



From News Source Sharers to Post Viewers: How Topic Diversity and Conspiracy Theories Shape Engagement With Misinformation During a Health Crisis

YUWEI CHUAI, University of Luxembourg, Luxembourg

JICHANG ZHAO, Beihang University, China

GABRIELE LENZINI, University of Luxembourg, Luxembourg

Online engagement with misinformation threatens societal well-being, particularly during health crises when susceptibility to misinformation is heightened in a multi-topic context. Here, we focus on the COVID-19 pandemic and address a critical gap in understanding engagement with multi-topic misinformation on social media at two user levels: news source sharers (who post news items) and post viewers (who engage with news posts). To this end, we analyze 7273 fact-checked source news items and their associated posts on X through the lens of topic diversity and conspiracy theories. We find that false news, especially those containing conspiracy theories, exhibits higher topic diversity than true news. At news source sharer level, false news has a longer lifetime and receives more posts on X than true news, with conspiracy theories further extending its longevity. However, topic diversity does not significantly influence news source sharers' engagement. At post viewer level, contrary to news source sharer level, posts characterized by heightened topic diversity receive more reposts, likes, and replies. Notably, post viewers tend to engage more with misinformation containing conspiracy narratives: false news posts that contain conspiracy theories, on average, receive 40.8% more reposts, 45.2% more likes, and 44.1% more replies compared to those without conspiracy theories. Our findings suggest that news source sharers and post viewers exhibit distinct engagement patterns on X, offering valuable insights into refining misinformation interventions at these two user levels.

CCS Concepts: • **Human-centered computing** → **Collaborative and social computing; Empirical studies in HCI**; • **Information systems** → **Social networks**.

Additional Key Words and Phrases: Misinformation, topic diversity, conspiracy theories, COVID-19, online engagement

ACM Reference Format:

Yuwei Chuai, Jichang Zhao, and Gabriele Lenzini. 2025. From News Source Sharers to Post Viewers: How Topic Diversity and Conspiracy Theories Shape Engagement With Misinformation During a Health Crisis. *Proc. ACM Hum.-Comput. Interact.* 9, 7, Article CSCW298 (November 2025), 53 pages. <https://doi.org/10.1145/3757479>

1 Introduction

The spread of misinformation – i. e., general false or misleading information created either intentionally or unintentionally – on social media platforms has emerged as one of the most significant global risks in recent years, posing far-reaching consequences for societies and individuals [10, 21]. Health misinformation, compared to misinformation in other domains, carries unique dangers due to its potential to cause harm to human bodies and lives [6]. This issue is particularly concerning during public health crises, such as the COVID-19 pandemic, when heightened emotions,

Authors' Contact Information: **Yuwei Chuai**, University of Luxembourg, Luxembourg, yuwei.chuai@uni.lu; **Jichang Zhao**, Beihang University, China, jichang@buaa.edu.cn; **Gabriele Lenzini**, University of Luxembourg, Luxembourg, gabriele.lenzini@uni.lu.



This work is licensed under a Creative Commons Attribution 4.0 International License.

© 2025 Copyright held by the owner/author(s).

ACM 2573-0142/2025/11-ARTCSCW298

<https://doi.org/10.1145/3757479>

uncertainty, and increased information consumption create fertile ground for the dissemination of false or misleading information [7, 8, 11, 28, 71]. Additionally, as a distinct type of misinformation, conspiracy theories are deliberately fabricated and attempt to explain the causes of significant public events as secret plots by powerful and malicious groups [17, 31, 38, 72]. Throughout the COVID-19 pandemic, misinformation and conspiracy theories regarding the virus, its origin, and potential treatments or prevention strategies have spread widely on social media, undermining pandemic control efforts [24, 70, 75]. For instance, widespread anti-vaccine conspiracy theories on social media have strengthened beliefs in the ineffectiveness and side effects of vaccinations, contributing to individuals' vaccination hesitancy and refusal, thereby increasing the risk of COVID-19 infection [22, 54, 70].

Notably, during health crises, various aspects of society become interconnected. The discussions traverse multiple domains and demonstrate strong interconnectedness, contributing to a rich and diverse informational landscape [45, 80]. For example, a comprehensive analysis of 13.9 million COVID-19 posts on X (formerly Twitter) reveals discussions spanning more than 20 topics across society, encompassing areas such as politics, economy, and racism [4]. Yet, an understanding of how users engage with misinformation within the multi-topic context during health crises is still in its infancy. Vosoughi et al. [78] examined a large-scale of true and false news posts distributed on X from 2006 to 2017, finding that false news posts are more novel and attract more replies expressing surprise compared to true news posts. The novelty in their study was measured by the topic similarity between the true (false) news post a user reposted and the previous posts to which the user had been exposed before the repost. However, other research shows that social media users often have specific topic preferences and tend to consume similar content, especially within groups of like-minded individuals, i. e., echo chambers [13]. Additionally, social media recommendation algorithms frequently feed users content based on their view histories and increase echo chambers, further reinforcing users' preferences [14, 23]. Reconciling these seemingly conflicting findings, one possible explanation is that false news may strategically connect users' interests with other topics to pursue (hidden) agendas. Particularly, in the context of health crises marked by high uncertainty, diverse topics are more interconnected within discussions around crises [4]. Therefore, an analysis of topic interconnectedness in true and false news items during health crises is warranted.

Previous research has shown that the diversity of topics within information content (i. e., multi-topic information) plays an important role in shaping perceptions and assessments of credibility through the mechanism of associative inference [40]. Associative inference is an adaptive process that enables individuals to link related information and form novel connections, even in the absence of direct experience, thereby increasing individuals' susceptibility to online misinformation [40]. Moreover, the associative inference around online misinformation can evolve within a reinforcing cycle. Individuals collaboratively construct a misleading version of reality and amplify alleged evidence. This facilitates subsequent actions that, in turn, further reinforce the manufactured reality [56]. Therefore, the interconnected and associative multi-topic discussions during health crises underscore the potential role of topic diversity in driving engagement with misinformation on social media. Based on this rationale, we address the following research questions to explore topic diversity in misinformation (RQ1.1) and examine whether it shapes user engagement (RQ1.2):

- *RQ1.1: Does misinformation exhibit higher topic diversity compared to true information?*
- *RQ1.2: Is topic diversity positively associated with user engagement, and is this association stronger for misinformation than for true information?*

Moreover, the integration of conspiracy theories within misinformation narratives serves as an effective strategy to deliberately interconnect various topics [31, 45]. This deliberate approach contributes to an increase in individuals' beliefs in misinformation by fostering illusory pattern

perception and promoting associative inference [40, 76]. In the early months of the COVID-19 pandemic, social media platforms were inundated with various conspiracy theories related to the origins, spread, and treatment of the virus [24]. The propagation of misinformation with coordinated conspiracies has promoted distrust in public health authorities and skepticism about the safety and effectiveness of COVID-19 vaccines [65, 70]. While the proliferation of conspiracy theories on social media has raised significant concerns, little is known about how the incorporation of conspiracy theories in misleading narratives differentiates individuals' engagement on social media compared to other types of misinformation [37]. Given this, we further investigate the following research questions to analyze whether misinformation containing conspiracy theories exhibits higher topic diversity (RQ2.1) and receives more engagement (RQ2.2) compared to misinformation without conspiracy theories:

- *RQ2.1: Does misinformation containing conspiracy theories exhibit higher topic diversity compared to misinformation without conspiracy theories?*
- *RQ2.2: Does misinformation containing conspiracy theories receive more engagement compared to misinformation without conspiracy theories?*

To address our research questions, we perform a comprehensive analysis of online engagement with multi-topic misinformation during the early stages of the COVID-19 pandemic, focusing on topic diversity and conspiracy theories. It is worth noting that we examine engagement at two levels: *news source sharer level* and *post viewer level*. News source sharers create posts on X that link to source news items, while post viewers encounter these news posts within the platform and then engage by reposting, liking, or replying. The two user groups may have different motivations for engagement [2, 31, 75]. Additionally, while previous studies have extensively explored post viewer engagement on social media [11, 57, 71, 78], research on news source sharer engagement remains limited. Understanding this under-investigated area is crucial to complement our knowledge of news-sharing behavior.

Methodology: We conduct our analysis based on a COVID-19 dataset consisting of 7273 fact-checked source news claims and their corresponding posts on X. These source news claims span six languages and cover the initial period of COVID-19 from January to September 2020. At news source sharer level, the engagement metrics include the number of posts linking to a specific news item and the lifetime of the news item from the creation of its first post to its latest post. At post viewer level, the engagement metrics include repost count, like count, and reply count. To gain a comprehensive understanding of the topic spectrum surrounding COVID-19 discussions, we additionally use over 15 million daily sampled posts at the early stage of COVID-19 to train a Word2Vec model and collect topic keywords. Subsequently, we use the topic keywords to calculate topic diversity and identify conspiracy theories. Finally, we employ multiple regression models to investigate our research questions.

Contributions: Our work reveals that elevated topic diversity has been a characteristic of COVID-19 misinformation, particularly for those accompanied by conspiracy narratives. This characteristic becomes more significant within false news posts on X compared to their source news claims. At post viewer level, heightened topic diversity and the integration of conspiracy theories significantly correlate with increased engagement (i. e., reposts, likes, and replies) with misinformation on X. For instance, false news posts that contain conspiracy theories, on average, receive 40.8% more reposts, 45.2% more likes, and 44.1% more replies compared to false news posts without conspiracy theories. However, at news source sharer level, topic diversity and conspiracy theories have no significant associations with the number of posts that source news items receive. This difference suggests that news source sharers and post viewers have distinct engagement patterns. Our findings offer valuable insights into understanding engagement with

multi-topic misinformation from news source sharers to post viewers on social media during health crises. Additionally, by highlighting the different engagement patterns at news source sharer and post viewer levels in terms of topic diversity and conspiracy theories, this research informs the development of targeted interventions and strategies at both user levels to mitigate engagement with false information and foster a more resilient society.

2 Background and Related Work

2.1 Misinformation on Social Media

Social media platforms have transformed information exchange, enabling public users to create and engage with online content globally and instantaneously. However, the absence of quality control mechanisms inherent in traditional media has led to a significant rise in misinformation on social media. Particularly, the spread of misinformation on social media has posed far-reaching consequences across many critical societal domains ranging from politics to public health, thus becoming a concerning global issue in recent years [21, 22, 30, 54]. Consequently, social media platforms are under increasing pressure to implement effective measures to reduce the spread of misinformation. Understanding how users engage with misinformation and contribute to its viral spread on social media can help social media platforms design targeted interventions. Prior research has shown that online misinformation spreads farther, faster, deeper, and more broadly compared to true information, as evidenced by the examination of online posts fact-checked by third-party organizations [78]. Nevertheless, another study, based on community fact-checked posts on X, indicates that misinformation receives fewer reposts compared to true information [19]. This contradiction highlights that not every piece of misinformation spreads faster than true information. Instead, there are likely crucial factors behind the veracity of information that make misinformation unique and contribute to its online diffusion.

Several works have focused on understanding what characterizes misinformation and contributes to its viral spread. For instance, misinformation is more likely to incorporate negative emotions, such as anger, which has been identified as an important driver for information diffusion [8, 11, 33, 62]. Additionally, there is a partisan asymmetry that misinformation tends to be right-leaning and promote conservative positions [12]. Social media users are more likely to believe and share information that is congruent to their political ideologies [29, 36, 50, 59, 63]. Therefore, right-leaning users are more likely to share misinformation on social media, compared to left-leaning users [49]. More importantly, uncertain situations, such as public emergencies or crises, create fertile ground for online misinformation to thrive. During such times, people are eager to know more information about the situation and are thus more likely to believe and share related misinformation, which can undermine trust in control policies and cause detrimental harm to societies and individuals.

2.2 Misinformation During Public Events/Crises

During public events or crises, characterized by heightened emotions, uncertainty, and a surge in information consumption, falsehoods that cater to specific demands and offer related information to contextualize the situation become rampant and easily spread [11]. The preservation of democracy is crucial during political elections. However, democracy can be undermined by the widespread dissemination of dedicated online misinformation aimed at pursuing specific political agendas. It has been found that people are exposed to misinformation more often during elections than usual [1, 12]. Election-related misinformation is often polarized, which can increase people's susceptibility and affect their voting behaviors [49, 60, 61]. Additionally, climate change poses challenges at multiple levels, with any mitigation efforts carrying significant economic and political implications [41]. Consequently, the debate surrounding global climate change has persisted for years and has

become increasingly polarized on social media [25–27]. This polarized environment fosters climate change denial and misinformation, further complicating efforts to address and control climate change [25, 41].

The spread of misinformation during health crises is particularly severe, as it can directly harm human health and even cost lives [6]. Throughout the COVID-19 pandemic, misinformation on a wide variety of COVID-19 related topics has circulated widely on social media platforms. For example, the widespread misinformation on the COVID-19 vaccine reduces users' beliefs in its effectiveness, impeding the fight against the spread of COVID-19 [22, 43, 54]. A global survey has revealed that a substantial part of people consider COVID-19 misinformation as highly reliable [64]. Understanding how misinformation spreads and attracts users during such public events and crises is crucial for developing effective intervention strategies for emergency management.

2.3 Engagement With COVID-19 Misinformation on Social Media

The engagement with misinformation during health crises, particularly COVID-19, poses significant threats to crisis control and public safety. Several studies have scrutinized the factors that can influence engagement with COVID-19 misinformation on social media. Recognizing emotions as a significant driver of online information diffusion [11, 57, 58], research in this body has predominantly focused on the role of emotions during COVID-19 [5, 8, 28, 71]. For example, Solovev and Pröllochs [71] conducted a large-scale computational analysis on the role of moral emotions in the spread of COVID-19 misinformation. They found that false posts with higher other-condemning emotions (e.g., anger) are more viral (i.e., more reposts) than true posts. Based on experimental design, Freiling et al. [28] investigated the role of anxiety in misinformation belief and sharing during COVID-19, revealing that anxiety is a driving factor in believing and sharing both true and false claims. Wang and Rzeszotarski [79] put the focus on the images and found that online posts containing COVID-19 misinformation images receive similar interactions with posts containing random images. However, COVID-19 misinformation images are shared for longer periods than non-misinformation ones. Additionally, COVID-19 misinformation, mimicking the format and language features of news and scientific reports, is likely to receive more likes [48]. For the concern that social bots may share a large amount of misinformation during COVID-19, Teng et al. [75] analyzed different types of users in the engagement with misinformation. They found that the role of social bots in misinformation sharing is limited, and individuals with human-like behavior play a more prominent role.

Knowing the potential role of topics in engagement with misinformation, Ngai et al. [48] focused on the specific topics in the engagement with COVID-19 vaccine misinformation. They found that safety concerns are the most significant content theme but are negatively associated with likes and shares. However, during COVID-19, various aspects of society became interconnected. COVID-19 misinformation covered a broad spectrum of topics, including public health, politics, medicine, vaccines, and so forth [4, 80]. The multi-topic misinformation about COVID-19 carries high uncertainty and easily causes confusion to individuals, thus posing threats to people and society [80]. It has been revealed that individuals can adaptively link multiple related topics and make novel connections through the mechanism of associative inference, which increases the susceptibility to misinformation [40].

2.4 Multi-Topic Context and Conspiracy Theories During COVID-19

The discussion on COVID-19 during the pandemic not only focused on the direct health impacts but also spilled over into various social, economic, and political domains. Within the social domain, at the beginning of the pandemic, rumors and conspiracy theories about the origin of the virus in China were rampant on social media, exploiting anti-Asian racism and hate speech on social media

[35]. Using COVID-19 as a weapon, malicious misinformation fostered online hate to spread quickly beyond the control of social media platforms [77]. The rising anti-Asian racism during the COVID-19 pandemic knocked Asian businesses [35, 44, 74]. In the economic domain, the pandemic hit the global economy, particularly impacting sectors such as travel and leisure [81]. Additionally, the scientific uncertainty surrounding the COVID-19 pandemic fueled the weaponization of COVID-19 for politicization [39]. Some individuals and groups have sought to exploit the pandemic to spread extremist ideologies and radicalize others [15]. The high degree of politicization within COVID-19 information contributed to polarization in attitudes towards the pandemic and even to political ideologies [32]. For instance, the debate over vaccination became increasingly polarized during the pandemic, expanding far beyond the medical discourse and entering the domain of organizational politics [16]. Taken together, the COVID-19 pandemic has been incorporated into many other societal domains to pursue specific interests, yielding significant spillover effects on the stability of societies. In this paper, we use topic diversity to denote the extent to which the information involves multiple topics and explore its impact on the engagement with misinformation on social media during COVID-19.

