of being true. Recently, some researchers have proposed that informants in rumors can be further divided into ambiguous or concrete source cues and be used to evaluate the veracity of rumors (41–46). For instance, the source cue in the rumor, “A doctor of Peking Union Medical College Hospital recently published an article to confirm that clearing nose with normal saline could prevent COVID-19 virus,” refers to an ambiguous person—“a doctor.” By comparison, the source cue in the rumor “Zhong Nanshan announced that sequela of COVID-19 was more severe than that of a severe acute respiratory syndrome (SARS)” is concrete—Zhong Nanshan (a distinguished respiratory specialist in China).

Hypothesis 3 (H3): Information without visuals such as pictures or videos is more likely to be false.

Coronavirus disease 2019 rumors are often presented with visuals such as pictures and videos (14, 15, 47–51). Visuals are expected to represent reality because people tend to believe “seeing is believing” and “a picture is worth a thousand words” (52). Previous studies have found that pictures were presented with truthful information and sometimes were used to increase the perceived authenticity of rumors (14, 33, 53, 54). However, Zhang et al. (33) found that the presence of pictures did not differentiate true health rumors from false ones.

Hypothesis 4 (H4): Information that elicits positive or negative emotions and contains a call for action is more likely to be false.

Coronavirus disease 2019 rumors may have a long-term impact on people’s emotional states by eliciting negative emotions such as anxiety, helplessness, anger, and discomfort (55–57). Nonetheless, there is no consensus on whether elicited emotions are linked with the veracity of rumors. (5) reported that sentiments predicted the veracity of rumors on Twitter. Similarly, Zhang et al. (33) found that dreadful health rumors that described fearsome, disappointing, and undesirable events or outcomes were more prevalent and were more likely to be accurate than wish rumors that described potential positive consequences. In contrast, Chua and Banerjee (34) argued that sentiments elicited by rumors could not predict the veracity.

Furthermore, providing false information is not the only negative impact of false rumors. Many COVID-19 rumors may end with a call for action, such as transmitting the message to friends or relatives (39). This call for action may also motivate receivers to act against COVID-19 policies, cause panic among

the public, and urge receivers not to take vaccines (13, 21, 45, 58–60). However, whether a call for action is linked to the veracity of rumors is not yet known.

Method

In this study, we collected rumors listed on *piyao.org.cn* (hereafter known as *piyao*), an official rumor-debunking platform run by the Xinhua News of China and the Cyberspace Administration of China. This website collects rumors from 31 other major rumor-debunking platforms in China, including Zhuoyaoji (捉妖记), Wenzhoupiyao (温州辟谣), and real-time refutation of COVID-19 rumors (新冠实时辟谣). According to the 47th China Statistical Report on Internet Development (61), *piyao* releases and refutes most rumors spreading in China and attracts 100 million visits yearly. Like Snopes.com, *piyao* allows online users to submit rumors. With the help of professionals, researchers, and reporters, *piyao* sets the record straight on every rumor that it collects by rating it as “true,” “false,” or “undetermined” (33, 62). In addition, facing the COVID-19 pandemic, *piyao* added a section dedicated to “COVID-19 rumor-debunking” and has become an authoritative platform for debunking COVID-19 rumors. A rumor on *piyao* is presented in its original form (including a headline and a statement, and sometimes visuals if it contains) along with a veracity rating, checked facts, and detailed analyses of the information.

In total, 1,685 COVID-19 rumors were collected from *piyao* in January 2022. All the rumors had headlines and statements. We assume that veracity ratings from *piyao* represent the truth. Three hundred eighty-nine rumors were categorized as “undetermined” and, thus, excluded, leaving 1,296 rumors in the final data analysis. Table 1 presents the examples of rumors and their content characteristics. Table 2 shows the coding schemes. Two coders (JZ and CF) undertook the coding in two phases. In the first phase, they worked independently on 168 randomly-selected rumors and then resolved their disagreements (if any) through in-person discussions. The two coders demonstrated almost perfect inter-rater reliability for all the measures as indicated by Cohen’s kappa (k) (place: $k = 0.95, p < 0.001$; time: $k = 0.98, p < 0.001$; source cue: $k = 0.95, p < 0.001$; visual: $k = 1.00, p < 0.001$; emotion: $k = 0.95, p < 0.001$; call for action: $k = 0.96, p < 0.001$). In the second phase, they coded the remaining rumors that were assigned to them randomly.

Results

Tables 3–5 present the descriptive statistics of our data. In total, there were 82% false rumors and 18% true rumors out of 1,296 rumors in our sample. The Shapiro–Wilk test showed that the headline ($W = 0.95, p < 0.001$) and the statement ($W = 0.70, p < 0.001$) are not normally distributed. We first conducted

TABLE 1 Examples of COVID-19 rumors and their content characteristics.

