Chapter 2

The Emergence of COVID-19 Misinformation: Conception and Message Characteristics

Hai Liang

https://orcid.org/0000-0003-1779-9552

Ran Wei

https://orcid.org/0000-0002-2474-5280

Anfan Chen

https://orcid.org/0000-0002-7406-0415

Abstract

This chapter provides an operational definition of misinformation and relates it to the current research on the infodemic. To identify and trace the emergence of debunked yet widely circulating misinformation surrounding the COVID-19 pandemic, we used the computational method to analyze a set of big corpus from Sina Weibo, including temporal-spatial and topical patterns, sources, and message characteristics. Findings show that falsehoods about COVID-19 started with a burst in the early stage of the pandemic and followed by several peaks of rapid increase. This pattern repeated periodically. The origins of COVID-19 was the most prevalent topic, followed by information publicity and compliance measures. In terms of sources, ordinary individual and male weibo users accounted for most of the posted misinformation, whereas institutional and female users were more likely to publish debunking information. Finally, misinformation appeared to be shorter, with no URLs, and more attractive for wider circulation, making it a challenge to flatten the curve.

Keywords: misinformation, debunking, topic, temporal recurrence, Weibo

Misinformation about COVID-19 Defined

A hallmark of the COVID-19 pandemic has been the rising tide of misinformation—information that is false, factually incorrect, or misleading concerning various aspects of the virus. Although the taxonomies of false, inaccurate, or misleading information about COVID-19 have expanded to dis- and mal-information or fake news, to the best of our knowledge, there is no such a comprehensive and conclusive definition of misinformation. For example, in computational social science, misinformation refers to incorrect or misleading messages (Lazer et al., 2018). This definition has been criticized as problematic because it leads to inaccurate beliefs and greater disagreement over specific aspects of the information (e.g., truth, reliability, verifiability).

Our study focuses on COVID-19 misinformation, which differs from disinformation and mal-information. According to documents compiled by UNESCO (2018, p. 7), misinformation refers to information that is misleading but created or disseminated "without manipulative or malicious intent." As Southwell et al. (2018, p. 1) suggested, such information can be "deliberately promoted and accidentally shared" online. Disinformation, on the other hand, is orchestrated information that deliberately attempts to cause confusion or manipulate the public for economic or political gains, whereas mal-information refers to information that is true but weaponized to harm a targeted individual or institution.

Misinformation also differs from fake news in terms of the format of presentation and style. Misinformation is commonly presented as posted or blogged claims on social media, while fake news is typically presented in the form of news reports with a headline, lead, text, and accompanying photo or graphics. In the political context, Nyhan and Reifler (2010, p. 305) defined misinformation as "the presence of objectively incorrect or false information, which is not supported by clear evidence and expert opinion." Kim and Tandoc (2022) expanded the term to claims that have been debunked or corrected by expert scientific consensus that is either completely false or misleading.

Based on the above review, we conceptualize COVID-19 misinformation as unverified messages that contain false claims or factually incorrect information about COVID-19 that appear to be scientifically based. We call it “specious information” in Chinese (“似是而非的資訊”). In other words, we view misinformation concerning the COVID-19 pandemic as something like half-truths, a mixture of true, scientific information with false, incorrect claims or assertions without proven evidence. For example, BCG (Bacillus Calmette–Guérin vaccine) protects against the coronavirus; alcohol kills the virus. In other cases, assertions portray some claims as being absolutely correct. Take the misinformation on the origin of the virus as an example.

Operationally, as Kim and Tandoc (2022) argued, deciding what is scientifically true or false in a claim is complicated due to the inherent uncertainty of science. Likewise, spotting and classifying COVID-19 misinformation is also challenging and beyond the scope of our study. For instance, labeling a message as misinformation may cause concern about threats to civic freedoms, freedom of the press, and outright censorship, among others. In this chapter, instead of attempting to classify information about COVID-19 on social media, we take the combined approach of data mining analytics with manual content analyses: First, we collect a set of infodemic messages that had been identified by professional fact-checking websites or organizations; second, we integrate manual and computer-assisted content analysis to obtain a keyword list, which was frequently used to correct/refute/debunk online falsehood, and proved to be a effective instrumental in identifying infodemic messages in our dataset.