Particularly, at the onset of the COVID-19 pandemic, a myriad of conspiracy theories emerged, providing alternative explanations for the origins, spread, and treatment of the virus [24]. Examples include claims suggesting that the COVID-19 virus is a man-made bioweapon [24, 46]. Conspiracy theories often revolve around the notion of secret plots orchestrated by malevolent and powerful groups, deliberately crafted to achieve specific outcomes. This characteristic of conspiracy theories facilitates the connection of multiple topics through an illusory perception pattern, potentially fostering associative inference among individuals [40, 45, 76]. The integration of conspiracy theories within misinformation became more pronounced during the COVID-19 pandemic, marked by a significant increase in hyperlinks from misinformation websites to conspiracy theories [31]. Despite this observation, the question of whether misinformation containing conspiracy theories receives more engagement than other types of misinformation remains to be explored. In light of these considerations, this paper tries to address the current research gaps by investigating the roles of topic diversity and conspiracy theories in the engagement with misinformation during health crises.

3 Data and Methods

3.1 Dataset: MM-COVID

The data for our analysis is sourced from a publicly available repository on GitHub [42]. This repository provides an extensive COVID-19 dataset that consists of 8825 source news claims and their corresponding engagement on X. Initially, the repository gathers these news claims and their veracity verdicts from popular third-party fact-checking organizations worldwide, including Snopes and Poynter, as well as official health websites. To keep the quantity of the source claims in each language, the repository filters six languages: English, Spanish, Portuguese, Hindi, French, and Italian. This enables a broad and diverse examination of misinformation across linguistic and cultural contexts. Subsequently, online posts that explicitly reference these news sources are collected from X through its advanced search API. Notably, while the news sources in the dataset are linked to the original news URLs, precisely parsing the original news pages and obtaining their content is challenging due to the diverse range of news items collected from various languages and web pages. Instead, the dataset provides news claims, representing the main opinions and topics.

Table 1. COVID-19 topic spectrum. The 25 topics across 10 broad themes are included.

| Theme | Topic |
|-----------------------------------|---|
| Source (origin) | Outbreak, alternative causes |
| Prevention | Social distancing, disinfecting and cleanliness |
| Symptoms | Symptoms |
| Spread and growth | Modes of transmission, spread of cases, death reports |
| Treatment and recovery | Drugs and vaccines, therapies, alternative methods, testing |
| Impact on health care sector | Impact on hospitals and clinics, health policy, frontline workers |
| Government response | Travel restrictions, financial measures, lockdown regulations |
| Impact on the economy and markets | Shortage of products, panic buying, stock markets, employment, impact on business |
| Political impacts | Political impacts |
| Racism | Racism |

Data preprocessing: We begin by translating all the non-English texts in source news claims and their associated posts into English using Google Translate API via Googletrans.¹ Google Translate has a high performance in translating between English and non-English languages, retaining overall meaning in 82.5% of the translations [66, 73]. Subsequently, we perform text preprocessing by removing non-alphabetic and non-numerical characters, stop words, and URLs. We lemmatize all the words using the English transformer pipeline on Spacy and only consider news claims and posts that contain at least one lemmatized word. Additionally, we exclude posts that were posted before January 1, 2020. Due to the absence of release dates for source news items in the dataset, we use the first post that links to the specific source news item as its release date. In the end, we obtain 8822 news claims and associated 85,978 posts. Specifically, 59,363 posts are linked to 1543 false news claims, and 26,615 posts are linked to 7279 true news claims.

3.2 COVID-19 Topic Spectrum

A previous study analyzed over 13 million English COVID-19 posts and identified 26 topics within 10 broad themes, including source, prevention, spread and growth, treatment and recovery, impact on the economy and markets, impact on health care sector, and government response, political impact, and racism [4]. This research also provided keywords corresponding to each topic. To ensure a comprehensive understanding of the topic spectrum surrounding COVID-19 discussions, we utilize these keywords as seeds to identify similar words and construct a topic lexicon based on a Word2Vec model. Upon reevaluating the associations between topic categories and their themes, we find that the keywords in the topic “Hotspots and location” relate solely to places, such as “cities” and “location,” not adequately reflecting the theme “Spread and growth” in terms of transmission modes or case numbers. Therefore, in the main analysis, we remove the topic “Hotspots and location” and focus on the remaining 25 topics across 10 themes (Table 1).²

¹Googletrans library is available at <https://github.com/ssut/py-googletrans>. It is a free and unlimited Python library that implements Google Translate API and uses the same servers that translate.google.com uses.

²In a subsequent robustness check, we conduct the analysis again using all the original topics and keywords. The results are robust and consistently support our findings. The full estimation results are reported in Suppl. S3.

Table 2. The examples of topic seeds (bioweapon, vaccine, election, and quarantine) and their 10 most similar words.

| bioweapon | vaccine | election | quarantine |
|-------------------|--------------------|-----------------|----------------|
| biowarfare | vaccines | elections | isolation |
| biologicalweapon | vaccination | primaries | quarantined |
| bioweapons | vaccin | voting | selfquarantine |
| bioengineered | vax | 2020election | quarantine |
| biologicalwarfare | vaccine | electoral | selfisolation |
| bioterrorism | cure | election2020 | quarantining |
| engineered | inoculation | referendum | quarantine |
| weaponized | treatments | vote | lockdown |
| wuhanlab | coronavirusvaccine | votes | quarantine |
| biowar | antibodies | generalelection | isolating |

Word2Vec model: To train the Word2Vec model, we additionally collect a large-scale COVID-19 dataset from a public repository on Kaggle.³ This dataset contains 26,854,078 posts in multiple languages, posted from March 9 to April 30, 2020 on X. Given the size of the dataset, we only filter English posts and do not involve the translation process. By combining this dataset with MM-COVID, we obtain a total of 15,059,238 English posts. After removing non-alphabetic and non-numerical characters as well as URLs in the posts, we utilize the refined corpus to train a Word2vec model with a vector size of 200 using the gensim library.⁴ The examples of the keywords and their ten most similar words are shown in Table 2.

Topic lexicon: We expand our set of topic keywords by incorporating words that exhibit cosine similarities of more than 0.6 with at least one of the initial topic seeds. This process yields 828 additional topic keywords, and they are automatically categorized into the same topics as their corresponding topic seeds. Additionally, we conduct a manual review of these keywords and remove ambiguous terms and abbreviations. For example, we remove the keyword “reason” from the topic “Alternative causes,” as it does not specify a particular alternative cause for COVID-19, whereas the keyword “chinavirus” is an effective term for this topic category. Similarly, we remove ambiguous abbreviations such as “pres” in the topic “Political impacts.” Ultimately, we retain 90.6% of the 828 keywords and integrate them with the topic seeds to create an extensive topic lexicon.⁵

Topic identification: We utilize the topic lexicon to identify topics discussed in both news claims and their associated posts. As a result, 82.4% of news claims contained at least one topic. Additionally, our analysis considers both the content of the source claims and the content of the corresponding posts when identifying topics in online posts. Consequently, 91.4% of posts contain at least one topic.

3.3 Topic Diversity

Topic diversity indicates the extent to which the content involves multiple topics. To give a sense of topic diversity, we construct topic co-occurrence networks for true and false news claims. Specifically, we calculate the frequency of topic co-occurrence of any two topics out of 25 topics in true and false news claims, and retain the co-occurrence edges with frequencies exceeding the

³The dataset is available at <https://www.kaggle.com/datasets/smid80/coronavirus-covid19-tweets>.

⁴The model is trained on the High Performance Computing platform at the University of Luxembourg.

⁵We repeat our analysis using the keywords without manual validation. The results remain consistent, further enhancing the robustness of our findings. The full estimation results are reported in Suppl. S3.

median threshold. In the network of true news claims (Fig. 1a), the number of edges is 107, while in the network of false news claims (Fig. 1b), the number of edges is 167. Compared to true news claims, false news claims contain 56.1% more co-occurrence edges and are more interconnected, indicating higher interconnectedness and suggesting greater topic diversity in false news claims. For example, the topic “Political impacts” is connected to 5 topics in true news claims, whereas it connects to 10 additional topics, such as “Death reports” and “Drugs and vaccines,” in false news claims (Fig. 1c).

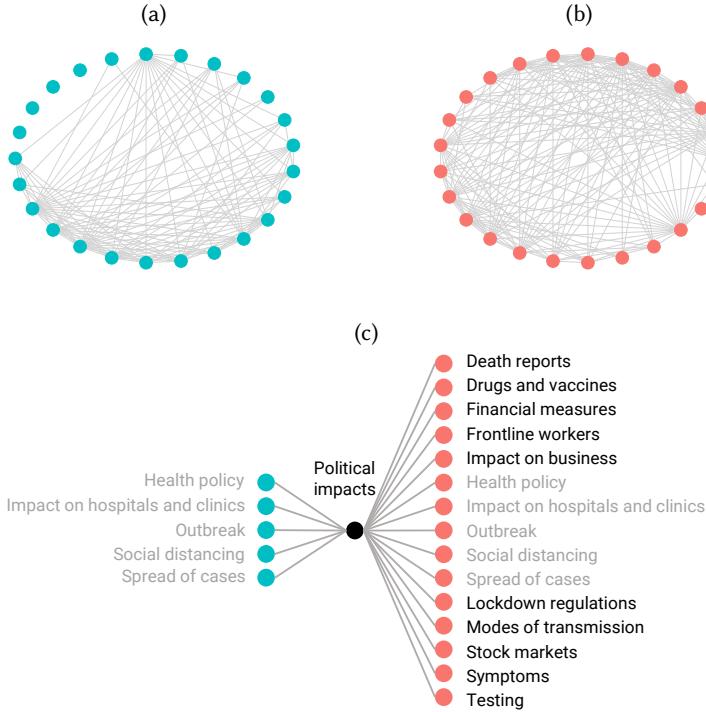


Fig. 1. Topic co-occurrence in true and false news claims. (a) Topic co-occurrence network in true news claims. The number of edges is 107. (b) Topic co-occurrence network in false news claims. The number of edges is 167. (c) An example of topic co-occurrences for “Political impacts.” Shown are the topic “Political impacts” and its co-occurring topics in true and false news claims.

Building upon previous research that used the Gini coefficient to measure the inequality of topic distribution in the content [45], we quantitatively measure topic diversity (*Diversity*) as:

$$Diversity = 1 - Gini = 1 - \frac{\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|}{2n^2 \bar{x}}, \quad (1)$$

where x_i denotes the number of words associated to topic i in each news claim (post), and n is the number of topics. In our analysis, we exclude news items and posts that do not relate to any topic, resulting in 7273 source news claims with corresponding 70,904 posts (see details in Table 4).

3.4 COVID-19 Conspiracy Theories

The outbreak of COVID-19 has fueled widespread conspiracy theories with misleading claims about its origin, treatment, and other related topics. We summarize the conspiracy theories identified in previous research [24, 46, 70] and use a set of keywords and phrases to identify conspiracy theories

Table 3. COVID-19 conspiracy theories and corresponding keywords (phrases). “()” indicates that the word can be separated into a phrase. “{} and {}” indicates one of the words in the first set {} needs to appear with one of the words in the second set {}. “5G” and “Bill Gates” are case sensitive.

| Conspiracy theory | Keywords |
|-------------------|---|
| Origin | bat, batsoup, weapon, biowar, lab, conspiracy, chinese(virus, biochemical, bioterrorism, biowarfare, china(virus, wildlife, ccp(virus, chinese(flu, china(coronavirus, chinese(coronavirus, chinesevirus19, chines(virus, manmade, wuhan(coronavirus, wuhan(coronavirus, wuflu, wuhan(virus, chinaliedpeopledie |
| Vaccines | {vaccine, vax, vaccination, vaccin, vaccine} and {anti, infertile, do not work, don't work, does not work, doesn't work, autism, autoimmune} |
| 5G | 5G |
| BigPharma | big pharma, fauci pharma, gates pharma |
| BillGates | Bill Gates |
| Exaggeration | does not exist, doesn't exist, exaggerated, inflated, refuse |
| FilmYourHospital | filmyourhospital, film your hospital, empty hospital, empty bed |
| GMO | gmo, genetically modified |

embedded in false news claims and subsequent posts. Specifically, we consider eight COVID-19 conspiracy theory categories that connect to at least one topic in our analysis: the sources and origins of the virus that caused the outbreak of COVID-19 (VirusOrigin), the connection between 5G and COVID-19 (5G), the role of big pharmaceutical companies (BigPharma) and Bill Gates (BillGates) in the spread of COVID-19, the claim that the scope of the pandemic was exaggerated (Exaggeration), the claim that hospitals were actually empty (FilmYourHospital), the idea that the genetically modified crops led to the emergence of COVID-19 (GMO), and the skepticism regarding the ineffectiveness and side-effects of vaccines (Vaccines).

We combine the conspiracy-specific terms adopted by prior studies [24, 46] and compile an extensive list of keywords and phrases to identify the eight conspiracy theories (Table 3). Subsequently, we match these keywords with the lemmatized words in each news claim or post, and for phrases, we directly match them with the original translated text. We find that 89 false news claims and 4265 false news posts incorporate keywords (phrases) indicating conspiracy theories. Given that the shares of conspiracy posts range from 0.6% to 18% in previous studies [24], the 6% share of posts containing conspiracy theories in our dataset is reasonable. Upon a manual check of the 89 source news items that contain keywords or phrases related to conspiracy theories, we find that the accuracy is 92.1%. Therefore, the selected keywords (phrases) can effectively identify conspiracy theories in our dataset.

3.5 User Engagement on X

On the platform of X, users can interact with posts they view through reposting, liking, or replying. Existing research typically uses the number of reposts to study engagement with misinformation on social media [10, 11, 19, 58, 78]. Additionally, the number of likes and the number of replies are also important indicators of engagement with online posts [10, 51]. However, few studies examine engagement with misinformation from the perspective of news source sharing. Multiple posts on X can link to the same source news items, which increases the visibility of the news items. In this paper, we consider user engagement at two levels: news source sharer level and post viewer

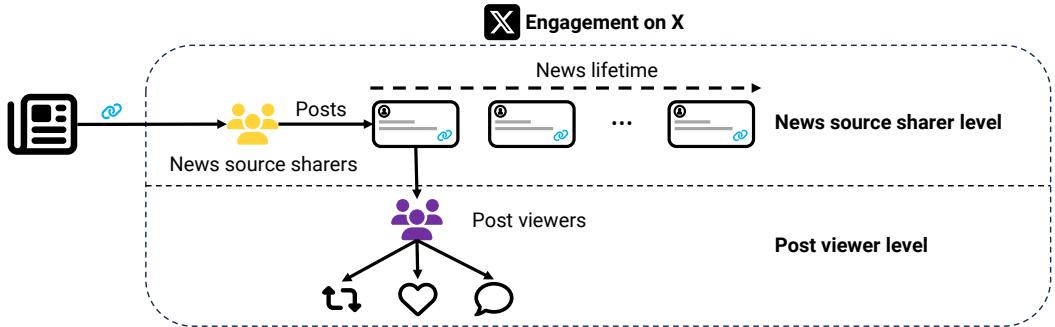


Fig. 2. Engagement on X at news source sharer level and post viewer level. The news source sharers view original news items and share them on X through posts containing news links. Subsequently, post viewers engage with these news posts by reposting, liking, or replying.

level (Fig. 2). At news source sharer level, news items are shared by news source sharers via posts on X. The engagement metrics include the number of posts and the lifetime of news items from the creation of the first post to the creation of the latest post. At post viewer level, engagement is measured by how viewers interact with these posts. Post viewers can repost, like, and reply to posts that are created by news source sharers and contain links to the original news items.

3.6 Empirical Models

We employ multiple explanatory regression models to explore our research questions separately. To this end, we define the following variables to measure *engagement at news source sharer level and post viewer level*.

- The engagement at news source sharer level:
 - *NewsLifetime*: The duration (in hours) from the creation of the first post to the creation of the latest post on X for the specific news item.
 - *NewsPostCount*: A count variable indicating the number of posts associated with the specific news item.
- The engagement at post viewer level:
 - *RepostCount*: A count variable indicating the number of reposts received by the specific post.
 - *LikeCount*: A count variable indicating the number of likes received by the specific post.
 - *ReplyCount*: A count variable indicating the number of replies received by the specific post.

Additionally, we define the following *key variables* to address our research questions:

- *Falsehood*: A dummy variable indicating whether the news claim is false (= 1) or true (= 0).
- *Conspiracy*: A dummy variable indicating whether the false news claim or post contains conspiracy theories (= 1) or not (= 0).
- *Diversity*: The topic diversity of the news claim or post.

We also control for the variables from *content characteristics* that can affect user engagement:

- *Words*: A continuous-like count variable indicating the number of words in the text of the news claim or post. This control variable is commonly used to account for the content richness [45].
- *Media*: A dummy variable indicating whether the post has media elements (= 1) or not (= 0). Previous research has found that media elements, such as images and videos, can facilitate user engagement [85].
- Emotions are considered to be the main driver of online information diffusion [11, 57, 58]. Our analysis considers six cross-cultural emotions – anger, disgust, fear, joy, sadness, and surprise –

out of the eight basic emotions [11, 55]. We employ the commonly used NRC emotion lexicon to count emotion words in news claims and posts. Notably, 92.1% news claims and 95% posts contain at least one emotion word. Subsequently, we calculate the weight of each emotion in the specific text content based on the number of corresponding emotion words. If the news claims (posts) have no emotion words, all the weights of emotions are set to 0.