Content characteristics		Description	Example
Information length	Headline	The number of Chinese characters in a rumor headline	Attention! Attention! A technique to get! Medical doctors confirmed that normal saline to prevent COVID-19. (注意! 注意! 教大家一个新技能! 医生证实生理盐水可以预防新冠病毒。)
	Statement	The number of Chinese characters in a rumor statement	Surprise! A Peking Union Medical College Hospital doctor recently published a paper confirming that clearing the nose with normal saline could prevent the COVID-19 virus. To protect your family, please retransmit this information to your relatives and friends with your fingers. (惊喜! 北京协和医院的一个医生近日发文, 可以用生理盐水清洗鼻子, 这样能有效预防新型冠状病毒感染。为了让你的家人远离新冠病毒, 动动你的手指, 转发给你的家人和朋友。)
Information specification	Place	Concrete or ambiguous place	Peking Union Medical College Hospital (北京协和医院)
	Time	Concrete or ambiguous time	recently (近日)
	Source cue	Concrete or ambiguous source cues	a doctor (一个医生)
	Visual	Presence of pictures or videos along with the text	pictures/videos
Information affectiveness	Emotion	A positive or negative emotion that expresses hope, happiness, relief, disgust, anger, fear, and anxiety	protect your family (保护家人)
	Call for action	A call for action such as sending this message to friends	Please retransmit this information to your relatives and friends (动动你的手指, 转发给你的家人和朋友)

TABLE 2 Coding schemes for the content characteristics of COVID-19 rumors.

Variable	Coding scheme
Veracity	True rumors = 1, False rumors = 0
Headline	Number of Chinese characters in the headline
Statement	Number of Chinese characters in the statement
Place	Concrete place = 1, Ambiguous place = 0
Time	Concrete time = 1, Ambiguous time = 0
Source cue	Concrete source cue = 1, Ambiguous source cue = 0
Visual	With visual = 1, Without visual = 0
Emotion	With emotion =1, Without emotion = 0
Call for action	With a call for action=1, Without a call for action = 0

preliminary analyses to examine the associations between the veracity of rumors and each independent variable. For the headline and the statement, we ran the Wilcoxon rank-sum test to examine whether true and false rumors showed significant differences in the number of Chinese characters. We found that the headline of the false rumor contained significantly more Chinese characters ($Mdn = 12$) than the true rumor ($Mdn =$

$8, p < 0.001, r = -0.42$). The statements of the false rumor ($Mdn = 78$) contained significantly fewer Chinese characters than the true rumor ($Mdn = 120, p < 0.001, r = -0.53$). In addition, separate Pearson's chi-squared test showed that there were significant associations between the veracity of rumors and whether the information contained a concrete place ($X^2 = 37.85, p < 0.001$) with an odds ratio of 2.45, a concrete time ($X^2 = 48.44, p < 0.001$) with an odds ratio of 2.74, a concrete source cue ($X^2 = 39.06, p < 0.001$) with an odds ratio of 2.46, a visual ($X^2 = 12.65, p < 0.001$) with an odds ratio of 0.51, an emotion ($X^2 = 148.24, p < 0.001$) with an odds ratio of 0.17, and a call for action ($X^2 = 140.17, p < 0.001$) with an odds ratio of 0.17.

However, the above preliminary analyses could not include all the independent variables simultaneously. We, thus, further used logistic regression to examine the relationship between the veracity of rumors and eight independent variables in a single model. Our analysis met all the assumptions of using the logistic regression, including: (1) the response variable (veracity) is binary, (2) the observations are independent, (3) there is no multicollinearity among explanatory variables as indicated by low variance inflation factor (VIF) values (headline: 5.00; statement: 3.88; place: 1.82; time: 2.16; source cue: 1.34; visual: 1.16; emotion:1.08; and call for action: 1.82), (4) there are no

TABLE 3 Descriptive statistics of information length.

Variable	Full dataset (n = 1,296)			True rumors (n = 232)			False rumors (n = 1,064)		
	Range	Mean	Median	Range	Mean	Median	Range	Mean	Median
Headline	[5, 28]	13.16	12	[5, 17]	8.73	8	[6, 28]	14.12	12
Statement	[17, 330]	92.38	80	[78, 330]	143.03	120	[17, 201]	81.37	78

TABLE 4 Descriptive statistics of information specification.

Variable	Full dataset (n = 1,296)		True rumors (n = 232)		False rumors (n = 1,064)	
	% (Value=1)	% (Value=0)	% (Value=1)	% (Value=0)	% (Value=1)	% (Value=0)
Place	38.67 (501)	61.33 (795)	46.65 (108)	53.35 (124)	36.88 (392)	63.12 (672)
Time	42.32 (548)	57.68 (748)	63.31 (147)	36.69 (85)	37.77 (402)	62.23 (662)
Source cue	35.71 (463)	64.29 (833)	56.68 (132)	43.32 (100)	31.24 (333)	68.76 (731)
Visual	31.04 (402)	68.96 (894)	16.72 (39)	83.28 (193)	34.16 (364)	65.84 (700)

TABLE 5 Descriptive statistics of information affectiveness.