The Emergence and Topical Trends of COVID-19 Misinformation on Weibo

Since January 2020 when the COVID-19 pandemic broke out, popular social media platforms made the presence of COVID-19 misinformation prevalent in China (Chen et al., 2020; Yang et al., 2021). For instance, prominent misinformation narratives circulated widely on social media platforms. These included unsubstantiated claims about the origins of the virus; whether it was intended to be a bioweapon; and whether Bill Gates engineered the virus for profit. Although the government, news media, and social media platform operators have made intensive efforts to dispel and correct falsehoods, a large amount of misinformation still circulating on Chinese social media (Chen, 2022).

In this chapter, for practical considerations about data accessibility in all of the four cities in our study, we use large-scale datasets from Sina Weibo (hereafter as Weibo), China's most popular Twitter-like micro-messaging site, to assess the emergence and prevalence of misinformation and then examine its temporal-spatial patterns, sources, and content features. The Beijing-headquartered Weibo is a commercial site run by Sina Corporation; its micro-blogging service is accessible to the general public worldwide for self-publishers (e.g., blogging or content creators). As China's political center, cultural center, and IT center, Beijing is the world's most populous capital city. More than 21 million residents live in the capital city, where social media users from all walks of life concentrate.

Procedurally, a few misinformation-identification approaches have been developed to examine information authenticity either automatically (i.e., based on computer processing) or manually (i.e., based on human judgment, such as experts or platforms. see Zhou & Zafarani, 2020). At least four types of approaches have been proposed to detect online misinformation (Zhou & Zafarani, 2020; Grinberg et al., 2019; Allen et al., 2021):

1) **knowledge-based approach**, which detects misinformation by judging if the referred knowledge in the message (text) is consistent with facts. For example, comparing the

knowledge extracted from to-be-verified news content with known facts, with experts or crowd-sources fact-checking methods;

2) **style-based approach**, showing how the information is formatted and framed. Malicious entities prefer to write “fake” news in a special style to encourage others to read and convince them to trust, such as exaggeration and swearing;
3) **network-based approach**, mapping the misinformation by how it’s spread and how actors are connected online, such as investigating and using the information related to the dissemination of false messages;
4) **source-based approach**, evaluating the credibility of information sources or citations.

In this study, we mixed the knowledge-based approach with the expert fact-checking method in detecting the emergence of misinformation about COVID-19 on Weibo.

Findings

The Emergence of Misinformation about COVID-19 on Weibo

To estimate the emergence, prevalence, and temporal-spatial patterns of COVID-19 misinformation, we collect data from Weibo, which has 560 million monthly active users (Sina Weibo, 2021). The advanced search function of Weibo, as well as a list of representative keywords manually summarized from 2,970 COVID-19 misinformation cases from two Chinese fact-checking websites—Tencent Jiaozhen (news.qq.com/Original/jzhjym.htm) and Weibo Piyao (weibo.com/weibopiyao), are implemented in retrieving COVID-19 misinformation-related posts on Weibo published between January 1, 2020, and February 6, 2022. After removing duplicates, we obtain a total of 285,695 original posts published by 81,829 unique users.

All posts are retrieved using misinformation-related keywords, including those that debunked misleading or false posts. To filter out debunking posts, we manually summarize 34 seed keywords that are related to debunking narratives, such as “rumor”, “don’t believe”, “fake”, “not true”, “conspiracy”, and “spread the truth” (for the full list, see Appendix 2.A). Using a method called “word embeddings” (Garg et al., 2018; Kroon et al., 2021) and a pre-trained word embeddings model by Tencent AI Lab (Tencent, 2022), we expand the initial list to 399 keywords that are semantically correlated and similar to the seed words. To validate the performance of the keyword-based classification method, we randomly sample 1,000 posts and manually code if a post is debunking or misinformation. The results show that the keyword-based approach has achieved a satisfactory performance, with an agreement with human annotation of Krippendorff’s alpha coefficient at 0.786. Accordingly, our dataset include 192,542 misinformation posts (posted by 61,700 users) and 93,423 debunking posts (posted by 20,129 users).