Finally, the posters on social media play an important role in driving engagement with their posts [11, 78]. We use the following four variables to account for *poster characteristics*:

- *Verified*: A dummy variable indicating whether the post account is verified (= 1) or not (= 0).
- *AccountAge*: A continuous-like count variable indicating the age (in days) of the post account.
- *Followers*: A continuous-like count variable indicating the number of followers of the post account.
- *Followees*: A continuous-like count variable indicating the number of followees of the post account.

The descriptive statistics for the above variables are shown in Table 4.

Topic diversity (RQ1.1 & RQ2.1): We use *Diversity* as the dependent variable to analyze topic diversity in true and false news from original news claims to online posts. In terms of claims, the linear regression model that incorporates the main independent variables of *Falsehood* and *Conspiracy* is specified as:

$$Diversity_i = \beta_0 + \beta_1 Falsehood_i + \beta_2 Conspiracy_i + \beta_3 Words_i + \alpha' Emotions_i + u_{\text{lang}} + u_{\text{time}}, \quad (2)$$

where *Emotions* indicates six basic emotions. Additionally, u_{lang} denotes language-specific fixed effects, and u_{time} denotes month-year fixed effects.

Given that each news item can have multiple posts, we need to control for the possible heterogeneity among news items when analyzing the topic diversity in terms of online posts. Therefore, we specify a liner regression model incorporating news-specific random effects:

$$\begin{aligned} Diversity_i = & \beta_0 + \beta_1 Falsehood_i + \beta_2 Conspiracy_i + \beta_3 Words_i + \beta_4 Media_i + \beta_5 Verified_i \\ & + \beta_6 AccountAge_i + \beta_7 Followers_i + \beta_8 Followees_i + \alpha' Emotions_i \\ & + u_{\text{lang}} + u_{\text{time}} + v_{\text{news}}, \end{aligned} \quad (3)$$

where v_{news} indicates news-specific random effects. Additionally, we incorporate *Media*, poster characteristics as control variables. In the above two models, the coefficient estimate of *Falsehood* (β_1) is to address RQ1.1, and the coefficient estimate of *Conspiracy* (β_2) is to address RQ2.1.

Engagement at news source sharer level – news lifetime and post count (RQ1.2 & RQ2.2): The first layer of engagement is from source news claims to news posts on X. We examine how topic diversity and conspiracy theories influence engagement with misinformation in terms of news lifetime and post count. First, we take *NewsLifetime* as the dependent variable and specify a linear regression model as follows:

$$\begin{aligned} Lifetime_i = & \beta_0 + \beta_1 Falsehood_i + \beta_2 Conspiracy_i + \beta_3 Diversity_i + \beta_4 Falsehood_i \times Diversity_i \\ & + \beta_4 Words_i + \alpha' Emotions_i + u_{\text{lang}} + u_{\text{time}}, \end{aligned} \quad (4)$$

where the interaction term of *Falsehood* \times *Diversity* is to estimate the moderating effect of *Falsehood* on *Diversity*. Subsequently, we use a negative binomial regression model to explain *PostCount*. The dependent variable is $\log(E(PostCount_i|x_i))$, and the independent variables remain consistent with Eq. (4).

Engagement at post viewer level – reposts, likes, and replies (RQ1.2 & RQ2.2): To examine the effects of topic diversity and conspiracy theories on social engagement (i. e., reposts, likes, and replies) with misinformation on X, we specify a negative binomial regression model with

Table 4. Dataset overview. Reported are mean values or count numbers for the variables (standard deviations in parentheses). Column (1) includes all the news claims and associated posts. 51.7% in English, 21.6% in Spanish, 4.9% in Portuguese, 15.7% in Hindi, 3.2% in French and 3.0% in Italian. Column (2) includes false news claims (posts) that contain conspiracy theories.

| | (1) All | (2) CTs |
|-----------------------------|-----------------------------|----------------------------|
| News Count | 7,273 | 82 |
| Falsehood | 18.9% | 100% |
| <u>Engagement metrics</u> | | |
| Lifetime | 315.617 (900.617) | 1,968.141 (1,626.337) |
| Post Count | 70,904 | 4,265 |
| Repost Count | 1,300,597 | 161,823 |
| Like Count | 3,943,848 | 452,106 |
| Reply Count | 380,224 | 107,363 |
| <u>News characteristics</u> | | |
| Dates | 01/02/20 – 09/14/20 | 01/07/20 – 09/08/20 |
| Diversity | 0.083 (0.045) | 0.089 (0.038) |
| Words | 15.614 (6.364) | 11.841 (6.739) |
| Anger | 0.058 (0.109) | 0.040 (0.093) |
| Disgust | 0.045 (0.108) | 0.059 (0.153) |
| Fear | 0.183 (0.208) | 0.149 (0.183) |
| Joy | 0.065 (0.128) | 0.097 (0.215) |
| Sadness | 0.096 (0.139) | 0.089 (0.135) |
| Surprise | 0.045 (0.099) | 0.026 (0.080) |
| <u>Post characteristics</u> | | |
| Dates | 01/02/20 – 09/14/20 | 01/09/20 – 09/09/20 |
| Diversity | 0.104 (0.052) | 0.111 (0.050) |
| Words | 26.372 (10.192) | 27.006 (8.876) |
| Anger | 0.068 (0.100) | 0.069 (0.090) |
| Disgust | 0.074 (0.134) | 0.064 (0.105) |
| Fear | 0.195 (0.177) | 0.157 (0.146) |
| Joy | 0.062 (0.106) | 0.088 (0.163) |
| Sadness | 0.129 (0.133) | 0.095 (0.106) |
| Surprise | 0.052 (0.091) | 0.050 (0.088) |
| Media | 21.9% | 13.6% |
| Verified | 18.5% | 7.4% |
| AccountAge | 2,460.659 (1,457.714) | 2,259.347 (1,393.434) |
| Followers | 369,041,439 (1,755,545.224) | 56,847.547 (1,399,857.795) |
| Followees | 3,031.517 (15,099.135) | 2,361.965 (15,535.361) |

language-specific fixed effects, month-year fixed effects, and news-specific random effects:

$$\begin{aligned}
 \log(E(y_i|x_i)) = & \beta_0 + \beta_1 Falsehood_i + \beta_2 Conspiracy_i + \beta_3 Diversity_i + \beta_4 Falsehood_i \times Diversity_i \\
 & + \beta_3 Words_i + \beta_4 Media_i + \beta_5 Verified_i + \beta_6 AccountAge_i + \beta_7 Followers_i \\
 & + \beta_8 Followees_i + \alpha' \mathbf{Emotions}_i + u_{\text{lang}} + u_{\text{time}} + v_{\text{news}},
 \end{aligned} \tag{5}$$

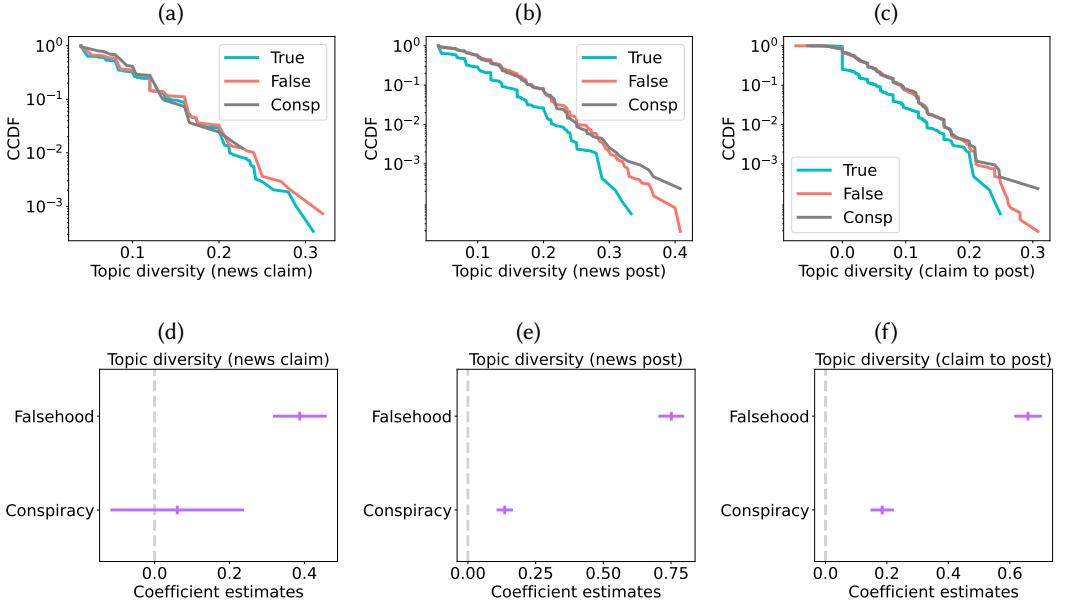


Fig. 3. The topic diversity from news claims to news posts. (a)–(c) The Complementary Cumulative Distribution Functions (CCDFs) for (a) topic diversity at the source claim level, (b) topic diversity at the post level, and (c) the changes in topic diversity from claims to posts. Three categories of CCDFs are included in each figure, i.e., true news (True), false news (False), and CT news (Consp). (d)–(f) The coefficient estimates of the independent variables – *Falsehood* and *Conspiracy*. Other independent variables are included during estimation but omitted in the visualization for better readability. Shown are mean values with error bars representing 95% Confidence Intervals (CIs). The dependent variables are (d) topic diversity in news claims, (e) topic diversity in news posts, and (f) the changes in topic diversity from claims to posts, respectively. See Table S1 in Suppl. S1 for full estimation results.

where x_i indicates all the independent variables, and y_i signifies the dependent variables, including $RepostCount_i$, $LikeCount_i$ and $ReplyCount_i$. In the above models for engagement, the coefficient estimates of *Falsehood* (β_1), *Diversity* (β_3), and *Falsehood* \times *Diversity* (β_4) are to address RQ1.2, and the coefficient estimate of *Conspiracy* (β_2) is to address RQ2.2. All the continuous variables in these models are z-standardized.

4 Empirical Results

4.1 Topic Diversity (RQ1.1 & RQ2.1)

4.1.1 Descriptive statistics. We start by examining topic diversity from news claims to news posts (Figs. 3a–3c). Fig. 3a shows that false news claims (mean of 0.086) exhibit 4.9% higher topic diversity compared to true news claims (mean of 0.082, $KS = 0.072$, $p < 0.001$). This difference is significantly magnified (43.0% higher) at the post level, where the average topic diversity in false news posts is 0.113, compared to 0.079 in true news posts (Fig. 3b, $KS = 0.333$, $p < 0.001$). Additionally, we examine the paired changes in topic diversity from claims to their corresponding posts (Fig. 3c), finding that topic diversity in false news items (increase of 0.031) increases 158.3% more than that in true news items (increase of 0.012, $KS = 0.396$, $p < 0.001$). For false news items containing conspiracy theories (CTs), the topic diversity at the claim level has no statistically significant difference compared to the overall false news items (Fig. 3a, $KS = 0.138$, $p = 0.095$). However, Fig.

3c indicates that the most significant increases in topic diversity from claims to posts are more likely to be in CT news items, compared to overall false news items ($KS = 0.044, p < 0.001$). This may lead to the significant difference in topic diversity between CT news and general false news at the post level (Fig. 3b, $KS = 0.036, p < 0.001$).

4.1.2 Regression results. We further conduct regression analysis to estimate the effects of *Falsehood* and *Conspiracy* on topic diversity (Figs. 3d–3f). In Fig. 3d, the coefficient estimate of *Falsehood* is significantly positive ($coef. = 0.388, p < 0.001$). This implies that the topic diversity in false news is, on average, 0.388 standard deviations higher than in true news at the claim level, consistent with the observations in Fig. 3a. Additionally, the coefficient estimate of *Conspiracy* is not statistically significant, which indicates that the topic diversity in false news claims containing conspiracy theories is not significantly different from that in false news claims without conspiracy theories (Fig. 3d). Subsequently, we explore the topic diversity between true and false news at the post level. In Fig. 3e, the coefficient estimate of *Falsehood* is significantly positive ($coef. = 0.751, p < 0.001$), suggesting that the topic diversity in false news is, on average, 0.751 standard deviations higher than in true news at the post level.

Notably, as observed in the CCDF in Fig. 3b and the coefficient estimates of *Falsehood* in Figs. 3d–3e, the difference between true and false news posts is larger than that between true and false news claims. Additionally, the coefficient estimate of *Conspiracy* in Fig. 3e becomes significantly positive ($coef. = 0.136, p < 0.001$). This suggests that false news posts containing conspiracy theories exhibit even higher topic diversity (0.136 standard deviations more) than false news posts without conspiracy theories. There are two possible reasons for the larger difference in topic diversity between true and false news at the post level compared to the claim level. One is that news claims with higher topic diversity receive more posts on X. Another is that false news has a bigger increase in topic diversity from the claim level to the post level, compared to true news. Given the second reason, we further examine the paired changes in topic diversity from the source news claims to the corresponding posts. In Fig. 3f, the positive coefficient estimate of *Falsehood* is statistically significant ($coef. = 0.659, p < 0.001$). This suggests that the increase in topic diversity in false news from the claim level to the post level is significantly larger (0.659 standard deviations more) than in true news. This finding could partially explain the enlarged difference in topic diversity between true and false news posts, compared to their source claims. In addition, the coefficient estimate of *Conspiracy* in Fig. 3f is significantly positive ($coef. = 0.185, p < 0.001$). This means that the topic diversity in false news that contains conspiracy theories increases 0.185 standard deviations more than that in false news without conspiracy theories from the claim level to the post level. This finding could partially explain why the coefficient estimate of *Conspiracy* in Fig. 3d is not statistically significant, whereas it becomes statistically significant in Fig. 3e.

In summary, we find that false news claims exhibit higher topic diversity compared to true news claims, and this difference is magnified at the post level (RQ1.1). Additionally, false news posts containing conspiracy theories have higher topic diversity than those without conspiracy theories, while no significant difference is observed at the claim level (RQ2.1).

4.2 News Source Sharer Engagement: News Lifetime and Post Count (RQ1.2 & RQ2.2)

4.2.1 Descriptive statistics. In terms of engagement at news source sharer level, we examine news lifetime that indicates the duration (in hours) from the creation time of the first post to the creation time of the last post for each news item on X. Fig. 4a shows that false news items have significantly longer lifetimes (mean of 1463.757 hours) than true news items (mean of 47.951 hours, $KS = 0.805, p < 0.000$). CT news items have longer lifetimes (mean of 1968.141 hours) than false news items ($KS = 0.167, p < 0.05$). Additionally, Fig. 4b shows that false news claims receive more posts (mean

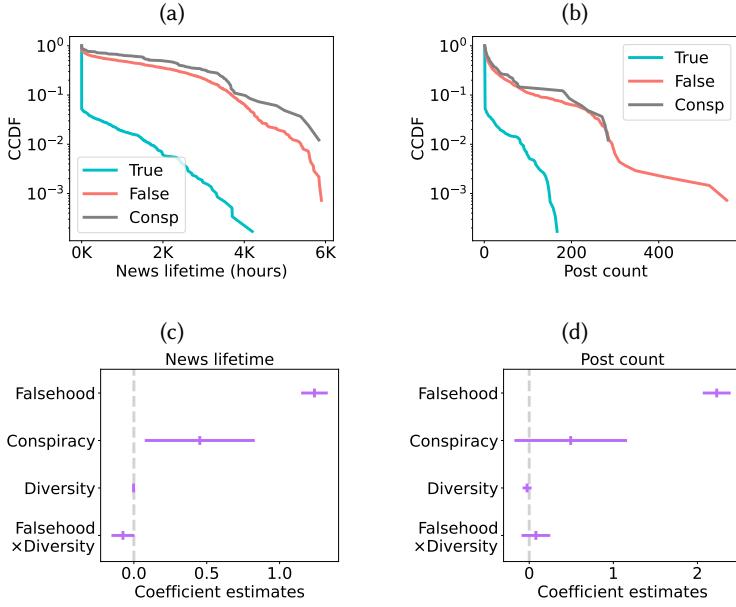


Fig. 4. The engagement at news source sharer level – news lifetime and post count. (a)–(b) The CCDFs for (a) news lifetime and (b) post count. Three categories of CCDFs are included in each figure, i.e., true news (True), false news (False), and CT news (Consp). (c)–(d) The coefficient estimates of the independent variables – *Falsehood*, *Conspiracy*, *Diversity*, and *Falsehood* \times *Diversity*. Other independent variables are included during estimation but omitted in the visualization for better readability. Shown are mean values with error bars representing 95% CIs. The dependent variables are (c) news lifetime and (d) post count, respectively. See Table S2 in Suppl. S1 for full estimation results.

of 38.015) than true news claims (mean of 3.159, $KS = 0.806, p < 0.001$). However, the number of posts received by CT news items (mean of 44.976) does not differ significantly from that received by false news items ($KS = 0.113, p = 0.257$).