Variable	Full dataset (n = 1,296)		True rumors (n = 232)		False rumors (n = 1,064)	
	% (Value=1)	% (Value=0)	% (Value=1)	% (Value=0)	% (Value=1)	% (Value=0)
Emotion	59.88 (776)	40.12 (520)	24.73 (57)	75.27 (175)	67.01 (713)	32.99 (351)
Call for action	58.95 (764)	41.05 (532)	26.74 (62)	73.26 (170)	65.93 (701)	34.07 (363)

TABLE 6 Logistic regression findings.

Predictors	Odds Ratios	95% CI	p	FDR corrected p
Headline	0.37	0.30–0.44	<0.001	<0.001
Statement	1.11	1.09–1.13	<0.001	<0.001
Place	20.83	9.60–48.98	<0.001	<0.001
Time	22.31	9.63–57.92	<0.001	<0.001
Source	0.64	0.31–1.17	0.214	0.244
Call	0.06	0.02–0.12	<0.001	<0.001
Visual	1.41	0.53–3.73	0.492	0.492
Emotion	0.15	0.08–0.29	<0.001	<0.001

N = 1,296, R² = 0.78.

extreme outliers, and (5) the sample size was determined based on the number of independent variables. We found that a minimum sample size of 500 yields reliable and valid sample estimates (63). Thus, our sample size is sufficient.

Table 6 shows the findings of the logistic regression with all the eight independent variables as predictors for the veracity of COVID-19 rumors in China. The statistical significance of each predictor was corrected using the false discovery rate (FDR) with the Benjamini–Hochberg method. We found that the number

of Chinese characters in a rumor’s headline ($p < 0.001$) and statement ($p < 0.001$) were both significantly related to the veracity, supporting H1. The presence of a concrete place ($p < 0.001$) or time of events ($p < 0.001$) was also significantly related to the veracity of rumors, supporting H2. Nonetheless, source cues ($p = 0.214$) and visual ($p = 0.492$) were associated with the veracity with limited significance, thus not supporting H3. Furthermore, we found that a call for action ($p < 0.001$) was significantly related to the veracity, supporting H4.

Discussion

The Internet is a vital and convenient platform for spreading information about the COVID-19 pandemic (7, 33, 64). However, there is also a sharp increase in COVID-19 rumors. The ratio of true rumors varies across studies in different contexts and topics. Our data showed that false COVID-19 rumors (82%) were much more prevalent than true rumors (18%) in China. This finding is consistent with Zhang et al. (33), who revealed that 75.1% of 453 health-related information was false. It contradicted Gelfert (65) that claimed rumors were often based on facts and were, thus, usually true.

Importantly, our findings support most of our hypotheses. First, although the length of the headline and statement are significant indicators of the veracity of rumors, we found opposite effects. The longer the headline, the more likely the rumor was false. By contrast, the shorter the statement, the more likely the rumor was false. Our finding about the length of the headline is consistent with prior studies (25, 66–69). However, our finding regarding the length of the statement contradicts previous studies (70–72), which revealed that deceptive statements were longer than truthful ones because deceivers attempt to provide more information to increase the perceived credibility of the information. Similarly, Zhang et al. (33) showed that false health-related rumors were longer than true ones because rumormongers have learned that longer statements could reduce uncertainty.

We speculate that there are two possible explanations for these contradictory findings. One possibility is that rumormongers have learned to take advantage of their target readers' dependence on the Internet and smartphones for information acquisition. The long headline is, thus, used to increase the attractiveness and saliency of the information to catch readers' attention. However, as the screen size of smartphones is limited, rumormongers may tend to avoid lengthy details so readers can finish reading the complete statement instead of a piece of fragmented information on one page. The other possibility is that since rumormongers have not witnessed the events, it is difficult for them to provide detailed information. Previous studies have shown that rumormongers could only add peripheral information, but not key details (73–75).

Second, we found that the presence of concrete place and time of events are good indicators of true rumors. Prior studies have revealed that deceptive information included vague spatial and temporal information (10, 39), while true information contained more details. Our findings are consistent with these studies. However, our results differed from Zhang et al. (33) that revealed the presence of a place name was related to the veracity of a rumor with limited significance. However, one crucial difference between Zhang et al. (33) and our study is that

we categorized the “place” into two types: an ambiguous place and a concrete place. We found that only when a concrete place is included in a rumor, it is more likely to be true.

Similarly, we found that rumors containing a concrete time of events such as “on Wednesday” or “in October” rather than an ambiguous time such as “these days” or “recently” are more likely to be true. These findings are consistent with previous studies on deception, which have demonstrated rumormongers may feel strain, guilt, and restlessness in the process (20–22). So, they tended to provide ambiguous information to maintain their distance from the false information receivers (76, 77). Moreover, rumormongers may give such details in a fuzzy, nebulous manner (11, 76) because information receivers could easily use a concrete place and time of events to verify and debunk false rumors.