Figure 2.1 presents the temporal trends of all misinformation posts between 2020 and 2022, from which several trends emerged. First, most misinformation messages were posted at the early stage of the pandemic in 2020 and the number of posts declined gradually over the following six months. Second, several peaks can be observed, which coincided with different waves of COVID-19 outbreaks: Phase 1 between January 29, 2020, and April 9, 2020 (initial outbreaks of Delta virus, a variant of SARS-CoV-2, in China), Phase 2 between June 20, 2020, and September 11, 2020 (the first wave in the United States), Phase 3 between October 2, 2020, and February 1, 2021 (2020 United States presidential election), and Phase 4 between November 28, 2021, and January 5, 2022 (Omicron, another variant of SARS-CoV-2, uprising with a global outbreak).

[Insert Figure 2.1 Here]

Topics of COVID-19 Misinformation Posts on Weibo

To categorize the collected posts, we adopt topic modeling to detect the hidden topics that emerge from a large collection of posts. Specifically, we conducted a Latent Dirichlet allocation (i.e. LDA) model (Blei, 2012) on all collected posts. Based on the results of the

LDA modeling process, we chose the optimal k value of 30 topics for our model. Then, the automatically generated keywords of the 30 topics were read carefully and labeled by a native Chinese speaker and communication expert. Finally, we manually recoded the 30 extracted topics into 13 topical frames (see Appendix 2.B).

As presented in Figure 2.2, among all original posts related to COVID-19 misinformation, the most prevalent topics are concerned with the origin of COVID-19 (23.63%), publicity such as government announcements or updates on the outbreaks (16.68%), and prevention measures such as compliance (9.15%), followed by vaccination (8.67%), quarantine policies and measures (7.23%), international aspects of the pandemic (7.17%), international aid (7.14%), daily confirmed cases (5.74%), virus testing (5.04%), updates on pandemic variants (4.94%), virus spreading (3.58%), and the social impacts of the pandemic (2.56%). The rest, coded as “other”, account for 4.46%. The topical features of Weibo seem to be more diverse and wide-ranging compared to other studies (e.g., Zeng & Chan, 2021) which reported that the most prevailing topics in Chinese pieces of misinformation were about travel restrictions and travelers who spread the virus.

[Insert Figure 2.2 Here]

The topics of misinformation over the two years are matched with the timeline of the outburst of posts on Weibo. The evolving trends of ups and downs in topics are shown in Figure 2.3. The origin of COVID-19 is a recurring topic and the proportion generally increased over the entire period of two years. Additionally, among the most posted topics, misinformation about preventive measures is more consistent while misinterpretation or faulty posting on official publicity (e.g., updates, daily briefings, and policy announcements) decreases over time.

[Insert Figure 2.3 Here]

Sources of Misinformation on Weibo

Because all Weibo posts we collect are original after excluding reposts, we consider the users who posted the misinformation messages (e.g., bloggers, content creators, or content curators) as a proxy of the source of misinformation according to the categories of account they hold: (1) government accounts, including a variety of departments and branches of the Chinese government at all levels from national to village-level; (2) media accounts, including all sorts of media organizations, professional journalists, editors, and TV hosts/hostess, and web-based freelancers such as bloggers, publishers or self-appointed hosts (3) verified individual users (aka influential content creators), those who actively create information for others to consume and whose identities have been verified by Weibo; (4) corporation accounts, including accounts affiliated to commercial entities; and (5) unverified individual users (e.g., the rest, which is the majority of Weibo users).

As Table 2.1 shows, most of the misinformation sources are ordinary individual account holders who were either anonymous or whose identities were not verified. They account for nearly two-thirds of the sources of posted COVID-19 misinformation on Weibo (60.58%), followed by verified individual account holders (23.81%). The rest are sourced by holders or creators of government accounts (7.54%), corporation accounts (5.58%), and media accounts (2.49%).

Although the two types of individual account holders proportionally account for 84.39% of all misinformation sources, they only generate 52.92% of misinformation messages. Surprisingly, government and media account holders have posted 40.96% of misinformation messages. In a closer look at these sources, we note that the majority of those government accounts are run by local governments at the grassroots (e.g., township party committee or party secretary of a village). The high percentage of misinformation posted from those accounts suggests that the local government account holders may lack the know-how in filtering pandemic information and sorting out infodemic messages. Similarly, a great number of media accounts are created and operated by so-called free-lance content creators, not the official media. They are sources of misinformation because they may be either misled by falsehoods or motivated by business incentives to unintentionally or intentionally promote misinformation messages.