4.2.2 Regression results. We conduct regression analysis to further rigorously examine the effects of *Falsehood*, *Conspiracy*, and *Diversity* on news lifetime and post count (Figs. 4c–4d). In Fig. 4c, the coefficient estimate of *Falsehood* is significantly positive ($coef. = 1.239, p < 0.001$). This suggests that, on average, false news items have lifetimes of 1.239 standard deviations longer for receiving posts on X compared to true news items. Additionally, the coefficient estimate of *Conspiracy* is also significantly positive ($coef. = 0.452, p < 0.05$). This means that, on average, the lifetimes for false news items containing conspiracy theories are 0.452 standard deviations longer than that for false news items without conspiracy theories. However, the coefficient estimates for *Diversity* and *Falsehood* \times *Diversity* are not statistically significant. This suggests that topic diversity has no significant impact on the lifetimes of news items, irrespective of their veracity.

Additionally, in Fig. 4d, the coefficient estimate of *Falsehood* is significantly positive ($coef. = 2.230, p < 0.001$), indicating that the number of posts referencing a specific source news item is significantly associated with its veracity. On average, false news items are linked by $e^{2.230} - 1 = 8.3$ times more posts on X compared to true news items. Notably, the coefficient estimate of *Conspiracy* is not statistically significant for post count, which implies that the integration of conspiracy theories has no additional impact on the number of posts received by false news items. The coefficient estimate of *Diversity* is not statistically significant. It suggests that the topic diversity

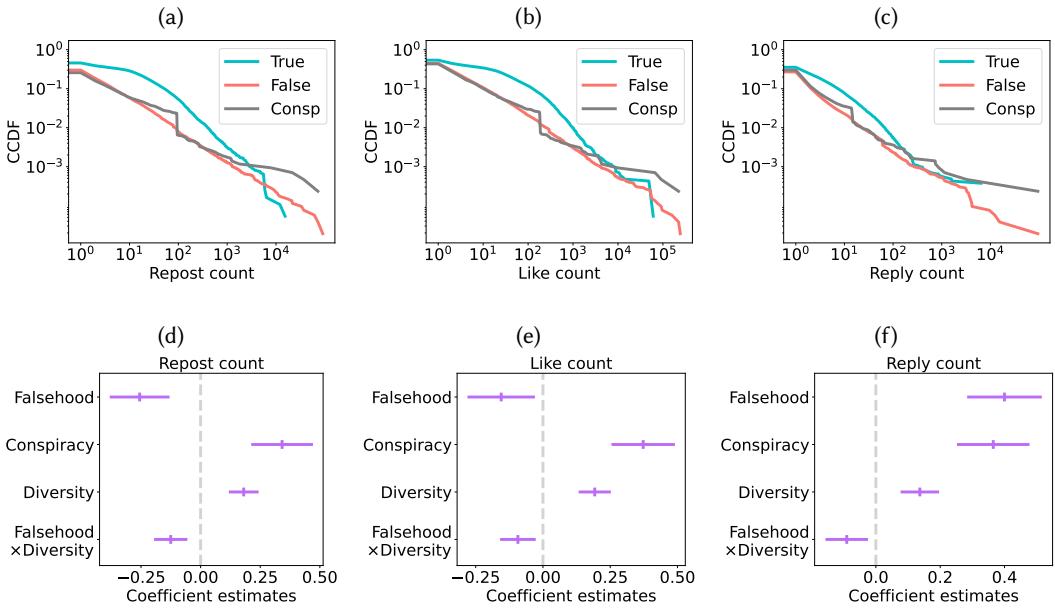


Fig. 5. The engagement at post viewer level – repost count, like count, and reply count. (a)–(c) The CCDFs for (a) repost count, (b) like count, and (c) reply count. Three categories of CCDFs are included in each figure, i. e., true news (True), false news (False), and CT news (Consp). (d)–(f) The coefficient estimates of the independent variables – *Falsehood*, *Conspiracy*, *Diversity*, and *Falsehood* \times *Diversity*. Other independent variables are included during estimation but omitted in the visualization for better readability. Shown are mean values with error bars representing 95% CIs. The dependent variables are (d) repost count, (e) like count, and (f) reply count, respectively. See Table S3 in Suppl. S1 for full estimation results.

of a source news claim has no significant association with the number of online posts it receives. This finding does not support the first reason for the enlarged difference in topic diversity between true and false news posts, compared to that between true and false news claims (see Section 4.1.2). Furthermore, the coefficient estimate of the interaction term *Falsehood* \times *Diversity* is also not statistically significant, reinforcing the finding that topic diversity has no significant effect on post count, irrespective of the veracity of news claims.

In summary, we find that, at news source sharer level, misinformation with higher topic diversity does not receive more engagement in terms of news lifetime and post count (RQ1.2). However, false news items containing conspiracy theories tend to receive a higher number of posts on X compared to false news items without conspiracy theories but do not have significantly longer lifetimes (RQ2.2).

4.3 Post Viewer Engagement: Reposts, Likes, and Replies (RQ1.2 & RQ2.2)

4.3.1 Descriptive statistics. In terms of engagement at post viewer level, we examine the numbers of reposts, likes, and replies for news posts. To this end, we first conduct descriptive analysis without incorporating the defined control variables from content and poster characteristics. We find that, on average, true news posts receive more engagement on X than false news posts in terms of reposts (Fig. 5a; true: 29.396, false: 14.403, $KS = 0.240, p < 0.001$), likes (Fig. 5b; true: 84.765, false: 45.234, $KS = 0.235, p < 0.001$), and replies (Fig. 5c; true: 6.294, false: 5.031, $KS = 0.133, p < 0.001$). However, it is worth noting that news posts that receive the most significant reposts, likes,

or replies belong to the false category. Additionally, CT news posts receive more reposts (mean of 37.942, $KS = 0.043, p < 0.001$) and replies (mean of 25.173, $KS = 0.031, p < 0.01$) than overall false news posts, while without statistically significant difference in like count ($KS = 0.017, p = 0.194$).

4.3.2 Regression results. Incorporating multiple control variables related to content and poster characteristics that can affect engagement with news posts, we conduct regression analysis to estimate the effects of *Falsehood*, *Conspiracy*, and *Diversity* on the numbers of reposts, likes, and replies (Figs. 5d–5f). In Fig. 5d, the coefficient estimate of *Falsehood* is significantly negative ($coef. = -0.256, p < 0.001$), which means that false news posts receive fewer reposts than true news posts. On average, false news posts receive 22.6% fewer reposts than true news posts. Similarly, the coefficient estimate of *Falsehood* in Fig. 5e is significantly negative ($coef. = -0.155, p < 0.05$). This means that, on average, false news posts receive 14.4% fewer likes than true news posts. However, the coefficient estimate of *Falsehood* in Fig. 5f is significantly positive ($coef. = 0.400, p < 0.001$). This means that, on average, false news posts receive 49.2% more replies than true news posts.

Additionally, the coefficient estimates of *Diversity* in Fig. 5d ($coef. = 0.181, p < 0.001$), Fig. 5e ($coef. = 0.192, p < 0.001$), and Fig. 5f ($coef. = 0.137, p < 0.001$) are significantly positive. This indicates that heightened topic diversity is associated with increased online engagement in reposts, likes, and replies. Specifically, a one standard deviation increase in topic diversity is associated with 19.8% more reposts, 21.2% more likes, and 14.7% more replies. However, the coefficient estimates of *Falsehood* \times *Diversity* in Fig. 5d ($coef. = -0.125, p < 0.001$), Fig. 5e ($coef. = -0.093, p < 0.01$), and Fig. 5f ($coef. = -0.091, p < 0.01$) are significantly negative.⁶ This suggests that the positive association of topic diversity with engagement on X is not significantly stronger within the context of false news posts compared to true news posts. On the contrary, the coefficient estimates of *Conspiracy* in Fig. 5d ($coef. = 0.342, p < 0.001$), Fig. 5e ($coef. = 0.373, p < 0.001$), and Fig. 5f ($coef. = 0.365, p < 0.001$) are significantly positive. This means that the integration of conspiracy theories within false news posts is positively linked to engagement at post viewer level. Specifically, false news posts that contain conspiracy theories receive 40.8% more reposts, 45.2% more likes, and 44.1% more replies compared to false news posts without conspiracy theories.

Clarification for the discrepancies between descriptive statistics and regression results: We observe two discrepancies between descriptive analysis and regression analysis: (i) In the descriptive analysis, Fig. 5c shows that true news posts receive more replies than false news posts, whereas Fig. 5f indicates that false news posts receive more replies than true news posts. (ii) No significant difference between CT news posts and overall false news posts is observed in the CCDF of like count (Fig. 5b), while Fig. 5e indicates that false news posts containing conspiracy theories receive more likes than those without conspiracy theories. These discrepancies suggest that the falsehood may correlate with other variations in the control variables, potentially complicating the findings in the descriptive analysis and underscoring the necessity of incorporating control variables that can affect engagement.

For instance, emotions, included as control variables in our regression models, have been well documented as a significant driver in facilitating online post sharing [8, 11, 33, 62, 71, 78]. We find that falsehood is positively correlated with anger ($r = 0.038, p < 0.001$) and surprise ($r = 0.051, p < 0.001$) but negatively correlated with joy ($r = -0.020, p < 0.001$; see Table S6 in Suppl. S2). This suggests that false news posts express more anger and surprise but less joy compared to true news posts. Within the estimation results of our regression models, the coefficient estimates of *Anger* for repost count ($coef. = 0.049, p < 0.001$), like count ($coef. = 0.059, p < 0.001$), and reply count ($coef. = 0.040, p < 0.01$) are consistently positive and statistically significant (see Table S3

⁶The coefficient estimates of *Diversity* across repost count, like count, and reply count remain significantly positive in the subset analysis for false news posts. See Table S4 in Suppl. S1 for full estimation results.

in Suppl. S1 for full estimation results). Similarly, the coefficient estimates of *Surprise* for repost count ($coef. = 0.093, p < 0.001$), like count ($coef. = 0.120, p < 0.001$), and reply count ($coef. = 0.121, p < 0.001$) are also consistently positive and statistically significant. Our results suggest that news posts expressing higher levels of anger or surprise tend to receive more reposts, likes, and replies, aligning with findings from previous research [11, 78]. However, the coefficient estimate of *Joy* is not statistically significant for repost count, while significantly positive for reply count ($coef. = 0.077, p < 0.001$). This suggests that news posts expressing more joy are more likely to receive replies but do not receive more reposts. In light of this, we find two opposite effects: falsehood contributes to an increase in replies for false news posts, while the lower expression of joy in false news posts, compared to true news posts, links to fewer replies. The aggregation of these two opposite effects in the descriptive analysis could partially explain the discrepancy observed between Fig. 5c and Fig. 5f.

In summary, at post viewer level, news posts with higher topic diversity receive more reposts, likes, and replies; however, this association is not stronger for false news posts (RQ 1.2). Additionally, false news posts containing conspiracy theories receive more reposts, likes, and replies compared to false news posts without conspiracy theories (RQ 2.2).

4.4 Robustness Checks

To ensure the robustness of our findings, we conduct several additional checks: (i) We examine the variance inflation factors (VIFs) among the independent variables at both news claim and news post levels (see Table S7 in Suppl. S2). All the VIFs are significantly lower than the critical threshold of 4, which indicates that multicollinearity is not an issue for our analysis. (ii) To test whether our findings are robust with potential noise in topic keywords, we conduct our analysis again using the original 26 topics and associated keywords without manual validation (see details in Suppl. S3). (iii) To ensure that our findings are not influenced by the topic granularity, we recalculate topic diversity at the theme level and repeat our analysis (see details in Suppl. S4). (iv) We further balance the samples between true and false news items through propensity score matching and reduce potential unobserved confounding factors (see details in Suppl. S5). (v) We extend conspiracy theories by including QAnon and repeat our analysis (see details in Suppl. S6).⁷ (vi) Given that news posts can also debunk linked false news items, we exclude potential debunking posts and repeat our analysis (see details in Suppl. S7). (vii) We use an alternative topic modeling method (BERTopic) to cluster topics again and repeat our analysis (see details in Suppl. S8). All the results are robust and consistently support our main findings.

4.5 Summary of Main Findings

Our analysis yields several key findings: (i) False news, especially when accompanied by conspiracy theories, exhibits higher topic diversity compared to true news. This discrepancy increases from source news claims to news posts on X. (ii) At news source sharer level, neither the news lifetime nor the number of news posts is significantly associated with topic diversity. However, false news has a longer lifetime and receives more posts compared to true news. Additionally, the integration of conspiracy theories is associated with a longer lifetime for false news but has no significant association with the number of false news posts. (iii) At post viewer level, posts with higher topic

⁷QAnon is a far-right conspiracy theory alleging that a cabal of politicians and media personalities are engaged in a war to undermine society [3, 37, 47, 52]. The QAnon discourse has become more diverse and evolved into a global phenomenon during the COVID-19 pandemic [34]. We adopt a list of QAnon-specific keywords from previous research [69] to identify QAnon conspiracy theory in our dataset. The performance of this keyword list in identifying QAnon discourse has been manually validated, showing high agreement with human assessments [3]. As a result, we identify 6 QAnon news claims and 302 QAnon news posts.

diversity receive more reposts, likes, and replies, while this association is not stronger in false news posts. (iv) The integration of conspiracy theories is linked to more engagement with COVID-19 misinformation at post viewer level. False news posts that contain conspiracy theories, on average, receive 40.8% more reposts, 45.2% more likes, and 44.1% more replies compared to false news posts without conspiracy theories.

5 Discussion

5.1 Relevance

The spread of misinformation surrounding health crises is a serious public concern, intersecting multiple domains in society. Here, we present an empirical analysis of 7273 fact-checked source news claims and their associated posts on X to explore how topic diversity and conspiracy theories influence engagement with multi-topic misinformation during the COVID-19 pandemic. Understanding the characteristics of misinformation and their effects on user engagement is a crucial and growing area of research for developing effective countermeasures and mitigating the spread of misinformation [10, 11, 20, 21, 78]. Our study makes two main contributions: (i) We analyze the characteristics of misinformation and its spread from the perspectives of topic diversity and conspiracy theories; (ii) We examine how these characteristics shape engagement with misinformation at two distinct user levels, i. e., news source sharer level and post viewer level. Specifically, we investigate online user engagement from the shares of source news items by news source sharers to the subsequent engagement with news posts by post viewers on X (i. e., reposts, likes, and replies). Our findings offer valuable insights for further research and practical strategies for combating the spread of online misinformation during health crises and beyond.

5.2 Research Implications

5.2.1 Heightened topic diversity in misinformation/conspiracy theories. The observed high topic diversity in COVID-19 misinformation suggests that the interconnection of multiple topics becomes a characteristic of misinformation narratives during health crises. Here, the topic diversity reflects the interconnectedness of various topics within a single piece of misinformation, potentially enhancing its novelty. Different from our study, Vosoughi et al. [78] investigated the topic similarity between users' reposted true (or false) posts and their previously viewed posts. They found that false posts are more dissimilar from previously viewed posts compared to true posts. However, given users' topic interests and recommendation algorithms' tendency to promote similar content, users potentially consume more content that aligns with their interests [14, 23]. These seemingly conflicting findings could be partially reconciled by our insights on topic diversity: misinformation may strategically link familiar topics together with diverse, novel topics, thereby drawing users' attention to new information and increasing their susceptibility. Particularly, our finding that misinformation containing conspiracy theories exhibits the strongest topic diversity underscores this deliberate strategy in conspiracy narratives. Moreover, previous research reveals coordinated efforts among misinformation websites from multiple domains to hyperlink to each other, further supporting our interpretation [31].

Additionally, we find that the topic diversity in COVID-19 misinformation significantly increases from its source claims to subsequent online posts. This mutation/evolution may contribute to the adaptability of misinformation, allowing it to resonate with diverse audiences and contexts [82]. For instance, QAnon is a far-right political conspiracy theory that originated in 2017. It became more diverse during COVID-19, evolving and expanding beyond political themes to incorporate pandemic-related topics [34]. Theoretically, within the context of multi-topic misinformation, individuals tend to adaptively connect related information and make associative inferences to

reduce uncertainty and gain control in a way that is congruent to their prior knowledge [40, 45]. Understanding this cognitive process is essential for developing targeted interventions aimed at reducing susceptibility to multi-topic misinformation, especially when accompanied by conspiracy theories. Taken together, our findings underscore the importance of considering the multi-topic nature of misinformation in further research.