Third, we did not find any significant association between source cues and the veracity of rumors, which contradicts previous studies (9, 10, 71). Nonetheless, our finding is consistent with other studies that have been conducted in China. For example, Zhou et al. (24) and Jang et al. (39) revealed that rumormongers counterfeited celebrities' sayings with expressions such as “Professor Zhong Nanshan warned that...” or “Professor Zhang Wenhong said...” Despite containing specific source cues, these rumors were often false.

Fourth, we did not find any significant association between the use of visuals and the veracity of rumors. Previous studies argued that rumormongers used visuals to strengthen the perceived credibility of information. For example, Zhang et al. (33) revealed that the use of pictures was negatively related to the veracity of rumors. Nonetheless, visuals may not be related to the veracity of the rumors because rumormongers often included pictures or videos that were not associated with the events (78–80) or used fake pictures or videos (48, 76, 81) as evidence to imitate the practice of news reports.

Fifth, we showed that elicited emotions in rumors could differentiate between true and false rumors. This finding is inconsistent with Zhang et al. (33), which found that health-related rumors evoking positive emotions were more likely to be false than negative ones. Nonetheless, our finding is consistent with Li et al. (82), which demonstrated that heightened emotions were associated with the veracity of the information. Both the positive and negative emotions increase uncritical acceptance of information. These discrepancies may lie in the varied backgrounds of these studies. Our study focuses on COVID-19 rumors that often involve high emotional effectiveness, but Zhang et al. (33) focused on health-related rumors that may not contain a high ratio of emotions.

Finally, we found that the presence of a call for action is associated with false rumors. Previous studies have also shown that calling for counteraction to create chaos could be the purpose of some rumors (83, 84). Whether a rumor contains a call to retweet the information or a call to take counteraction

against governmental prevention and curative policies may present an obstacle to implementing these policies.

Implications and limitations

This study has several theoretical contributions. First, our findings show that information manipulation theory (37) may not apply to all the topics. Deceivers utilize different information manipulation techniques in various topics or fields. For instance, deceivers often create longer statements to enhance perceived credibility, while online COVID-19 rumors usually have shorter statements. Second, this study sheds light on the research field of “using online data to predict human behaviors” by offering a tentative linguistic approach. Our findings show that the content characteristics of rumors predict the veracity of rumors, which may facilitate information receivers to differentiate between false and true rumors. Third, this study extends our current knowledge of information identification by exploring the content characteristics of COVID-19 rumors. A few rules of thumb should be followed when evaluating the veracity ratings of online COVID-19 rumors, with particular attention to the length of headlines and statements, concrete place and time of events, and the presence of emotion and a call for action.

From a practical perspective, the findings of this study can be used as “rules of thumb” to fight against false COVID-19 rumors. Health agencies, organizations, and institutions may use better countermeasures to fight against false-rated rumors. For instance, they may analyze and summarize the content characteristics of each rumor after presenting factual information. This practice may enhance people’s skills in identifying false rumors. In addition, our findings may be used as references for evaluating factual information and for debunking putatively false rumors. We suggest that the content characteristics such as place and time of events should be carefully examined. Furthermore, our study focused exclusively on the content characteristics of COVID-19 rumors in contrast to most prior studies. Facing the prevalence of COVID-19 rumors online, our study offers a set of easy-to-follow guidelines for evaluating the veracity of these rumors by online users.

Our study has some limitations. First, our study only focuses on COVID-19 rumors in China. Our findings may not generalize to COVID-19 rumors spreading in other languages or other countries. Future studies may examine the content characteristics of COVID-19 rumors in different societies and cultures. Another limitation is that our study does not provide any causal links between the content characteristics and the veracity of rumors. In addition, as our study only examines a limited number of the content characteristics of COVID-19 rumors, further studies may explore more features that might be useful in distinguishing true and false rumors using text mining techniques.

Conclusion

Our study provides preliminary evidence on the use of the content characteristics of rumors as guidelines to fight against false COVID-19 rumors. We found that information receivers should pay particular attention to the place and time of events and evaluate whether these rumors include concrete or an ambiguous place and time of events. We also revealed that information receivers should check the length of the headline and the statement, assess whether the information elicits any emotion, and watch out for a call for action. Nonetheless, the presence of source cues or visuals may not help differentiate between true and false rumors.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

CF designed the study. JZ performed the research. CF and JZ analyzed the data. CF, JZ, and XK wrote the paper. All authors declare no competing interest and agree the final version of this manuscript.