[Insert Table 2.1 Here]

In addition, self-reported gender can be directly retrieved from Weibo profile, with 1 = male (n = 195,922, 68.58%), and 0 = female (n = 89,773, 31.42%). Based on the profile information, we also obtain the number of followers (M = 10,297, Mdn = 8,966, SD = 2,305), the number of followees (those who are followed by the engaged user) (M = 667, Mdn = 297, SD = 1,382), and the number of historical posts (M = 10,518, Mdn = 2,924, SD = 115,487) for each engaged user in our dataset. The results of these additional analyses suggest that male users were more likely to post misinformation posts than were female users (male = 33.30%, female = 27.50%), while female users were more likely to be associated with debunking posts than were male users (female = 72.50%, male = 66.70%), with a significant difference, χ^2 (1, N = 285,695) = 962.089, p < .001. Verified accounts were less likely to post misinformation than debunking messages (66.90% vs. 78.40%), while unverified accounts were more likely to post misinformation than debunking messages (33.10% vs. 21.60%), with a significant difference, χ^2 (1, N = 285,695) = 3971.300, p < .001. However, we did not find any significant differences in terms of the number of followers, followees, or historical posts of users.

Characteristics of Misleading COVID-19 Messages

Some textual features are extracted from the Weibo posts and have been examined in previous social media studies. Factors have included the length of the texts and whether the posts include pictures, hashtags, or URLs. We used existing tools to extract other latent features. For example, the emotion score of each post was calculated based on the methods by Zhang et al. (2017) and Liu (2010). To be specific, positive words were assigned the value equal to +1, while negative words with -1. The emotion scores of positive and negative words were adjusted in the case of transition words (e.g., but). The emotion score was also adjusted by weighting for each polarity score according to its adverbs of degree. To complete the calculation, the valence of each polarity word in a post was added to obtain an emotion score for each post.

A larger value indicates more positive emotions. The detection of uncivil posts is based on the word embedding method and a pre-trained word embedding model by Tencent AI Lab (Tencent, 2022). Specifically, we developed a list of 2,643 uncivil words from a list of 40 seed uncivil words obtained from Song et al. (2021).

Since the abovementioned variables were proposed to examine social media texts in general, they might not be specific to misinformation. We also compared the features of misinformation messages with debunking messages to see what features might be related to misinformation messages.

As Table 2.2 shows, in terms of message attributes, 36.60% of the messages included a picture, 64.50% carried hashtags, and only 0.64% included hyperlinks, suggesting most misinformation messages did not have a clear source. Nevertheless, they were the common features of social media posts related to COVID-19, at least on Weibo. Interestingly, one-third (33.05%) of the misinformation messages used uncivil words such as vulgarity and swearing. In addition, they were more emotional. Compared to debunking messages, misinformation messages were more likely to include positive emotions (12.38 vs. 10.35, $t = 21.0$, $p < .001$) and uncivil words (33.05% vs. 20.68%, χ^2 (1, N = 285,695) = 1,706.3, $p < .001$), but more likely to be shorter (199.80 vs. 229.60, $t = 21.9$, $p < .001$). These unique features of misinformation indicate that misinformation can be more eye-catching and thus spread more widely than debunking and other messages as indicated in previous studies (e.g., Vosoughi et al., 2018).

[Insert Table 2.2 Here]

Summary of Key Findings

In summary, using big data from Weibo and computational methods and with a focus on temporal and topical distribution, source, and message characteristics, a number of patterns of COVID-19 misinformation were uncovered in the present chapter:

- Most misinformation messages were posted at the early stage of the pandemic in 2020 and the number of posts declined gradually over the following six months.
- Several peaks in posting were observed, which coincided with different waves of COVID-19 outbreaks. That is, whenever there was a new outbreak, posting concering COVID-19 that was halftruths or total falsehood would increase in largest numbers.
- Misinformation concerning the CVODI-19 appeared to be shorter, with no URLs, and more attractive in visual appeal crafted for wider reach on social media.
- The origin of COVID-19 was the most prevalent topic, followed by false messages about updagtes on the pandemic and compliance measures.
- In terms of sources, ordinary individual and male users accounted for most of the posted misinformation, whereas institutional and female users were more likely to post debunking information.