5.2.2 Engagement patterns at news source sharer level and post viewer level. During health crises marked by high uncertainty, novel information can update users' understanding and knowledge of the current state, helping them manage anxiety and fostering their engagement intentions [7, 11, 28, 78]. Furthermore, sharing novel posts can make users feel that they possess rare and important information that others may lack, thereby boosting their self-esteem [11, 18]. In this light, our findings suggest that the high topic diversity and the incorporation of conspiracy theories in misinformation can enhance its perceived novelty, thereby driving engagement. However, through further examining how topic diversity and conspiracy theories shape engagement at news source sharer level and post viewer level, we find distinct engagement patterns at the two user levels.

Engagement patterns at news source sharer level: At news source sharer level, the observed longer lifetime and higher post count on X for false news items, compared to true news items, indicate the persistence and widespread reach of misinformation on social media during COVID-19. The persistence of false news items provides them with more opportunities to accumulate engagement on social media, potentially reaching a broader audience. Moreover, the lifetime and post count appear to be largely attributed to the veracity of news claims, as opposed to other variables including *Conspiracy* and *Diversity*. Given this, it is plausible that the prolonged lifetime and high post count for false news items mainly result from deliberately planned dissemination strategies by misinformation providers, involving means such as social bots and monetization [2, 46, 75]. This perspective also aligns with the prior finding that social bots amplify low-credibility content during the initial spreading stages [67]. Consequently, the role of individuals' self-motivated sharing behaviors in the persistence and diffusion of COVID-19 misinformation is limited at news source sharer level. However, massive posts repeating the same false claims with adaptability – i. e., through increased topic diversity – can still heighten individuals' susceptibility and beliefs. In this regard, social media platforms should take measures to restrict repeated posts linking to the same false news items and their mutations.

Engagement patterns at post viewer level: As opposed to engagement at news source sharer level, engagement at post viewer level (i. e., reposts, likes, and replies) on X might be largely driven by organic interactions from human users rather than social bots during COVID-19 [75]. Accordingly, we find that topic diversity and conspiracy theories become significant in driving engagement with misinformation at post viewer level in terms of reposts, likes, and replies. To efficiently reduce human users' engagement with misinformation at this level, social media platforms need to prioritize targeting misinformation that exhibits high topic diversity and incorporates conspiracy theories.

Additionally, our analysis reveals that replies differ from reposts and likes in their engagement with misinformation. Specifically, excluding the effects of topic diversity, conspiracy theories, and other control variables, false news posts receive more replies but fewer reposts and likes than true news posts. Distinct from reposts and likes, replies can not only support but also debunk false news posts. Given this, the higher number of replies for false news posts than true news posts can be triggered by debunking misinformation [83]. Our supplementary analysis of quotes shows that, although not statistically significant, false news posts also have a trend to receive more quotes than true news posts (see Suppl. S9 for full estimation results). This finding suggests that quotes serve a similar function as replies.

Taken together, our findings suggest that engagement with misinformation might be primarily driven by social bots at news source sharer level, while human users dominate engagement at post viewer level. This distinction highlights the need to examine misinformation engagement and develop intervention strategies at a multi-level framework.

5.3 Practical Implications

5.3.1 Implementing cross-platform fact-checking. Currently, social media providers are actively implementing measures to control the spread of misinformation on their platforms, including third-party professional fact-checking by experts and community-based fact-checking driven by crowds [10, 21]. However, these content moderation policies/features primarily function within individual platforms. Our findings suggest that misinformation from external sources, even if fact-checked on one platform, can still enter another platform through a large volume of posts. To prevent this spillover of misinformation across platforms, coordinated efforts among social media platforms, such as sharing fact-checking information, are needed.

5.3.2 Expanding fact-checks to misinformation with same links. Given that misinformation can be repeatedly shared on social media platforms through the same source links, an effective strategy is to directly moderate posts that contain previous fact-checked misinformation links. For example, X has rolled out a Community Notes feature that allows users to write fact-checking notes as additional context displayed on the corresponding misleading posts.⁸ This community-based fact-checking approach has proven effective in reducing the spread of misleading posts [9]. A recent update to Community Notes on X ensures that helpful notes that indicate posts' media elements, such as images or videos, are misleading will be visible on all posts containing the same media. Similarly, fact-checking notes about misinformation links could be expanded to cover all posts that contain the same links.

5.3.3 Empowering logical reasoning and critical thinking for social media users. To reduce users' susceptibility to misinformation with high topic diversity and promote rational information processing, tools that can enhance users' logical reasoning and critical thinking abilities should be adopted. The rise of Large Language Models (LLMs) presents a promising opportunity to help users conduct logical reasoning and identify potentially conflicting or fictional information. Although current LLMs may still have limitations in reasoning tasks, particularly for misinformation identification [53], this approach represents a promising direction as LLMs continue to improve [84].

5.4 Limitations and Further Research

Our study is subject to several limitations. First, our analysis is based on real-world social media data and, as is typical with observational studies, is limited to reporting associations between variables rather than causal claims. Future research could benefit from experimental designs and user studies to establish causal relationships and explore heterogeneity in the effects of topic diversity and conspiracy theories on engagement across different user groups.

Second, our findings on engagement can be influenced by social media recommendation algorithms and potential soft moderation implemented on misleading posts. Specifically, while our study mainly focuses on user-driven engagement at two levels and discusses the potential role of social bots, social media algorithms can selectively recommend posts and boost their visibility, potentially influencing user engagement and adding complexity to our results. Given that posts linking to the same news items can be treated similarly by recommendation algorithms or have similar audiences, the incorporation of news-specific random effects, i. e., random intercepts, in

⁸More details are available at <https://communitynotes.x.com/guide/en/about/introduction>

our models helps mitigate this concern. However, the precise role of recommendation algorithms warrants further exploration in future research.

Additionally, due to the absence of information indicating whether and when soft moderation was applied to certain misleading posts, we cannot control for the potential effect of soft moderation on subsequent engagement. Previous research has shown that soft moderation interventions – e.g., attaching warning tags – adopted by X during COVID-19 can reduce belief in misinformation and thus lower the spread/visibility of misleading posts [68]. Moreover, a recent study demonstrates that community-based fact-checking approach is effective in reducing the spread of misleading posts on X [9]. With the potential impact of such interventions in reducing engagement, we still find that false news has a longer lifetime than true news. This indicates that our results are not overestimated. On the other hand, prior research suggests that warning labels can also prompt interactions aimed at debunking false claims [83]. To address this concern, we exclude potential debunking posts based on debunk-specific keywords and conduct our analysis again. Our results are robust and consistently support our main findings.

Furthermore, the absence of detailed engagement timelines at post viewer level in our dataset hinders a more granular analysis of how topic diversity and conspiracy theories influence engagement with misinformation over time on social media. This limitation highlights an important area for further examination.

Finally, our analysis is based on a single curated public dataset that includes social context on one platform – X [42]. Although this dataset offers unique features, such as multilingual content, it may introduce potential selection bias during data collection. To the best of our knowledge, no other publicly available datasets are suitable for addressing our research questions. Moreover, current restrictions on social media platform APIs make it increasingly challenging for researchers to access comprehensive data. We advocate for greater openness from social media platforms to facilitate academic research and encourage further studies to validate and generalize our findings across broader datasets and other social media platforms.

6 Conclusion

The spread of health misinformation on social media, especially during health crises such as COVID-19, poses unprecedented threats to our society. It is crucial to understand the factors and mechanisms that contribute to the engagement with health misinformation on social media. Here, we provide a comprehensive analysis of how topic diversity and conspiracy theories shape engagement with COVID-19 misinformation from news source sharer level to post viewer level. We find that high topic diversity and the integration of conspiracy theories are characteristics of COVID-19 misinformation. Topic diversity has no significant association with the engagement at news source sharer level. However, topic diversity and conspiracy theories can significantly enhance engagement with COVID-19 misinformation at post viewer level. Our findings provide insights into understanding the engagement with health misinformation and highlight the engagement patterns for news source sharers and post viewers regarding topic diversity and conspiracy theories. These insights are valuable for developing targeted interventions at both user levels during health crises.

7 Ethics Statement

This research has received ethical approval from the Ethics Review Panel of the University of Luxembourg (ref. ERP 23-053 REMEDIS). All analyses are based on publicly available data. We declare no competing interests.

Acknowledgments

This research is supported by the Luxembourg National Research Fund (FNR) and Belgian National Fund for Scientific Research (FNRS), as part of the project REgulatory Solutions to MitigatE DISinformation (REMEDIS), grant ref. INTER_FNRS_21_16554939_Remedis.

References

- [1] Hunt Allcott and Matthew Gentzkow. 2017. Social media and fake news in the 2016 election. *Journal of Economic Perspectives* 31, 2 (2017), 211–236.
- [2] Cameron Ballard, Ian Goldstein, Pulak Mehta, Genesis Smothers, Kejsi Take, Victoria Zhong, Rachel Greenstadt, Tobias Lauinger, and Damon McCoy. 2022. Conspiracy brokers: Understanding the monetization of YouTube conspiracy theories. In *Proceedings of the ACM Web Conference 2022*. 2707–2718.
- [3] Dominik Bär, Nicolas Pröllochs, and Stefan Feuerriegel. 2023. Finding Qs: Profiling QAnon supporters on Parler. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 17. 34–46.
- [4] Ranganathan Chandrasekaran, Vikalp Mehta, Tejali Valkunde, and Evangelos Moustakas. 2020. Topics, trends, and sentiments of tweets about the COVID-19 pandemic: Temporal infoveillance study. *Journal of Medical Internet Research* 22, 10 (2020), e22624.
- [5] Marina Charquero-Ballester, Jessica G Walter, Ida A Nissen, and Anja Bechmann. 2021. Different types of COVID-19 misinformation have different emotional valence on Twitter. *Big Data & Society* 8, 2 (2021), 20539517211041279.
- [6] Canyu Chen, Haoran Wang, Matthew Shapiro, Yunyu Xiao, Fei Wang, and Kai Shu. 2022. Combating health misinformation in social media: Characterization, detection, intervention, and open issues. ArXiv.
- [7] Lucia L. Chen, Steven R. Wilson, Sophie Lohmann, and Daniela V. Negraia. 2023. What are you anxious about? Examining subjects of anxiety during the COVID-19 pandemic. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 17. 137–148.
- [8] Yuwei Chuai, Yutian Chang, and Jichang Zhao. 2022. What really drives the spread of COVID-19 tweets: A revisit from perspective of content. In *2022 IEEE 9th International Conference on Data Science and Advanced Analytics*. 1–10.
- [9] Yuwei Chuai, Moritz Pilarski, Thomas Renault, David Restrepo-Amariles, Aurore Troussel-Clément, Gabriele Lenzini, and Nicolas Pröllochs. 2024. Community-based fact-checking reduces the spread of misleading posts on social media. ArXiv.
- [10] Yuwei Chuai, Haoye Tian, Nicolas Pröllochs, and Gabriele Lenzini. 2024. Did the roll-out of Community Notes reduce engagement with misinformation on X/Twitter? *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW2 (2024), 1–52.
- [11] Yuwei Chuai and Jichang Zhao. 2022. Anger can make fake news viral online. *Frontiers in Physics* 10 (2022), 970174.
- [12] Yuwei Chuai, Jichang Zhao, Nicolas Pröllochs, and Gabriele Lenzini. 2025. Is fact-checking politically neutral? Asymmetries in how U.S. fact-checking organizations pick up false statements mentioning political elites. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 19. 403–429.
- [13] Matteo Cinelli, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi, and Michele Starnini. 2021. The echo chamber effect on social media. *Proceedings of the National Academy of Sciences* 118, 9 (2021), e2023301118.
- [14] Federico Cinus, Marco Minici, Corrado Monti, and Francesco Bonchi. 2022. The effect of people recommenders on echo chambers and polarization. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 16. 90–101.
- [15] Garth Davies, Edith Wu, and Richard Frank. 2023. A witch's brew of grievances: The potential effects of COVID-19 on radicalization to violent extremism. *Studies in Conflict & Terrorism* 46, 11 (2023), 2327–2350.
- [16] Ehsan Dehghan and Ashwin Nagappa. 2022. Politicization and radicalization of discourses in the alt-tech ecosystem: A case study on Gab Social. *Social Media + Society* 8, 3 (2022), 20563051221113075.
- [17] Karen M. Douglas and Robbie M. Sutton. 2018. Why conspiracy theories matter: A social psychological analysis. *European Review of Social Psychology* 29, 1 (2018), 256–298.
- [18] Karen M Douglas, Joseph E Uscinski, Robbie M Sutton, Aleksandra Cichocka, Turkyay Nefes, Chee Siang Ang, and Farzin Deravi. 2019. Understanding conspiracy theories. *Political Psychology* 40 (2019), 3–35.
- [19] Chiara Drolsbach and Nicolas Pröllochs. 2023. Diffusion of community fact-checked misinformation on Twitter. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–22.
- [20] Ullrich Ecker, Jon Roozenbeek, Sander van der Linden, Li Qian Tay, John Cook, Naomi Oreskes, and Stephan Lewandowsky. 2024. Misinformation poses a bigger threat to democracy than you might think. *Nature* 630, 8015 (2024), 29–32.
- [21] Ullrich KH Ecker, Stephan Lewandowsky, John Cook, Philipp Schmid, Lisa K Fazio, Nadia Brashier, Panayiota Kendeou, Emily K Vraga, and Michelle A Amazeen. 2022. The psychological drivers of misinformation belief and its resistance

to correction. *Nature Reviews Psychology* 1, 1 (2022), 13–29.

- [22] Adam M. Enders, Joseph Uscinski, Casey Klofstad, and Justin Stoler. 2022. On the relationship between conspiracy theory beliefs, misinformation, and vaccine hesitancy. *PLoS ONE* 17, 10 (2022), e0276082.
- [23] Kristen Engel, Shruti Phadke, and Tanushree Mitra. 2023. Learning from the ex-believers: Individuals' journeys in and out of conspiracy theories online. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–37.
- [24] Dmitry Erokhin, Abraham Yosipof, and Nadejda Komendantova. 2022. COVID-19 conspiracy theories discussion on Twitter. *Social Media + Society* 8, 4 (2022), 20563051221126051.
- [25] Max Falkenberg, Alessandro Galeazzi, Maddalena Torricelli, Niccolò Di Marco, Francesca Larosa, Madalina Sas, Amin Mekacher, Warren Pearce, Fabiana Zollo, Walter Quattrociocchi, and Andrea Baronchelli. 2022. Growing polarization around climate change on social media. *Nature Climate Change* 12, 12 (2022), 1114–1121.
- [26] Justin Farrell. 2019. The growth of climate change misinformation in US philanthropy: evidence from natural language processing. *Environmental Research Letters* 14, 3 (2019), 034013.
- [27] Helen Fischer, Markus Huff, and Nadia Said. 2022. Polarized climate change beliefs: No evidence for science literacy driving motivated reasoning in a U.S. national study. *American Psychologist* 77 (2022), 822–835.
- [28] Isabelle Freiling, Nicole M Krause, Dietram A Scheufele, and Dominique Brossard. 2023. Believing and sharing misinformation, fact-checks, and accurate information on social media: The role of anxiety during COVID-19. *New Media & Society* 25, 1 (2023), 141–162.
- [29] R. Kelly Garrett and Robert M. Bond. 2021. Conservatives' susceptibility to political misperceptions. *Science Advances* 7, 23 (2021), eabf1234.
- [30] Jon Green, William Hobbs, Stefan McCabe, and David Lazer. 2022. Online engagement with 2020 election misinformation and turnout in the 2021 Georgia runoff election. *Proceedings of the National Academy of Sciences* 119, 34 (2022), e2115900119.
- [31] Hans W. A. Hanley, Deepak Kumar, and Zakir Durumeric. 2023. A golden age: Conspiracy theories' relationship with misinformation outlets, news media, and the wider Internet. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–33.
- [32] P. Sol Hart, Sedona Chinn, and Stuart Soroka. 2020. Politicization and polarization in COVID-19 news coverage. *Science Communication* 42, 5 (2020), 679–697.
- [33] Christy Galletta Horner, Dennis Galletta, Jennifer Crawford, and Abhijeet Shirsat. 2021. Emotions: The unexplored fuel of fake news on social media. *Journal of Management Information Systems* 38, 4 (2021), 1039–1066.
- [34] Mohamad Hoseini, Philipe Melo, Fabricio Benevenuto, Anja Feldmann, and Savvas Zannettou. 2023. On the globalization of the QAnon conspiracy theory through Telegram. In *Proceedings of the 15th ACM Web Science Conference 2023*. 75–85.
- [35] Justin T. Huang, Masha Krupenkin, David Rothschild, and Julia Lee Cunningham. 2023. The cost of anti-Asian racism during the COVID-19 pandemic. *Nature Human Behaviour* 7, 5 (2023), 682–695.
- [36] Libby Jenke. 2023. Affective polarization and misinformation belief. *Political Behavior* (2023), 1–60.
- [37] Soojong Kim and Jisu Kim. 2023. The information ecosystem of conspiracy theory: Examining the QAnon narrative on Facebook. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–24.
- [38] Yubo Kou, Xinning Gui, Yunan Chen, and Kathleen Pine. 2017. Conspiracy talk on social media: Collective sensemaking during a public health crisis. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–21.
- [39] S. E. Kreps and D. L. Kriner. 2020. Model uncertainty, political contestation, and public trust in science: Evidence from the COVID-19 pandemic. *Science Advances* 6, 43 (2020), eabd4563.
- [40] Sian Lee, Haeseung Seo, Dongwon Lee, and Aiping Xiong. 2023. Associative inference can increase people's susceptibility to misinformation. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 17. 530–541.
- [41] Stephan Lewandowsky. 2021. Climate change disinformation and how to combat it. *Annual Review of Public Health* 42 (2021), 1–21.
- [42] Yichuan Li, Bohan Jiang, Kai Shu, and Huan Liu. 2020. MM-COVID: A multilingual and multimodal data repository for combating COVID-19 disinformation. ArXiv.
- [43] Sahil Loomba, Alexandre de Figueiredo, Simon J. Piatek, Kristen de Graaf, and Heidi J. Larson. 2021. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nature Human Behaviour* 5, 3 (2021), 337–348.
- [44] Michael Luca, Elizaveta Pronkina, and Michelangelo Rossi. 2024. The evolution of discrimination in online markets: How the rise in anti-Asian bias affected Airbnb during the pandemic. *Marketing Science* (2024).
- [45] Alessandro Miani, Thomas Hills, and Adrian Bangerter. 2022. Interconnectedness and (in)coherence as a signature of conspiracy worldviews. *Science Advances* 8, 43 (2022), eabq3668.
- [46] J. D. Moffitt, Catherine King, and Kathleen M. Carley. 2021. Hunting conspiracy theories during the COVID-19 pandemic. *Social Media + Society* 7, 3 (2021), 20563051211043212.