Funding

This study was funded by the National Social Science Fund of China (18BYY066), the Fundamental Research Funds for the Central Universities of Chongqing University (2021CDJSKZX10 and 2022CDJSKJC02), and Grants in Humanities and Social Sciences by Chongqing Municipal Education Commission (22SKJD006).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Lewandowsky S, van der Linden S. Countering misinformation and fake news through inoculation and prebunking. *Eur Rev Soc Psychol.* (2021) 32:348–84. doi: 10.1080/10463283.2021.1876983
- Marco-Franco JE, Pita-Barros P, Vivas-Orts D, Gonzalez-de-Julian S, Vivas-Consuelo D. COVID-19, fake news, and vaccines: should regulation be implemented? *Int J Environ Res Public Health.* (2021) 18:744. doi: 10.3390/ijerph18020744
- Pennycook G, Rand DG. Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, analytic thinking. *J Pers.* (2020) 88:185–200. doi: 10.1111/jopy.12476
- Horne CL. Internet governance in the “post-truth era”: analyzing key topics in “fake news” discussions at IGF. *Telecommun Policy.* (2021) 45:102150. doi: 10.1016/j.telpol.2021.102150
- Kwon D. Fake news: mars Edition, circa 1877. *Scientist.* (2018) 32:68–68.
- Sample C, Jensen MJ, Scott K, McAlaney J, Fitchpatrick S, Brockinton A, et al. Interdisciplinary lessons learned while researching fake news. *Front Psychol.* (2020) 11:537612. doi: 10.3389/fpsyg.2020.537612
- Silva LMS, Pinto BCT, Morado CN. Internet: impact of fake news on the biology teaching and learning process. *Rev Tecnol E Soc.* (2021) 17:203–22. doi: 10.3895/rt.s.v17n46.13629
- Ofcom. *Half of UK Adults Exposed to False Claims about Coronavirus.* (2020). Available online at: <https://www.ofcom.org.uk/about-of-what/stest/features-and-news/half-of-uk-adults-exposed-to-false-claims-about-coronavirus> (accessed Jan 20, 2022).
- Mitchell A, Oliphant JB. *Americans Immersed in Covid-19 News: Most Think Media Are Doing Fairly Well Covering It.* Pew Research Center. Available online at: <https://www.journalism.org/2020/03/18/americans-immersed-in-covid-19-news-most-think-media-are-doing-fairly-well-covering-it> (accessed January 20, 2022).
- Yang C, Zhou XY, Zafarani R. CHECKED: Chinese COVID-19 fake news dataset. *Soc Netw Anal Mining.* (2021) 11:58. doi: 10.1007/s13278-021-00766-8
- Banerjee S, Chua A. A theoretical framework to identify authentic online reviews. *Online Inf Rev.* (2014) 38:634–49. doi: 10.1108/OIR-02-2014-0047
- Wilby P. News and how to use it: what to believe in a fake news world. *Br J Rev.* (2021) 32:71–3. doi: 10.1177/0956474821998999b
- Schuetz SW, Sykes TA, Venkatesh V. Combating COVID-19 fake news on social media through fact checking: antecedents and consequences. *Eur J Inf Syst.* (2021) 30:376–88. doi: 10.2139/ssrn.3954501
- Song CG, Ning NW, Zhang YL, Wu B. A multimodal fake news detection model based on crossmodal attention residual and multichannel convolutional neural networks. *Inf Process Manag.* (2021) 58:102437. doi: 10.1016/j.ipm.2020.102437
- Zeng JF, Zhang Y, Ma X. Fake news detection for epidemic emergencies via deep correlations between text and images. *Sustain Cities Soc.* (2021) 66:102652. doi: 10.1016/j.scs.2020.102652
- Ahmad AR, Murad HR. The impact of social media on panic during the COVID-19 pandemic in Iraqi Kurdistan: online questionnaire study. *J Med Internet Res.* (2020) 22:19556. doi: 10.2196/19556
- Botosova L. Report of HLEG on fake news and online disinformation. *Media Lit Acad Res.* (2018) 1:74–7. Available online at: https://www.mlar.sk/wp-content/uploads/2019/01/MLAR_2018_2_NEWS_1_Report-of-Hleg-on-Fake-News-and-Online-Disinformation.pdf
- Subedi P, Thapa B, Pande A. Use of social media among intern doctors in regards to Covid-19. *Europasian J Med Sci.* (2020) 2:56–64. doi: 10.46405/ejms.v2i1.41
- Pulido Rodríguez C, Villarejo-Carballido B, Redondo-Sama G, Guo M, Ramis M, Flecha R. False news around COVID-19 circulated less on Sina Weibo than on Twitter. How to overcome false information? *Int Multidisciplinary J Soc Sci.* (2020) 9:107–28. doi: 10.17583/rimicis.2020.5386
- Bond GD, Lee AY. Language of lies in prison: linguistic classification of prisoners’ truthful and deceptive natural language. *Appl Cogn Psychol.* (2005) 19:313–29. doi: 10.1002/acp.1087
- Chiu MM, Oh YW. How fake news differs from personal lies. *Am Behav Sci.* (2021) 65:243–58. doi: 10.1177/0002764220910243