Implications

In this chapter, we have applied computational methods to analyze a set of massive social media corpus collected from Weibo in Beijing for two years to sketch out the emergence of misinformation surrounding COVID-19 and assess its temporal and topical features in terms of trending patterns, prevailing topics, sources, and content features of the misinformation. The following is a summary of our key findings:

First, temporal patterns of misinformation posts on Weibo show that falsehoods about COVID-19 started with a burst and then increased rapidly with several peaks, and that this pattern repeated periodically. Specifically, our data found that 5 out of 13 topics on COVID-19 misinformation resurged multiple times. Temporal recurrence of such misinformation messages and the recurring spread of misinformation may hinder falsehood spreaders from strategically keeping false rumors alive in hopes of influencing followers or followees. This finding lends critical implications for misinformation correction practices by authorities and medical agencies: countering the infodemic could not be accomplished with a single campaign. Persistent efforts are needed to counter or correct misinformation during and ever after a pandemic.

Second, the most posted topics of misinformation messages are about the original COVID-19, information publicity, and compliance. These findings are comparable to the previous study by Chen et al. (2020), indicating that the origination of COVID-19 was the locus of public attention in the infodemic. Although the top three topics were also the most prevalent in debunking messages, we observed significant differences in topics between posting misinformation and debunking messages. Even though 23.63% of misinformation posts were about COVID-19 origin, only 11.94% of debunking messages were related to the topic. On the contrary, 9.36% of debunking messages were about the virus spreading, and only 3.58% were misinformation related to it. This is especially critical in terms of the misinformation countering effectiveness, since this type of divergence and mismatch may temper the efficiency of misinformation correction efforts enacted by multiple users.

Third, individual accounts, no matter verified or not, tend to post more misinformation than debunking messages, whereas government and media accounts were more likely to post debunking than misinformation messages. Nevertheless, government and media accounts also posted more than 40% of misinformation messages in total. As the

Chinese public relies heavily on social media channels for updated information and social interactions to cope with the COVID-19 pandemic (Wei et al., 2021, see Chapter 4 for details), frequent encounters of misinformation on social media may significantly shape people's beliefs and behavior about the COVID-19 pandemic. This is consistent with previous studies in terms of information sources about misinformation diffusion and consumption, indicating that verified users, especially institutions or organizations may play more prominent roles in the diffusion of misinformation consumed by the Chinese public (Chen et al., 2020; Vraga & Bode, 2017). Moreover, the study also found a link between user gender and misinformation engagement—debunking posts were more likely to come from female users, while misinformation messages had a higher possibility from male users (Chen et al., 2020).

Finally, we found that a great number of the misinformation messages also included pictures and hashtags, which might increase the spreading of these messages. Nevertheless, the percentages were relatively low compared to debunking messages. In addition, misinformation messages were shorter and more likely to contain emotions and uncivil words than debunking messages.

To conclude, although we were unable to analyze COVID-19 misinformation data from all of our studied cities due to lack of access to big data from Facebook, Twitter, or Line, the findings based on Weibo data have several implications for policymakers in combating misinformation during a public crisis: 1) misinformation on popular social media channels appear to be more attractive (and thus more widely circulated) than other messages (e.g., debunking messages); 2) it is very difficult to reduce the proliferation of misinformation from the sources because, as the Weibo posts show, it comes from nowhere (no URLs to trace, and government and media sources account for a significant proportion of it). These two insights suggest that coping with misinformation on social media platforms is a big challenge to the government and public health authorities.

Building on these findings next chapter will examine the dynamic process of the diffusion of misinformation about COVID-19 to uncover the patterns of the spread on social media platforms with a computational approach. From the user perspective, we will also investigate the vulnerability of receivers with different demographic characteristics who encounter misinformation in Chapter 4.

References

Allen, J., Arechar, A., Pennycook, G., & Rand, D. (2021). Scaling up fact-checking using the wisdom of crowds. *Science Advances, 7*(36), 1-10. https://doi.org/10.1126/sciadv.abf4393

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM, 55*(4), 77-84. https://doi.org/10.1145/2133806.2133826

Chen, K., Chen, A., Zhang, J., Meng, J., & Shen, C. (2020). Conspiracy and debunking narratives about COVID-19 origins on Chinese social media: How it started and who is to blame. *Harvard Kennedy School (HKS) Misinformation Review*,*1*(8),1-30. https://doi.org/10.37016/mr-2020-50