[47] Ryan C Moore, Ross Dahlke, Peter L Forberg, and Jeffrey T Hancock. 2024. The private life of QAnon: A mixed methods investigation of Americans' exposure to QAnon content on the web. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW2 (2024), 1–34.

[48] Cindy Sing Bik Ngai, Rita Gill Singh, and Le Yao. 2022. Impact of COVID-19 vaccine misinformation on social media virality: Content analysis of message themes and writing strategies. *Journal of Medical Internet Research* 24, 7 (2022), e37806.

[49] Dimitar Nikolov, Alessandro Flammini, and Filippo Menczer. 2021. Right and left, partisanship predicts (asymmetric) vulnerability to misinformation. *Harvard Kennedy School Misinformation Review* (2021).

[50] Mathias Osmundsen, Alexander Bor, Peter Bjerregaard Vahlstrup, Anja Bechmann, and Michael Bang Petersen. 2021. Partisan polarization is the primary psychological motivation behind political fake news sharing on Twitter. *American Political Science Review* 115, 3 (2021), 999–1015.

[51] Orestis Papakyriakopoulos and Ellen Goodman. 2022. The impact of Twitter labels on misinformation spread and user engagement: Lessons from Trump's election tweets. In *Proceedings of the ACM Web Conference 2022*. 2541–2551.

[52] Antonis Papasavva, Max Aliapoulos, Cameron Ballard, Emiliano De Cristofaro, Gianluca Stringhini, Savvas Zannettou, and Jeremy Blackburn. 2022. The Gospel according to Q: Understanding the QAnon conspiracy from the perspective of canonical information. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 16. 735–746.

[53] Saumya Pareek, Niels van Berkel, Eduardo Velloso, and Jorge Goncalves. 2024. Effect of explanation conceptualisations on trust in AI-assisted credibility assessment. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW2 (2024), 1–31.

[54] Francesco Pierri, Brea L. Perry, Matthew R. DeVerna, Kai-Cheng Yang, Alessandro Flammini, Filippo Menczer, and John Bryden. 2022. Online misinformation is linked to early COVID-19 vaccination hesitancy and refusal. *Scientific Reports* 12, 1 (2022), 5966.

[55] Robert Plutchik. 1991. *The emotions*. University Press of America.

[56] Stephen Prochaska, Kayla Duskin, Zarine Kharazian, Carly Minow, Stephanie Blucker, Sylvie Venuto, Jevin D. West, and Kate Starbird. 2023. Mobilizing manufactured reality: How participatory disinformation shaped deep stories to catalyze action during the 2020 U.S. presidential election. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–39.

[57] Nicolas Pröllochs, Dominik Bär, and Stefan Feuerriegel. 2021. Emotions explain differences in the diffusion of true vs. false social media rumors. *Scientific Reports* 11, 1 (2021), 22721.

[58] Nicolas Pröllochs, Dominik Bär, and Stefan Feuerriegel. 2021. Emotions in online rumor diffusion. *EPJ Data Science* 10, 1 (2021), 51.

[59] Ashwin Rao, Fred Morstatter, and Kristina Lerman. 2022. Partisan asymmetries in exposure to misinformation. *Scientific Reports* 12, 1 (2022), 15671.

[60] Steve Rathje, Jay J. Van Bavel, and Sander van der Linden. 2021. Out-group animosity drives engagement on social media. *Proceedings of the National Academy of Sciences* 118, 26 (2021), e2024292118.

[61] Raquel Recuero, Felipe Bonow Soares, and Anatoliy Gruzd. 2020. Hyperpartisanship, disinformation and political conversations on Twitter: The Brazilian presidential election of 2018. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 14. 569–578.

[62] Claire E. Robertson, Nicolas Pröllochs, Kaoru Schwarzenegger, Philip Pärnamets, Jay J. Van Bavel, and Stefan Feuerriegel. 2023. Negativity drives online news consumption. *Nature Human Behaviour* 7, 5 (2023), 1–11.

[63] Ronald E. Robertson, Jon Green, Damian J. Ruck, Katherine Ognyanova, Christo Wilson, and David Lazer. 2023. Users choose to engage with more partisan news than they are exposed to on Google Search. *Nature* (2023), 1–7.

[64] Jon Roozenbeek, Claudia R Schneider, Sarah Dryhurst, John Kerr, Alexandra LJ Freeman, Gabriel Recchia, Anne Marthe Van Der Bles, and Sander Van Der Linden. 2020. Susceptibility to misinformation about COVID-19 around the world. *Royal Society open science* 7, 10 (2020), 201199.

[65] Matheus Schmitz, Goran Muric, and Keith Burghardt. 2023. Detecting anti-vaccine users on Twitter. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 17. 787–795.

[66] Paul Sebo and Sylvain de Lucia. 2024. Performance of machine translators in translating French medical research abstracts to English: A comparative study of DeepL, Google Translate, and CUBBIT. *PLoS ONE* 19, 2 (2024), e0297183.

[67] Chengcheng Shao, Giovanni Luca Ciampaglia, Onur Varol, Kai-Cheng Yang, Alessandro Flammini, and Filippo Menczer. 2018. The spread of low-credibility content by social bots. *Nature Communications* 9, 1 (2018), 1–9.

[68] Filipo Sharevski, Raniem Alsaadi, Peter Jachim, and Emma Pieroni. 2022. Misinformation warnings: Twitter's soft moderation effects on COVID-19 vaccine belief echoes. *Computers & Security* 114 (2022), 102577.

[69] Karishma Sharma, Emilio Ferrara, and Yan Liu. 2022. Characterizing online engagement with disinformation and conspiracies in the 2020 US presidential election. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 16. 908–919.

- [70] Karishma Sharma, Yizhou Zhang, and Yan Liu. 2022. COVID-19 vaccine misinformation campaigns and social media narratives. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 16. 920–931.
- [71] Kirill Solovev and Nicolas Pröllochs. 2022. Moral emotions shape the virality of COVID-19 misinformation on social media. In *Proceedings of the ACM Web Conference 2022*. 3706–3717.
- [72] Victor Suarez-Lledo and Javier Alvarez-Galvez. 2021. Prevalence of health misinformation on social media: Systematic review. *Journal of Medical Internet Research* 23, 1 (2021), e17187.
- [73] Breena R Taira, Vanessa Kreger, Aristides Orue, and Lisa C Diamond. 2021. A pragmatic assessment of Google Translate for emergency department instructions. *Journal of General Internal Medicine* 36, 11 (2021), 3361–3365.
- [74] Chuang Tang, Shaobo (Kevin) Li, Yi Ding, Ram D. Gopal, and Guanglei Zhang. 2024. Racial discrimination and anti-discrimination: The COVID-19 pandemic's impact on Chinese restaurants in North America. *Information Systems Research* 35, 3 (2024), 1274–1295.
- [75] Xian Teng, Yu-Ru Lin, Wen-Ting Chung, Ang Li, and Adriana Kovashka. 2022. Characterizing user susceptibility to COVID-19 misinformation on Twitter. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 16. 1005–1016.
- [76] Jan-Willem Van Prooijen, Karen M. Douglas, and Clara De Inocencio. 2018. Connecting the dots: Illusory pattern perception predicts belief in conspiracies and the supernatural. *European Journal of Social Psychology* 48, 3 (2018), 320–335.
- [77] N. Velásquez, R. Leahy, N. Johnson Restrepo, Y. Lupu, R. Sear, N. Gabriel, O. K. Jha, B. Goldberg, and N. F. Johnson. 2021. Online hate network spreads malicious COVID-19 content outside the control of individual social media platforms. *Scientific Reports* 11, 1 (2021), 11549.
- [78] Soroush Vosoughi, Deb Roy, and Sinan Aral. 2018. The spread of true and false news online. *Science* 359, 6380 (2018), 1146–1151.
- [79] Luping Wang and Jeffrey M. Rzeszotarski. 2023. Understanding motivational factors in social media news sharing decisions. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–30.
- [80] Xiangyu Wang, Min Zhang, Weiguo Fan, and Kang Zhao. 2022. Understanding the spread of COVID-19 misinformation on social media: The effects of topics and a political leader's nudge. *Journal of the Association for Information Science and Technology* 73, 5 (2022), 726–737.
- [81] Ying Wang, Hongwei Zhang, Wang Gao, and Cai Yang. 2023. Spillover effects from news to travel and leisure stocks during the COVID-19 pandemic: Evidence from the time and frequency domains. *Tourism Economics* 29, 2 (2023), 460–487.
- [82] Muheng Yan, Yu-Ru Lin, and Wen-Ting Chung. 2022. Are mutated misinformation more contagious? A case study of COVID-19 misinformation on Twitter. In *Proceedings of the 14th ACM Web Science Conference 2022*. 336–347.
- [83] Savvas Zannettou. 2021. "I Won the Election!": An empirical analysis of soft moderation interventions on Twitter. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 15. 865–876.
- [84] Liudmila Zavolokina, Kilian Sprenkamp, Zoya Katashinskaya, Daniel Gordon Jones, and Gerhard Schwabe. 2024. Think fast, think slow, think critical: Designing an automated propaganda detection tool. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–24.
- [85] Cheng Zhou, Kai Li, and Yanhong Lu. 2021. Linguistic characteristics and the dissemination of misinformation in social media: The moderating effect of information richness. *Information Processing & Management* 58, 6 (2021), 102679.

Supplementary Materials

for “From News Source Sharers to Post Viewers: How Topic Diversity and Conspiracy Theories Shape Engagement With Misinformation During a Health Crisis”

S1 Estimation Results for Main Analysis

The full estimation results for the analysis in the main paper are reported in Tables S1–S3. Specifically, Table S1 reports the estimation results for the topic diversity. Table S2 reports the estimation results for the engagement at news source sharer level. Table S3 reports the estimation results for the engagement at post viewer level. Additionally, Table S4 reports the estimation results for the subset analysis of post viewer engagement within false news posts.

Table S1. Estimation results for the topic diversity in news claims [Column (1)], the topic diversity in news posts [Column (2)], and the changes in topic diversity from news claims to associated posts [Column (3)]. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) News claim | (2) News post | (3) Claim → Post |
|-----------------------|---------------------|----------------------|----------------------|
| Falsehood | 0.388*** (0.036) | 0.751*** (0.024) | 0.659*** (0.022) |
| Conspiracy | 0.061 (0.089) | 0.136*** (0.015) | 0.185*** (0.019) |
| Words | 0.391*** (0.013) | 0.373*** (0.004) | 0.390*** (0.004) |
| Media | | 0.022*** (0.006) | 0.035*** (0.008) |
| Verified | | 0.037*** (0.008) | 0.004 (0.010) |
| AccountAge | | 0.003 (0.003) | 0.001 (0.003) |
| Followers | | -0.001 (0.003) | -0.019*** (0.004) |
| Followees | | -0.013*** (0.002) | -0.015*** (0.003) |
| Emotions | ✓ | ✓ | ✓ |
| Language | ✓ | ✓ | ✓ |
| MonthYear | ✓ | ✓ | ✓ |
| Intercept | -0.078 (0.110) | -0.653*** (0.039) | -0.718*** (0.047) |
| News-level REs | ✗ | ✓ | ✓ |
| <i>N</i> | 7273 | 70,904 | 70,904 |
| <i>R</i> ² | 0.179 | ✗ | ✗ |

Table S2. Estimation results for news lifetime [Column (1)] and post count [Column (2)]. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) News lifetime | (2) Post count |
|------------------------------|----------------------|----------------------|
| Falsehood | 1.239*** (0.046) | 2.230*** (0.083) |
| Conspiracy | 0.452* (0.189) | 0.493 (0.335) |
| Diversity | -0.002 (0.005) | -0.025 (0.026) |
| Falsehood \times Diversity | -0.075 (0.039) | 0.079 (0.085) |
| Words | -0.080*** (0.010) | -0.614*** (0.030) |
| Emotions | ✓ | ✓ |
| Language | ✓ | ✓ |
| MonthYear | ✓ | ✓ |
| Intercept | 2.491*** (0.282) | 2.060*** (0.198) |
| <i>N</i> | 7273 | 7273 |
| <i>R</i> ² | 0.481 | 0.188 |

Table S3. Estimation results for repost count [Column (1)], like count [Column (2)], and reply count [Column (3)]. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) Repost count | (2) Like count | (3) Reply count |
|------------------------------|----------------------|---------------------|----------------------|
| Falsehood | -0.256*** (0.063) | -0.155* (0.063) | 0.400*** (0.058) |
| Conspiracy | 0.342*** (0.065) | 0.373*** (0.059) | 0.365*** (0.056) |
| Diversity | 0.181*** (0.031) | 0.192*** (0.030) | 0.137*** (0.030) |
| Falsehood \times Diversity | -0.125*** (0.035) | -0.093** (0.033) | -0.091** (0.033) |
| Anger | 0.049*** (0.015) | 0.059*** (0.014) | 0.040** (0.014) |
| Disgust | -0.043** (0.016) | -0.020 (0.015) | -0.020 (0.016) |
| Fear | 0.035* (0.016) | 0.008 (0.015) | 0.038* (0.015) |
| Joy | 0.006 (0.016) | 0.030* (0.015) | 0.077*** (0.015) |
| Sadness | 0.016 (0.017) | -0.009 (0.015) | -0.002 (0.016) |
| Surprise | 0.093*** (0.015) | 0.120*** (0.015) | 0.121*** (0.014) |
| Words | 0.519*** (0.017) | 0.467*** (0.016) | 0.443*** (0.015) |
| Media | 0.788*** (0.028) | 0.596*** (0.027) | 0.432*** (0.027) |
| Verified | 3.037*** (0.035) | 3.071*** (0.034) | 2.367*** (0.033) |
| AccountAge | 0.091*** (0.012) | 0.084*** (0.011) | 0.041*** (0.012) |
| Followers | 0.175*** (0.015) | 0.180*** (0.014) | 0.217*** (0.013) |
| Followees | 0.546*** (0.024) | 0.383*** (0.022) | 0.207*** (0.015) |
| Language | ✓ | ✓ | ✓ |
| MonthYear | ✓ | ✓ | ✓ |
| Intercept | 0.129 (0.169) | 1.055*** (0.155) | -1.398*** (0.153) |
| News-level REs | ✓ | ✓ | ✓ |
| <i>N</i> | 70,904 | 70,904 | 70,904 |
| <i>R</i> ² | ✗ | ✗ | ✗ |

Table S4. Estimation results for repost count [Column (1)], like count [Column (2)], and reply count [Column (3)] in the subset of false news posts. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) Repost count | (2) Like count | (3) Reply count |
|-----------------------|---------------------|---------------------|---------------------|
| Conspiracy | 0.356*** (0.075) | 0.389*** (0.066) | 0.378*** (0.062) |
| Diversity | 0.073*** (0.021) | 0.134*** (0.019) | 0.053** (0.018) |
| Words | 0.502*** (0.025) | 0.416*** (0.022) | 0.454*** (0.021) |
| Media | 0.895*** (0.037) | 0.712*** (0.034) | 0.525*** (0.033) |
| Verified | 2.624*** (0.048) | 2.564*** (0.045) | 1.969*** (0.042) |
| AccountAge | 0.097*** (0.016) | 0.113*** (0.014) | 0.079*** (0.014) |
| Followers | 0.502*** (0.046) | 0.616*** (0.046) | 0.602*** (0.036) |
| Followees | 0.717*** (0.035) | 0.512*** (0.031) | 0.269*** (0.022) |
| Emotions | ✓ | ✓ | ✓ |
| Language | ✓ | ✓ | ✓ |
| MonthYear | ✓ | ✓ | ✓ |
| Intercept | 0.129 (0.215) | 1.290*** (0.190) | -0.451* (0.186) |
| News-level REs | ✓ | ✓ | ✓ |
| <i>N</i> | 52,271 | 52,271 | 52,271 |
| <i>R</i> ² | ✗ | ✗ | ✗ |

S2 Cross-Correlations & Variance Inflation Factors

We provide the cross-correlations to show how the independent variables relate to each other at the news claim and news post levels. Table S5 reports the cross-correlations among the variables in news claims. Table S6 reports the cross-correlations among the variables in news posts.