- Fuller CM, Biros DP, Wilson RL. Decision support for determining veracity via linguistic-based cues. *Decis Support Syst.* (2009) 46:695–703. doi: 10.1016/j.dss.2008.11.001
- Huang YK, Yang WL, Lin TM, Shih TY. Judgment criteria for the authenticity of Internet book reviews. *Libr Inf Sci Res.* (2012) 34:150–6. doi: 10.1016/j.lisr.2011.11.006
- Zhou XY, Zafarani R, Shu K, Liu H. Fake news: fundamental theories, detection strategies and challenges. In: *Proceedings of the Twelfth International Conference on Web Search and Data Mining.* Melbourne, VIC (2019). p. 836–7. doi: 10.1145/3289600.3291382
- Luo MF, Hancock JT, Markowitz DM. Credibility perceptions and detection accuracy of fake news headlines on social media: effects of truth-bias and endorsement cues. *Commun Res.* (2020) 49:171–95. doi: 10.1177/0093650220921321
- Pal A, Chua AY, Goh DHL. Does KFC sell rat? *Analysis of tweets in the wake of a rumor outbreak.* *Aslib J Inf Manag.* (2017) 69:660–73. doi: 10.1108/AJIM-01-2017-0026
- Qazvinian E, Rosengren D, Radev R, Mei Q. Rumour has it: Identifying misinformation in Microblogs. In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing.* Edinburgh (2011). p. 1589–99.
- Shin J, Jian L, Driscoll K, Bar F. The diffusion of misinformation on social media: temporal pattern, message, and source. *Comput Hum Behav.* (2018) 83:278–87. doi: 10.1016/j.chb.2018.02.008
- Zhang XC, Ghorbani AA. An overview of online fake news: characterization, detection, and discussion. *Inf Process Manag.* (2020) 57:e102025. doi: 10.1016/j.ipm.2019.03.004
- Zhou C, Li K, Lu Y. Linguistic characteristics and the dissemination of misinformation in social media: the moderating effect of information richness. *Information Proc Manag.* (2021) 58:102679. doi: 10.1016/j.ipm.2021.102679
- Zhou L, Zhang D. An ontology-supported misinformation model: toward a digital misinformation library. *IEEE Trans Syst Man Cybern Part A Syst Hum.* (2007) 37:804–13. doi: 10.1109/TSMCA.2007.902648
- Luca M, Zervas G. Fake it till you make it: reputation, competition, and yelp review fraud. *Manag Sci.* (2016) 2:3412–27. doi: 10.1287/mnsc.2015.2304
- Zhang Z, Zhang Z, Li H. Predictors of the authenticity of Internet health rumours. *Health Inf Libr J.* (2015) 32:195–205. doi: 10.1111/hir.12115
- Chua A, Banerjee S. Linguistic predictors of rumor veracity on the internet. In: *Proceedings of the International Multi Conference of Engineers and Computer Scientists.* Hong Kong: IMECS (2016).
- Kwon S, Cha M, Jung K. Rumor detection over varying time windows. *PLoS ONE.* (2017) 12:e0168344. doi: 10.1371/journal.pone.0168344
- Garcia K, Berton L. Topic detection and sentiment analysis in twitter content related to covid-19 from brazil and the usa. *Applied Soft Comput.* (2021) 101:107057. doi: 10.1016/j.asoc.2020.107057
- McCormack SA. Information manipulation theory. *Commun Monogr.* (1992) 59:1–16. doi: 10.1080/03637759209376245
- Zhao LJ, Wang XL, Wang JJ, Qiu XY, Xie WL. Rumor-propagation model with consideration of refutation mechanism in homogeneous social networks. *Discret Dynam Nat Soc.* (2014) 5:1–11. doi: 10.1155/2014/659273
- Jang Y, Park CH, Lee DG, Seo YS. Fake news detection on social media: a temporal-based approach. *CMC.* (2021) 69:3563–79. doi: 10.32604/cmc.2021.018901
- Ning P, Cheng P, Li J, Zheng M, Schwebel DC, Yang Y, et al. COVID-19–Related rumor content, transmission, and clarification strategies in China: descriptive study. *J Med Internet Res.* (2021) 23:e27339. doi: 10.2196/27339
- Morris DS, Morris JS, Francia PL. A fake news inoculation? *Fact checkers, partisan identification, and the power of misinformation.* *Polit Groups Identities.* (2020) 8:986–1005. doi: 10.1080/21565503.2020.1803935
- Atodiresei CS, Tanaselea A, Iftean A. Identifying fake news and fake users on Twitter. *Kes-2018.* (2018) 126:451–61. doi: 10.1016/j.procs.2018.07.279
- Field-Fote E. Fake news in science. *J Neurol Phys Ther.* (2019) 43:139–40. doi: 10.1097/NPT.0000000000000285
- Kehoe J. Synthesis the triumph of spin over substance: staying smart in a world of fake news and dubious data. *Harv Bus Rev.* (2018) 96:152–3. Available online at: <https://www.econbiz.de/Record/the-triumph-of-spin-over-substance-staying-smart-in-a-world-of-fake-news-and-dubious-data-kehoe-jeff/10011795270>
- Oh YJ, Ryu JY, Park HS. What’s going on in the Korean Peninsula? A study on perception and influence of South and North Korea-Related fake news. *Int J Commun.* (2020) 14:1463–51. Available online at: <https://www.researchgate.net/>