Chen, X. (2022, May 05). At the moment of the new crown pneumonia epidemic, do not disturb. *Science and Technology Daily.* http://digitalpaper.stdaily.com/http_www.kjrb.com/kjrb/html/2022-05/05/content_534647.htm?div=-1

Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences, 115*(16), 3635-3644. https://doi.org/10.1073/pnas.1720347115

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 US presidential election. *Science, 363*(6425), 374-378. https://doi.org/10.1126/science.aau2706

Kim, H., & Tandoc Jr., E. (2022). Consequences of online misinformation on COVID-19: Two potential pathways and disparity by eHealth literacy. *Frontiers in Psychology, 13*, 1-10. https://doi.org/10.3389/fpsyg.2022.783909

Kroon, A. C., Trilling, D., & Raats, T. (2021). Guilty by association: Using word embeddings to measure ethnic stereotypes in news coverage. *Journalism & Mass Communication Quarterly, 98*(2), 451-477. https://doi.org/10.1177/1077699020932304

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., ... & Zittrain, J. L. (2018). The science of fake news. *Science, 359*(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

Liu, B. (2010). Sentiment analysis and subjectivity. *Handbook of Natural Language Processing, 2*, 627-666. https://doi.org/10.1201/9781420085938-c26

Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior, 32*(2), 303-330. https://doi.org/10.1007/s11109-010-9112-2

Sina Weibo. (2021). *Weibo service user agreement*. https://weibo.com/signup/v5/protocol

Song, Y., Kwon, K. H., Xu, J., Huang, X., & Li, S. (2021). Curbing profanity online: A network-based diffusion analysis of profane speech on Chinese social media. *New Media & Society, 23*(5), 982-1003. https://doi.org/10.1177/1461444820905068

Southwell, B. G., Thorson, E. A., and Sheble, L. (2018). Introduction: Misinformation among Mass Audiences as a Focus for Inquiry. In B. G. Southwell, E. A. Thorson & L. Sheble (Eds.), *Misinformation and Mass Audiences* (pp. 1-12). University of Texas Press. https://doi.org/10.7560/314555-002

Tencent. (2022). *Tencent AI Lab Embedding Corpus for Chinese Words and Phrases: A corpus on continuous distributed representations of Chinese words and phrases*. https://ai.tencent.com/ailab/nlp/en/embedding.html

UNESCO. (2018). *Journalism, 'fake news' & disinformation: Handbook for journalism education and training*. Retrieved May 28, 2020, from https://unesdoc.unesco.org/ark:/48223/pf0000265552/PDF/265552eng.pdf.multi

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146-1151. https://doi.org/10.1126/science.aap955

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

Wei, L., Yao, E., and Zhang, H. (2021). Authoritarian responsiveness and political attitudes during COVID-19: Evidence from Weibo and a survey experiment. *Chinese Sociological Review*, 1-37. https://doi.org/10.1080/21620555.2021.1967737

Yang, K. C., Pierri, F., Hui, P. M., Axelrod, D., Torres-Lugo, C., Bryden, J., & Menczer, F. (2021). The Covid-19 infodemic: Twitter versus Facebook. *Big Data & Society*, *8*(1), 1-16. https://doi.org/10.1177/20539517211013861

Zeng, J. & Chan, C.-H. (2021). A cross-national diagnosis of infodemics: Comparing the topical and temporal features of misinformation around COVID-19 in China, India, the US, Germany and France. *Online Information Review, 45*(4), 709-728. https://doi.org/10.1108/OIR-09-2020-0417

Zhang, L., Xu, L., & Zhang, W. (2017). Social media as amplification station: Factors that influence the speed of online public response to health emergencies. *Asian Journal of Communication*, *27*(3), 322-338. https://doi.org/10.1080/01292986.2017.1290124

Zhou, X., & Zafarani, R. (2020). A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys (CSUR), 53*(5), 1-40. https://doi.org/10.1145/3395046

Account Types	*N* (%) of accounts	*N* (%) of posts
Ordinary individual account	37,376 (60.58%)	60,045 (31.19%)
Verified individual account	14,689 (23.81%)	41,832 (21.73%)
Government account	4,653 (7.54%)	41,230 (21.42%)
Corporation account	3,444 (5.58%)	11,795 (6.13%)
Media account	1,538 (2.49%)	37,607 (19.54%)
Total	61,700 (100%)	192,509 (100%)

Table 2.1. Posted misinformation messages about COVID-19 by types of sources (n = 61,700)