Additionally, we examine the Variance Inflation Factors (VIFs) among the independent variables to mitigate the concern about multicollinearity. The variance inflation factors are reported in Table S7. We find that the variance inflation factors for the independent variables are all close to one and well below the critical threshold of four, which indicates that multicollinearity is not an issue of our analysis.

Table S5. Correlations of variables in news claims

| | Falsehood | Conspiracy | Anger | Disgust | Fear | Joy | Sadness | Surprise | Words |
|------------|-----------|------------|--------|---------|--------|--------|---------|----------|-------|
| Falsehood | 1.000 | | | | | | | | |
| Conspiracy | 0.221 | 1.000 | | | | | | | |
| Anger | -0.010 | -0.018 | 1.000 | | | | | | |
| Disgust | 0.094 | 0.014 | 0.182 | 1.000 | | | | | |
| Fear | 0.030 | -0.018 | 0.159 | -0.027 | 1.000 | | | | |
| Joy | -0.007 | 0.026 | -0.120 | -0.118 | -0.216 | 1.000 | | | |
| Sadness | 0.082 | -0.005 | 0.094 | 0.108 | 0.283 | -0.185 | 1.000 | | |
| Surprise | -0.009 | -0.020 | 0.036 | 0.050 | -0.087 | -0.001 | 0.020 | 1.000 | |
| Words | -0.371 | -0.063 | 0.020 | -0.047 | 0.077 | 0.100 | -0.016 | 0.017 | 1.000 |

Table S6. Correlations of variables in news posts

| | Falsehood | Conspiracy | Anger | Disgust | Fear | Joy | Sadness | Surprise | Words |
|------------|-----------|------------|--------|---------|--------|--------|---------|----------|--------|
| Falsehood | 1.000 | | | | | | | | |
| Conspiracy | 0.151 | 1.000 | | | | | | | |
| Anger | 0.038 | 0.002 | 1.000 | | | | | | |
| Disgust | 0.122 | -0.019 | 0.139 | 1.000 | | | | | |
| Fear | 0.075 | -0.054 | 0.089 | -0.063 | 1.000 | | | | |
| Joy | -0.020 | 0.062 | -0.142 | -0.155 | -0.254 | 1.000 | | | |
| Sadness | 0.133 | -0.065 | 0.093 | 0.014 | 0.375 | -0.220 | 1.000 | | |
| Surprise | 0.051 | -0.007 | 0.021 | -0.060 | -0.124 | 0.057 | -0.015 | 1.000 | |
| Words | 0.155 | 0.016 | 0.060 | -0.017 | 0.046 | 0.033 | 0.055 | 0.057 | 1.000 |
| Media | -0.076 | -0.051 | -0.017 | -0.031 | 0.013 | 0.007 | 0.000 | -0.033 | 0.104 |
| Verified | -0.278 | -0.072 | -0.018 | -0.003 | -0.031 | 0.003 | -0.061 | -0.028 | 0.080 |
| AccountAge | -0.185 | -0.035 | -0.033 | -0.020 | -0.030 | -0.007 | -0.036 | -0.019 | -0.060 |
| Followers | -0.220 | -0.045 | 0.000 | -0.019 | -0.015 | 0.028 | -0.035 | -0.014 | 0.104 |
| Followees | -0.006 | -0.011 | -0.001 | 0.091 | -0.017 | -0.002 | -0.023 | -0.001 | -0.071 |

| Continued | Media | Verified | AccountAge | Followers | Followees |
|------------|-------|----------|------------|-----------|-----------|
| Falsehood | | | | | |
| Conspiracy | | | | | |
| Anger | | | | | |
| Disgust | | | | | |
| Fear | | | | | |
| Joy | | | | | |
| Sadness | | | | | |
| Surprise | | | | | |
| Words | | | | | |
| Media | 1.000 | | | | |
| Verified | 0.218 | 1.000 | | | |
| AccountAge | 0.047 | 0.321 | 1.000 | | |
| Followers | 0.116 | 0.429 | 0.227 | 1.000 | |
| Followees | 0.006 | 0.053 | 0.105 | 0.044 | 1.000 |

Table S7. VIFs

| | News | Post |
|------------|------|------|
| Falsehood | 1.24 | 1.22 |
| Conspiracy | 1.05 | 1.04 |
| Anger | 1.07 | 1.05 |
| Disgust | 1.07 | 1.09 |
| Fear | 1.18 | 1.25 |
| Joy | 1.10 | 1.14 |
| Sadness | 1.13 | 1.21 |
| Surprise | 1.01 | 1.03 |
| Words | 1.19 | 1.09 |
| Media | | 1.06 |
| Verified | | 1.41 |
| AccountAge | | 1.16 |
| Followers | | 1.27 |
| Followees | | 1.03 |
| Mean VIF | 1.12 | 1.15 |

S3 Robustness Check With Original 26 Topics

We adopt the original 26 topics and associated keywords directly from previous research [4] and repeat the analysis. Table S8 reports the estimation results for the topic diversity. Table S9 reports the estimation results for the engagement at news source sharer level. Table S10 reports the estimation results for the engagement at post viewer level. All the results are robust and consistently support our main findings.

Table S8. Estimation results for the topic diversity in news claims [Column (1)], the topic diversity in news posts [Column (2)], and the changes in topic diversity from news claims to associated posts [Column (3)]. The topic diversity is recalculated based on the original 26 topics. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) News claim | (2) News post | (3) Claim → Post |
|-----------------------|---------------------|----------------------|----------------------|
| Falsehood | 0.404*** (0.034) | 0.796*** (0.023) | 0.724*** (0.022) |
| Conspiracy | 0.030 (0.095) | 0.117*** (0.015) | 0.165*** (0.019) |
| Words | 0.417*** (0.013) | 0.393*** (0.003) | 0.412*** (0.004) |
| Media | | 0.008 (0.006) | 0.018* (0.008) |
| Verified | | 0.043*** (0.008) | 0.013 (0.010) |
| AccountAge | | 0.002 (0.003) | -0.001 (0.003) |
| Followers | | -0.001 (0.003) | -0.021*** (0.004) |
| Followees | | -0.015*** (0.002) | -0.018*** (0.003) |
| Emotions | ✓ | ✓ | ✓ |
| Language | ✓ | ✓ | ✓ |
| MonthYear | ✓ | ✓ | ✓ |
| Intercept | -0.078 (0.109) | -0.709*** (0.037) | -0.873*** (0.046) |
| News-level REs | ✗ | ✓ | ✓ |
| <i>N</i> | 7479 | 73,006 | 73,006 |
| <i>R</i> ² | 0.196 | ✗ | ✗ |

Table S9. Estimation results for news lifetime [Column (1)] and post count [Column (2)]. The topic diversity is recalculated based on the original 26 topics. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) News lifetime | (2) Post count |
|------------------------------|----------------------|----------------------|
| Falsehood | 1.232*** (0.045) | 2.220*** (0.082) |
| Conspiracy | 0.412* (0.186) | 0.486 (0.332) |
| Diversity | -0.002 (0.005) | -0.045 (0.026) |
| Falsehood \times Diversity | -0.069 (0.040) | 0.084 (0.081) |
| Words | -0.081*** (0.010) | -0.609*** (0.030) |
| Emotions | ✓ | ✓ |
| Language | ✓ | ✓ |
| MonthYear | ✓ | ✓ |
| Intercept | 2.534*** (0.274) | 2.062*** (0.197) |
| <i>N</i> | 7479 | 7479 |
| <i>R</i> ² | 0.478 | 0.185 |

Table S10. Estimation results for repost count [Column (1)], like count [Column (2)], and reply count [Column (3)]. The topic diversity is recalculated based on the original 26 topics. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) Repost count | (2) Like count | (3) Reply count |
|------------------------------|----------------------|---------------------|----------------------|
| Falsehood | -0.240*** (0.061) | -0.150* (0.061) | 0.407*** (0.057) |
| Conspiracy | 0.335*** (0.064) | 0.365*** (0.059) | 0.361*** (0.056) |
| Diversity | 0.160*** (0.030) | 0.167*** (0.029) | 0.118*** (0.029) |
| Falsehood \times Diversity | -0.048 (0.034) | -0.034 (0.032) | -0.042 (0.033) |
| Words | 0.517*** (0.017) | 0.467*** (0.016) | 0.442*** (0.015) |
| Media | 0.774*** (0.027) | 0.594*** (0.026) | 0.426*** (0.026) |
| Verified | 3.027*** (0.035) | 3.036*** (0.034) | 2.368*** (0.032) |
| AccountAge | 0.094*** (0.012) | 0.089*** (0.011) | 0.044*** (0.011) |
| Followers | 0.186*** (0.015) | 0.184*** (0.014) | 0.221*** (0.013) |
| Followees | 0.464*** (0.022) | 0.300*** (0.019) | 0.189*** (0.014) |
| Emotions | ✓ | ✓ | ✓ |
| Language | ✓ | ✓ | ✓ |
| MonthYear | ✓ | ✓ | ✓ |
| Intercept | 0.057 (0.165) | 1.020*** (0.151) | -1.447*** (0.150) |
| News-level REs | ✓ | ✓ | ✓ |
| <i>N</i> | 73,006 | 73,006 | 73,006 |
| <i>R</i> ² | ✗ | ✗ | ✗ |

S4 Robustness Check at Theme Level

We recalculate topic diversity at the theme level and repeat our analysis. Table S11 reports the estimation results for the topic diversity. Table S12 reports the estimation results for the engagement at news source sharer level. Table S13 reports the estimation results for the engagement at post viewer level. All the results are robust and consistently support our main findings.

Table S11. Estimation results for the topic diversity in news claims [Column (1)], the topic diversity in news posts [Column (2)], and the changes in topic diversity from news claims to associated posts [Column (3)]. The topic diversity is recalculated at the theme level. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) News claim | (2) News post | (3) Claim → Post |
|-----------------------|---------------------|----------------------|----------------------|
| Falsehood | 0.311*** (0.036) | 0.699*** (0.026) | 0.616*** (0.023) |
| Conspiracy | 0.033 (0.092) | 0.130*** (0.016) | 0.173*** (0.019) |
| Words | 0.354*** (0.013) | 0.337*** (0.004) | 0.342*** (0.004) |
| Media | | 0.030*** (0.007) | 0.040*** (0.008) |
| Verified | | 0.033*** (0.008) | -0.002 (0.010) |
| AccountAge | | 0.005* (0.003) | 0.004 (0.003) |
| Followers | | 0.002 (0.003) | -0.020*** (0.004) |
| Followees | | -0.014*** (0.002) | -0.015*** (0.003) |
| Emotions | ✓ | ✓ | ✓ |
| Language | ✓ | ✓ | ✓ |
| MonthYear | ✓ | ✓ | ✓ |
| Intercept | -0.101 (0.110) | -0.517*** (0.040) | -0.600*** (0.047) |
| News-level REs | ✗ | ✓ | ✓ |
| <i>N</i> | 7273 | 70,904 | 70,904 |
| <i>R</i> ² | 0.147 | ✗ | ✗ |

Table S12. Estimation results for news lifetime [Column (1)] and post count [Column (2)]. The topic diversity is recalculated at the theme level. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) News lifetime | (2) Post count |
|------------------------------|----------------------|----------------------|
| Falsehood | 1.235*** (0.046) | 2.238*** (0.084) |
| Conspiracy | 0.448* (0.188) | 0.485 (0.324) |
| Diversity | -0.005 (0.005) | -0.048 (0.026) |
| Falsehood \times Diversity | -0.043 (0.042) | 0.044 (0.081) |
| Words | -0.082*** (0.010) | -0.602*** (0.030) |
| Emotions | ✓ | ✓ |
| Language | ✓ | ✓ |
| MonthYear | ✓ | ✓ |
| Intercept | 2.490*** (0.282) | 2.076*** (0.199) |
| <i>N</i> | 7273 | 7273 |
| <i>R</i> ² | 0.480 | 0.188 |

Table S13. Estimation results for repost count [Column (1)], like count [Column (2)], and reply count [Column (3)]. The topic diversity is recalculated at the theme level. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) Repost count | (2) Like count | (3) Reply count |
|------------------------------|----------------------|---------------------|----------------------|
| Falsehood | -0.225*** (0.062) | -0.125* (0.062) | 0.415*** (0.057) |
| Conspiracy | 0.346*** (0.065) | 0.378*** (0.059) | 0.368*** (0.056) |
| Diversity | 0.159*** (0.030) | 0.175*** (0.028) | 0.125*** (0.028) |
| Falsehood \times Diversity | -0.124*** (0.034) | -0.094** (0.032) | -0.074* (0.031) |
| Words | 0.528*** (0.017) | 0.477*** (0.016) | 0.444*** (0.015) |
| Media | 0.786*** (0.028) | 0.596*** (0.027) | 0.431*** (0.027) |
| Verified | 3.036*** (0.035) | 3.070*** (0.034) | 2.366*** (0.033) |
| AccountAge | 0.091*** (0.012) | 0.084*** (0.011) | 0.041*** (0.012) |
| Followers | 0.175*** (0.015) | 0.180*** (0.014) | 0.217*** (0.013) |
| Followees | 0.546*** (0.024) | 0.383*** (0.022) | 0.207*** (0.015) |
| Emotions | ✓ | ✓ | ✓ |
| Language | ✓ | ✓ | ✓ |
| MonthYear | ✓ | ✓ | ✓ |
| Intercept | 0.098 (0.169) | 1.023*** (0.154) | -1.418*** (0.153) |
| News-level REs | ✓ | ✓ | ✓ |
| <i>N</i> | 70,904 | 70,904 | 70,904 |
| <i>R</i> ² | ✗ | ✗ | ✗ |

S5 Robustness Check With Propensity Score Matching

To reduce potential unobserved confounding factors and ensure the robustness of our results, we further conduct propensity score matching for true and false news claims. Specifically, we match true and false news claims based on the variables of emotions and the number of words, with settings including the caliper of 0.3, “common support”, and “noreplacement.” Consequently, we discard 30 false news claims and achieve all biases lower than 0.05. Subsequently, we repeat our analysis based on the matched true and false news claims and their associated posts on X. Table S14 reports the estimation results for the topic diversity. Table S15 reports the estimation results for the engagement at news source sharer level. Table S16 reports the estimation results for the engagement at post viewer level. All the results are robust and consistently support our main findings.