[publication/339526171_What's_Going_on_in_the_Korean_Peninsula_A_Study_on_Perception_and_Influence_of_South_and_North_Korea-Related_Fake_News](#)

46. Orso D, Federici N, Copetti R, Vetrugno L, Bove T. Infodemic and the spread of fake news in the COVID-19-era. *Eur J Emerg Med.* (2020) 27:327–32. doi: 10.1097/MEJ.0000000000000713
47. Giachanou A, Zhang GB, Rosso P. Multimodal fake news detection with textual, visual and semantic information. *Text Speech Dialogue.* (2020) 12284:30–38. doi: 10.1007/978-3-030-58323-1_3
48. Singh VK, Ghosh I, Sonagara D. Detecting fake news stories via multimodal analysis. *J Assoc Inf Sci Technol.* (2021) 72:3–17. doi: 10.1002/asi.24359
49. Singhal S, Shah RR, Chakraborty T, Kumaraguru P, Satoh S. SpotFake: a multimodal framework for fake news detection. In: *2019 IEEE Fifth International Conference on Multimedia Big Data.* (2019). p. 39–47. doi: 10.1109/BigMM.2019.00-44
50. Sperandio NE. The role of multimodal metaphors in the creation of the fake news category: a proposal for analysis. *Rev Estud Linguagem.* (2020) 28:777–99. doi: 10.17851/2237-2083.28.2.777-799
51. Xue JX, Wang YB, Tian YC, Li YF, Shi L, Wei L. Detecting fake news by exploring the consistency of multimodal data. *Inf Process Manag.* (2021) 58:102610. doi: 10.1016/j.ipm.2021.102610
52. Ball MS, Smith GW. *Analyzing Visual Data.* Newbury Park, CA: Sage (1992). doi: 10.4135/9781412983402
53. Marcella R, Baxter G, Walicka A. User engagement with political “facts” in the context of the fake news phenomenon: an exploration of information behaviour. *J Doc.* (2019) 75:1082–99. doi: 10.1108/JD-11-2018-0180
54. van Helvoort J, Hermans M. Effectiveness of educational approaches to elementary school pupils (11 or 12 years old) to combat fake news. *Media Lit Acad Res.* (2020) 3:38–47. Available online at: <https://surfsharekit.nl/public/546e69d3-5118-461b-a327-9049a427da3f>
55. Perez-Dasilva JA, Meso-Ayerdí K, Mendiguren-Galdospin T. Fake news and coronavirus: detecting key players and trends through analysis of Twitter conversations. *Prof Inf.* (2020) 29:197–207. doi: 10.3145/epi.2020.may.08
56. Taddicken M, Wolff L. ‘Fake News’ in science communication: emotions and strategies of coping with dissonance online. *Media Commun.* (2020) 8:206–17. doi: 10.17645/mac.v8i1.2495
57. Vafeiadis M, Xiao AL. Fake news: how emotions, involvement, need for cognition and rebuttal evidence (story vs. informational) influence consumer reactions toward a targeted organization. *Public Relat Rev.* (2021) 47:1–14. doi: 10.1016/j.pubrev.2021.102088
58. Igwebuike EE, Chimuanya L. Legitimizing falsehood in social media: a discourse analysis of political fake news. *Discourse Commun.* (2021) 15:42–58. doi: 10.1177/1750481320961659
59. Martel C, Pennycook G, Rand DG. Reliance on emotion promotes belief in fake news. *Cognitive Res. Principles Implicat.* (2020) 5:1–20. doi: 10.1186/s41235-020-00252-3
60. Morejon-Llamas N. Disinformation and media literacy from the institutions: the decalogues against fake news. *Revista Internacional De Relaciones Publicas.* (2020) 10:111–34. doi: 10.5783/RIRP-20-2020-07-111-134
61. China Internet Information Center. *The 47th China Statistical Report on Internet Development.* (2021). Available online at: http://www.cac.gov.cn/2021-02/03/c_1613923423079314.htm (accessed on May 16th, 2022).
62. Martin JD, Hassan F. News media credibility ratings and perceptions of online fake news exposure in five countries. *Journalism Stud.* (2020) 21:2215–33. doi: 10.1080/1461670X.2020.1827970
63. Bujang MA, Omar ED, Baharum NA. A review on sample size determination for cronbach's alpha test: a simple guide for researchers. *Malaysian J Med Sci.* (2018) 25:85–99. doi: 10.21315/mjms2018.25.6.9
64. Kanoh H. Why do people believe in fake news over the Internet? An understanding from the perspective of existence of the habit of eating and drinking. *Knowledge Based Intelligent Inform Eng Syst.* (2018) 126:1704–9. doi: 10.1016/j.procs.2018.08.107