	Misinformation	Debunking
	N (%)	*N* (%)
Pictures	70,455 (36.6%)	45,850 (49.1%)
Hashtag	124,196 (64.5%)	64,330 (68.8%)
URL	1,233 (0.64%)	255 (0.27%)
Uncivil words	63,632 (33.05%)	19,327 (20.68%)
	M (*Mdn*, *SD*)	*M* (*Mdn*, *SD*)
Emotion intensity	12.38 (7.86, 25.80)	10.35 (6.88, 23.40)
Text Length	199.80 (146.0, 295.80)	229.60 (145.00, 360.90)

Table 2.2. Content features of misinformation and debunking messages

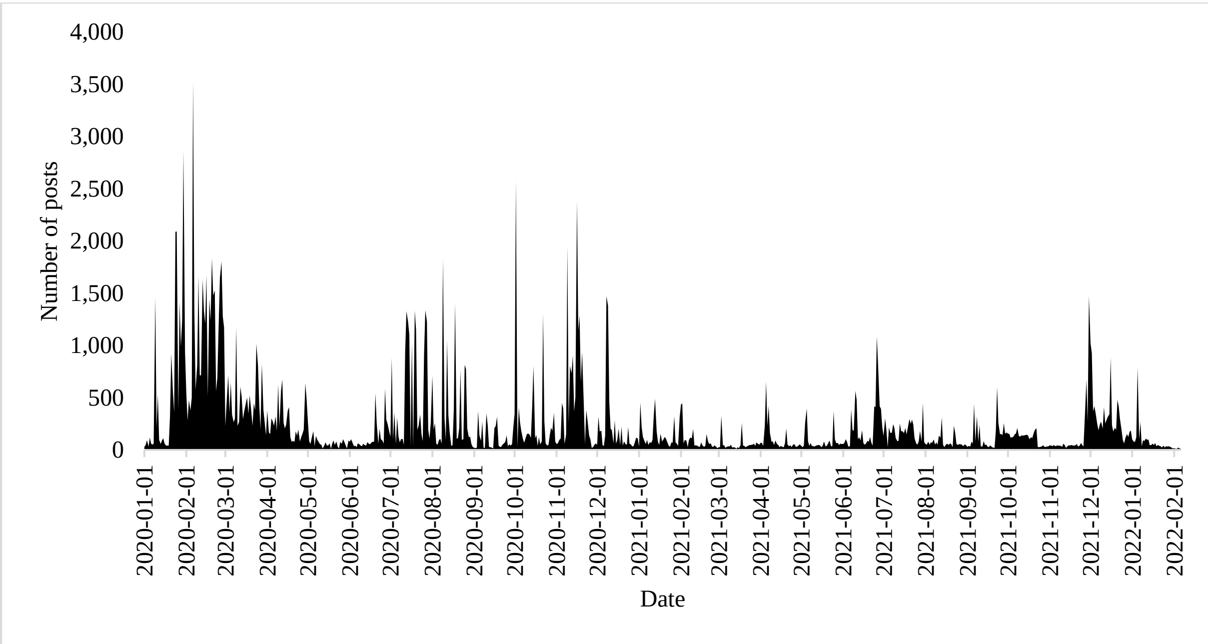

Figure 2.1. Time series of emergence of edentified Weibo posts pertaining to COVID-19 misinformation

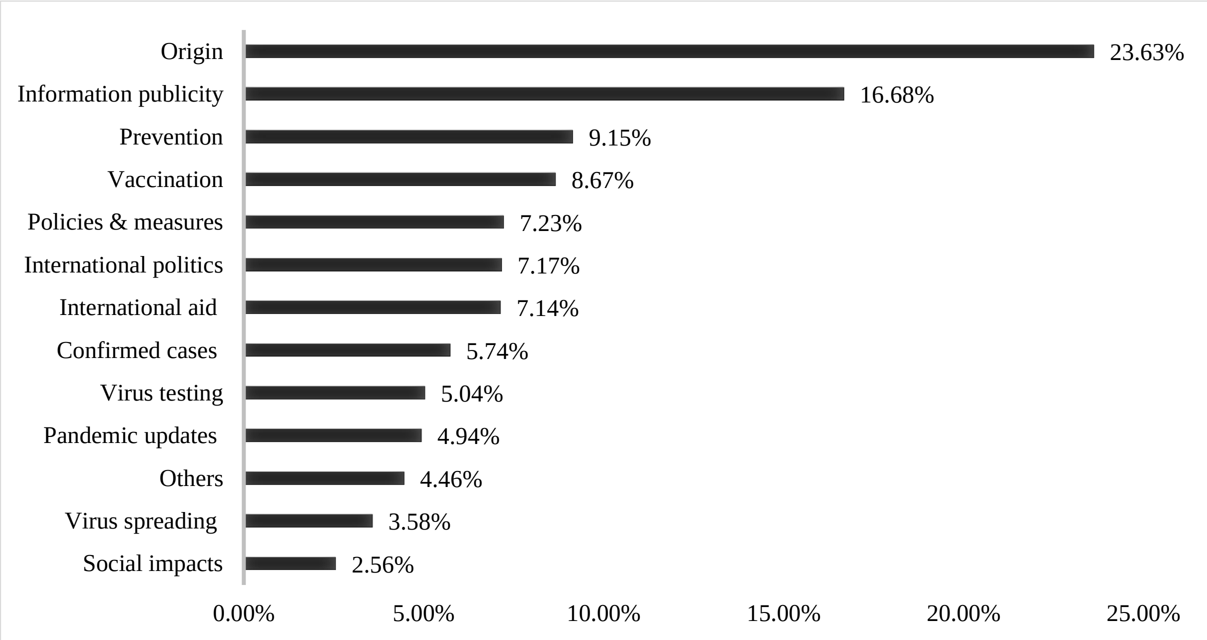

Figure 2.2. Distribution of posts of COVID-19 misinformation message by topical categories (*N* = 285,695)

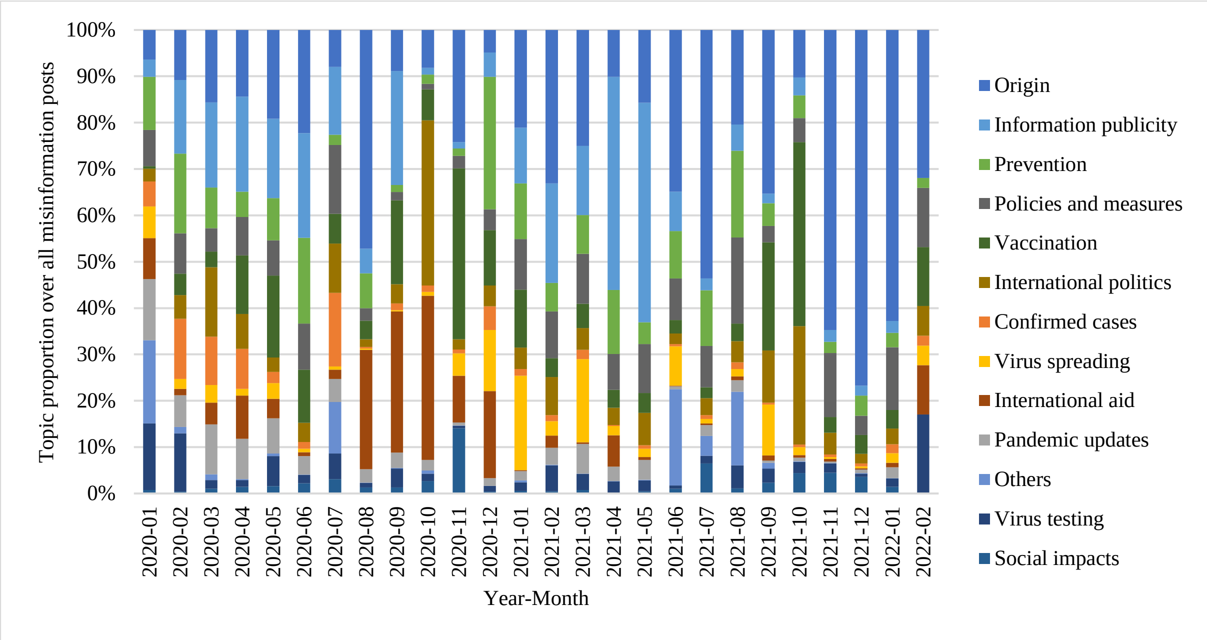

Figure 2.3. Time distribution of posts of COVID-19 misinformation message by categories