65. Gelfert A. Coverage-reliability, epistemic dependence, and the problem of rumor-base belief. *Philosophia.* (2013) 41:763–86. doi: 10.1007/s11406-012-9408-z
66. Bago B, Rand DG, Pennycook G. Fake News, Fast and slow: deliberation reduces belief in false (but not true) news headlines. *J Exp Psychol General.* (2020) 149:1608–13. doi: 10.1037/xge0000729
67. Efron DA, Raj M. Misinformation and morality: encountering fake-news headlines makes them seem less unethical to publish and share. *Psychol Sci.* (2020) 31:75–87. doi: 10.1177/0956797619887896
68. Pennycook G, Bear A, Collins ET, Rand DG. The implied truth effect: attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. *Management Sci.* (2020) 66:4944–57. doi: 10.1287/mnsc.2019.3478
69. Ross R, Rand D, Pennycook G. Beyond “fake news”: Analytic thinking and the detection of inaccuracy and partisan bias in news headlines. *Judg Decis Mak.* (2021) 16:484–504. doi: 10.31234/osf.io/cgsx6
70. Agarwala V, Sultana HP, Malhotra S, Sarkar A. Analysis of classifiers for fake news detection. *2nd Int Conf Recent Trends Adv Comput Interactive Innovation.* (2019) 165:377–83. doi: 10.1016/j.procs.2020.01.035
71. Gracia ADV, Martinez VC. Fake news during the Covid-19 pandemic in Spain: a study through Google Trends. *Revista Latina De Comun Soc.* (2020) 78:169–82. doi: 10.4185/RLCS-2020-1473
72. Roman-San-Miguel A, Sanchez-Gey-Valenzuela N, Elias-Zambrano R. Fake news during the COVID-19 State of Alarm. Analysis from the “political” point of view in the Spanish press. *Revista Latina De Comun Soc.* (2020) 78:359–92. doi: 10.4185/RLCS-2020-1481
73. Brasoveanu AMP, Andonie R. Semantic fake news detection: a machine learning perspective. *Adv Comput Intelligence Iwann.* (2019) 11506:656–67. doi: 10.1007/978-3-030-20521-8_54
74. Zhang CW, Gupta A, Kauten C, Deokar AV, Qin X. Detecting fake news for reducing misinformation risks using analytics approaches. *Eur J Operational Res.* (2019) 279:1036–52. doi: 10.1016/j.ejor.2019.06.022
75. Zhou XY, Zafarani R. Fake news detection: an interdisciplinary research. *Companion World Wide Web Conference.* (2019) 2019, 1292–1292. doi: 10.1145/3308560.3316476
76. Hauch VI, Blandón-Gitlin MJ, Sporer SL. Are computers effective lie detectors? A meta-analysis of linguistic cues to deception. *Personality Soc Psychol Rev Official J Soc Personal Soc Psychol.* (2015) 19:307. doi: 10.1177/1088868314556539
77. Depaulo BM, Lindsay JJ, Malone BE, Muhlenbruck L, Charlton K, Cooper H. Cues to deception. *Psychol Bull.* (2003) 129:74–118. doi: 10.1037/0033-2909.129.1.74
78. Qi P, Cao J, Yang TY, Guo JB, Li JT. Exploiting multi-domain visual information for fake news detection. In: *2019 19th IEEE International Conference on Data Mining.* (2019), 518–27. doi: 10.1109/ICDM.2019.00062
79. Kurfi MY, Msughter AE, Mohamed I. Digital images on social media and proliferation of fake news on covid-19 in Kano, Nigeria. *Galactica Media J Media Stud.* (2021) 3:103–24. doi: 10.46539/gmd.v3i1.111
80. Nygren T, Guath M, Axelsson CAW, Frau-Meigs D. Combatting visual fake news with a professional fact-checking tool in education in France, Romania, Spain and Sweden. *Information.* (2021) 12:1–25. doi: 10.3390/info12050201
81. Shu K, Mahudeswaran D, Liu H. FakeNewsTracker: a tool for fake news collection, detection, and visualization. *Comput Mathemat Organiz Theory.* (2019) 25:60–71. doi: 10.1007/s10588-018-09280-3
82. Li MH, Chen ZQ, Rao LL. emotion, analytical thinking and susceptibility to misinformation During the COVID-19 outbreak. *Comput Human Behav.* (2022) 133:107295. doi: 10.1016/j.chb.2022.107295
83. Calvert C, McNeff S, Vining A, Zarate S. Fake news and the first amendment: reconciling a disconnect between theory and doctrine. *Univ. Cincinnati Law Rev.* (2018) 86:99–138. Available online at: <https://scholarship.law.uc.edu/uclr/vol86/iss1/3>
84. Sengupta E, Nagpal R, Mehrotra D, Srivastava G. ProBlock: A novel approach for fake news detection. *Cluster Comput J Netw Softw Tools App.* (2021) 24:3779–95. doi: 10.1007/s10586-021-03361-w