Table S14. Estimation results for the topic diversity in news claims [Column (1)], the topic diversity in news posts [Column (2)], and the changes in topic diversity from news claims to associated posts [Column (3)]. The estimations are based on samples obtained after propensity score matching. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) News claim | (2) News post | (3) Claim → Post |
|-----------------------|---------------------|----------------------|----------------------|
| Falsehood | 0.400*** (0.036) | 0.745*** (0.024) | 0.640*** (0.022) |
| Conspiracy | 0.037 (0.091) | 0.137*** (0.015) | 0.186*** (0.019) |
| Words | 0.387*** (0.013) | 0.369*** (0.004) | 0.383*** (0.004) |
| Media | | 0.021*** (0.006) | 0.034*** (0.008) |
| Verified | | 0.041*** (0.008) | 0.009 (0.010) |
| AccountAge | | 0.003 (0.003) | 0.001 (0.003) |
| Followers | | -0.001 (0.003) | -0.019*** (0.004) |
| Followees | | -0.010*** (0.002) | -0.011*** (0.003) |
| Emotions | ✓ | ✓ | ✓ |
| Language | ✓ | ✓ | ✓ |
| MonthYear | ✓ | ✓ | ✓ |
| Intercept | -0.075 (0.111) | -0.644*** (0.039) | -0.705*** (0.047) |
| News-level REs | ✗ | ✓ | ✓ |
| <i>N</i> | 7243 | 69,366 | 69,366 |
| <i>R</i> ² | 0.181 | ✗ | ✗ |

Table S15. Estimation results for news lifetime [Column (1)] and post count [Column (2)]. The estimations are based on samples obtained after propensity score matching. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) News lifetime | (2) Post count |
|------------------------------|----------------------|----------------------|
| Falsehood | 1.234*** (0.046) | 2.230*** (0.083) |
| Conspiracy | 0.479* (0.189) | 0.501 (0.337) |
| Diversity | 0.000 (0.005) | -0.024 (0.026) |
| Falsehood \times Diversity | -0.073 (0.040) | 0.077 (0.086) |
| Words | -0.082*** (0.010) | -0.615*** (0.030) |
| Emotions | ✓ | ✓ |
| Language | ✓ | ✓ |
| MonthYear | ✓ | ✓ |
| Intercept | 2.572*** (0.275) | 2.077*** (0.197) |
| <i>N</i> | 7243 | 7243 |
| <i>R</i> ² | 0.482 | 0.188 |

Table S16. Estimation results for repost count [Column (1)], like count [Column (2)], and reply count [Column (3)]. The estimations are based on samples obtained after propensity score matching. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) Repost count | (2) Like count | (3) Reply count |
|------------------------------|----------------------|---------------------|----------------------|
| Falsehood | -0.244*** (0.063) | -0.134* (0.063) | 0.411*** (0.058) |
| Conspiracy | 0.360*** (0.065) | 0.391*** (0.059) | 0.370*** (0.057) |
| Diversity | 0.183*** (0.031) | 0.193*** (0.030) | 0.133*** (0.030) |
| Falsehood \times Diversity | -0.125*** (0.035) | -0.088** (0.033) | -0.088** (0.033) |
| Words | 0.512*** (0.017) | 0.459*** (0.016) | 0.442*** (0.015) |
| Media | 0.788*** (0.028) | 0.597*** (0.027) | 0.432*** (0.027) |
| Verified | 3.075*** (0.036) | 3.118*** (0.035) | 2.405*** (0.033) |
| AccountAge | 0.088*** (0.012) | 0.081*** (0.011) | 0.038** (0.012) |
| Followers | 0.172*** (0.015) | 0.177*** (0.014) | 0.216*** (0.013) |
| Followees | 0.635*** (0.026) | 0.467*** (0.024) | 0.250*** (0.017) |
| Emotions | ✓ | ✓ | ✓ |
| Language | ✓ | ✓ | ✓ |
| MonthYear | ✓ | ✓ | ✓ |
| Intercept | 0.061 (0.169) | 0.974*** (0.155) | -1.475*** (0.154) |
| News-level REs | ✓ | ✓ | ✓ |
| <i>N</i> | 69,366 | 69,366 | 69,366 |
| <i>R</i> ² | ✗ | ✗ | ✗ |

S6 Robustness Check With Inclusion of QAnon Conspiracy Theory

We further include QAnon to extend conspiracy theories and repeat our analysis. QAnon is a far-right conspiracy theory alleging that a cabal of politicians and media personalities are engaged in a war to undermine society [3, 37, 52]. The QAnon discourse has become more diverse and evolved into a global phenomenon during the COVID-19 pandemic [34]. We adopt a list of QAnon-specific keywords from previous research [69] to identify QAnon conspiracy theory in our dataset. The performance of this keyword list in identifying QAnon discourse has been manually validated, showing high agreement with human assessments [3]. As a result, we identify 6 QAnon news claims and 302 QAnon news posts. Table S17 reports the estimation results for the topic diversity. Table S18 reports the estimation results for the engagement at news source sharer level. Table S19 reports the estimation results for the engagement at post viewer level. All the results are robust and consistently support our main findings.

Table S17. Estimation results for the topic diversity in news claims [Column (1)], the topic diversity in news posts [Column (2)], and the changes in topic diversity from news claims to associated posts [Column (3)]. QAnon conspiracy theory is included during estimation. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) News claim | (2) News post | (3) Claim → Post |
|-----------------------|---------------------|----------------------|----------------------|
| Falsehood | 0.389*** (0.036) | 0.751*** (0.024) | 0.659*** (0.022) |
| Conspiracy | 0.030 (0.087) | 0.124*** (0.015) | 0.175*** (0.018) |
| Words | 0.391*** (0.013) | 0.373*** (0.004) | 0.389*** (0.004) |
| Media | | 0.022** (0.006) | 0.034*** (0.008) |
| Verified | | 0.037*** (0.008) | 0.005 (0.010) |
| AccountAge | | 0.003 (0.003) | 0.001 (0.003) |
| Followers | | -0.001 (0.003) | -0.019*** (0.004) |
| Followees | | -0.013*** (0.002) | -0.015*** (0.003) |
| Emotions | ✓ | ✓ | ✓ |
| Language | ✓ | ✓ | ✓ |
| MonthYear | ✓ | ✓ | ✓ |
| Intercept | -0.075 (0.110) | -0.652*** (0.039) | -0.717*** (0.047) |
| News-level REs | ✗ | ✓ | ✓ |
| <i>N</i> | 7273 | 70,904 | 70,904 |
| <i>R</i> ² | 0.179 | ✗ | ✗ |

Table S18. Estimation results for news lifetime [Column (1)] and post count [Column (2)]. QAnon conspiracy theory is included during estimation. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) News lifetime | (2) Post count |
|------------------------------|----------------------|----------------------|
| Falsehood | 1.234*** (0.046) | 2.232*** (0.083) |
| Conspiracy | 0.505** (0.183) | 0.443 (0.330) |
| Diversity | -0.002 (0.005) | -0.025 (0.026) |
| Falsehood \times Diversity | -0.074 (0.039) | 0.080 (0.085) |
| Words | -0.080*** (0.010) | -0.613*** (0.030) |
| Emotions | ✓ | ✓ |
| Language | ✓ | ✓ |
| MonthYear | ✓ | ✓ |
| Intercept | 2.487*** (0.282) | 2.065*** (0.198) |
| <i>N</i> | 7273 | 7273 |
| <i>R</i> ² | 0.481 | 0.188 |

Table S19. Estimation results for repost count [Column (1)], like count [Column (2)], and reply count [Column (3)]. QAnon conspiracy theory is included during estimation. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) Repost count | (2) Like count | (3) Reply count |
|------------------------------|----------------------|---------------------|----------------------|
| Falsehood | -0.251*** (0.063) | -0.150* (0.063) | 0.403*** (0.058) |
| Conspiracy | 0.271*** (0.063) | 0.300*** (0.058) | 0.312*** (0.055) |
| Diversity | 0.181*** (0.031) | 0.192*** (0.030) | 0.137*** (0.030) |
| Falsehood \times Diversity | -0.125** (0.035) | -0.092** (0.033) | -0.090** (0.033) |
| Words | 0.520*** (0.017) | 0.468*** (0.016) | 0.443*** (0.015) |
| Media | 0.788*** (0.028) | 0.596*** (0.027) | 0.432*** (0.027) |
| Verified | 3.037*** (0.035) | 3.071*** (0.034) | 2.367*** (0.033) |
| AccountAge | 0.092*** (0.012) | 0.085*** (0.011) | 0.041*** (0.012) |
| Followers | 0.175*** (0.015) | 0.179*** (0.014) | 0.217*** (0.013) |
| Followees | 0.545*** (0.024) | 0.383*** (0.022) | 0.207*** (0.015) |
| Emotions | ✓ | ✓ | ✓ |
| Language | ✓ | ✓ | ✓ |
| MonthYear | ✓ | ✓ | ✓ |
| Intercept | 0.130 (0.169) | 1.055*** (0.155) | -1.396*** (0.153) |
| News-level REs | ✓ | ✓ | ✓ |
| <i>N</i> | 70,904 | 70,904 | 70,904 |
| <i>R</i> ² | ✗ | ✗ | ✗ |

S7 Robustness Check With Exclusion of Potential Debunking Posts

News posts can also debunk the linked news items. We use keywords of “misinformation,” “misleading,” “fake,” “false,” and “disinformation” to identify potential debunking posts, resulting in 4316 news posts. We exclude these potential debunking posts and repeat our analysis. Table S20 reports the estimation results for the topic diversity. Table S21 reports the estimation results for the engagement at news source sharer level. Table S22 reports the estimation results for the engagement at post viewer level. All the results are robust and consistently support our main findings.

Table S20. Estimation results for the topic diversity in news claims [Column (1)], the topic diversity in news posts [Column (2)], and the changes in topic diversity from news claims to associated posts [Column (3)]. The potential debunking posts are excluded during estimation. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) News claim | (2) News post | (3) Claim → Post |
|-----------------------|---------------------|----------------------|----------------------|
| Falsehood | 0.364*** (0.037) | 0.773*** (0.025) | 0.717*** (0.023) |
| Conspiracy | 0.065 (0.095) | 0.124*** (0.016) | 0.171*** (0.020) |
| Words | 0.381*** (0.013) | 0.370*** (0.004) | 0.384*** (0.004) |
| Media | | 0.021** (0.007) | 0.033*** (0.008) |
| Verified | | 0.049*** (0.009) | 0.018 (0.011) |
| AccountAge | | 0.002 (0.003) | 0.001 (0.003) |
| Followers | | -0.005 (0.003) | -0.026*** (0.004) |
| Followees | | -0.009*** (0.002) | -0.010** (0.003) |
| Emotions | ✓ | ✓ | ✓ |
| Language | ✓ | ✓ | ✓ |
| MonthYear | ✓ | ✓ | ✓ |
| Intercept | -0.104 (0.117) | -0.636*** (0.039) | -0.697*** (0.048) |
| News-level REs | ✗ | ✓ | ✓ |
| <i>N</i> | 7097 | 66,588 | 66,588 |
| <i>R</i> ² | 0.173 | ✗ | ✗ |

Table S21. Estimation results for news lifetime [Column (1)] and post count [Column (2)]. The potential debunking posts are excluded during estimation. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) News lifetime | (2) Post count |
|------------------------------|----------------------|----------------------|
| Falsehood | 1.307*** (0.049) | 2.234*** (0.087) |
| Conspiracy | 0.401* (0.203) | 0.463 (0.344) |
| Diversity | -0.004 (0.005) | -0.023 (0.026) |
| Falsehood \times Diversity | -0.065 (0.044) | 0.077 (0.093) |
| Words | -0.078*** (0.010) | -0.619*** (0.029) |
| Emotions | ✓ | ✓ |
| Language | ✓ | ✓ |
| MonthYear | ✓ | ✓ |
| Intercept | 2.599*** (0.288) | 2.117*** (0.198) |
| <i>N</i> | 7097 | 7097 |
| <i>R</i> ² | 0.496 | 0.191 |

Table S22. Estimation results for repost count [Column (1)], like count [Column (2)], and reply count [Column (3)]. The potential debunking posts are excluded during estimation. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) Repost count | (2) Like count | (3) Reply count |
|------------------------------|----------------------|---------------------|----------------------|
| Falsehood | -0.207** (0.065) | -0.050 (0.065) | 0.525*** (0.060) |
| Conspiracy | 0.295*** (0.069) | 0.320*** (0.063) | 0.365*** (0.059) |
| Diversity | 0.177*** (0.032) | 0.191*** (0.030) | 0.138*** (0.030) |
| Falsehood \times Diversity | -0.118*** (0.036) | -0.086* (0.034) | -0.101** (0.033) |
| Words | 0.527*** (0.017) | 0.471*** (0.016) | 0.454*** (0.015) |
| Media | 0.768*** (0.029) | 0.581*** (0.028) | 0.416*** (0.027) |
| Verified | 3.071*** (0.038) | 3.127*** (0.037) | 2.420*** (0.035) |
| AccountAge | 0.097*** (0.013) | 0.083*** (0.012) | 0.036** (0.012) |
| Followers | 0.171*** (0.015) | 0.174*** (0.014) | 0.208*** (0.013) |
| Followees | 0.604*** (0.025) | 0.440*** (0.023) | 0.238*** (0.016) |
| Emotions | ✓ | ✓ | ✓ |
| Language | ✓ | ✓ | ✓ |
| MonthYear | ✓ | ✓ | ✓ |
| Intercept | 0.160 (0.174) | 1.050*** (0.159) | -1.452*** (0.159) |
| News-level REs | ✓ | ✓ | ✓ |
| <i>N</i> | 66,588 | 66,588 | 66,588 |
| <i>R</i> ² | ✗ | ✗ | ✗ |

S8 Robustness Check With BERTopic

To ensure the generalization of our results across the topic modeling methods, we use BERTopic, a popular unsupervised topic modeling tool, as an alternative method to automatically cluster topics.⁹ The BERTopic model generates 1526 topic clusters for the whole news claims and posts. We use the topic clusters and their corresponding probabilities to calculate topic diversity again and repeat the analysis. Table S23 reports the estimation results for the topic diversity. Table S24 reports the estimation results for the engagement at news source sharer level. Table S25 reports the estimation results for the engagement at post viewer level. All the results are robust and consistently support our main findings.

Table S23. Estimation results for the topic diversity in news claims [Column (1)], the topic diversity in news posts [Column (2)], and the changes in topic diversity from news claims to associated posts [Column (3)]. The topic diversity is calculated based on topics generated by BERTopic. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) News claim | (2) News post | (3) Claim → Post |
|-----------------------|---------------------|----------------------|----------------------|
| Falsehood | -0.038 (0.042) | -0.042 (0.028) | 0.059** (0.021) |
| Conspiracy | -0.150 (0.127) | 0.046* (0.020) | 0.045* (0.018) |
| Words | 0.089*** (0.013) | 0.144*** (0.005) | 0.110*** (0.004) |
| Media | | 0.016 (0.009) | 0.017* (0.007) |
| Verified | | 0.016 (0.011) | -0.016 (0.009) |
| AccountAge | | 0.001 (0.003) | 0.000 (0.003) |
| Followers | | -0.009* (0.004) | -0.007* (0.004) |
| Followees | | -0.019*** (0.003) | -0.016*** (0.003) |
| Emotions | ✓ | ✓ | ✓ |
| Language | ✓ | ✓ | ✓ |
| MonthYear | ✓ | ✓ | ✓ |
| Intercept | -0.162 (0.160) | 0.222*** (0.051) | -0.042 (0.044) |
| News-level REs | ✗ | ✓ | ✓ |
| <i>N</i> | 7273 | 70,904 | 70,904 |
| <i>R</i> ² | 0.041 | ✗ | ✗ |

⁹<https://maartengr.github.io/BERTopic/index.html>

Table S24. Estimation results for news lifetime [Column (1)] and post count [Column (2)]. The topic diversity is calculated based on topics generated by BERTopic. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) News lifetime | (2) Post count |
|------------------------------|----------------------|----------------------|
| Falsehood | 1.209*** (0.044) | 2.175*** (0.084) |
| Conspiracy | 0.405* (0.186) | 0.553 (0.350) |
| Diversity | -0.026*** (0.005) | -0.178*** (0.027) |
| Falsehood \times Diversity | -0.302*** (0.034) | -0.117 (0.068) |
| Words | -0.086*** (0.010) | -0.596*** (0.029) |
| Emotions | ✓ | ✓ |
| Language | ✓ | ✓ |
| MonthYear | ✓ | ✓ |
| Intercept | 2.394*** (0.275) | 2.029*** (0.202) |
| <i>N</i> | 7273 | 7273 |
| <i>R</i> ² | 0.508 | 0.192 |

Table S25. Estimation results for repost count [Column (1)], like count [Column (2)], and reply count [Column (3)]. The topic diversity is calculated based on topics generated by BERTopic. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) Repost count | (2) Like count | (3) Reply count |
|------------------------------|---------------------|----------------------|----------------------|
| Falsehood | -0.134* (0.060) | 0.000 (0.060) | 0.521*** (0.056) |
| Conspiracy | 0.351*** (0.065) | 0.387*** (0.059) | 0.375*** (0.056) |
| Diversity | 0.132*** (0.024) | 0.197*** (0.023) | 0.211*** (0.024) |
| Falsehood \times Diversity | -0.058* (0.028) | -0.094*** (0.026) | -0.173*** (0.027) |
| Words | 0.537*** (0.016) | 0.492*** (0.015) | 0.450*** (0.014) |
| Media | 0.783*** (0.028) | 0.591*** (0.027) | 0.432*** (0.027) |
| Verified | 3.041*** (0.035) | 3.076*** (0.034) | 2.369*** (0.033) |
| AccountAge | 0.092*** (0.012) | 0.086*** (0.011) | 0.043*** (0.012) |
| Followers | 0.174*** (0.015) | 0.181*** (0.014) | 0.218*** (0.013) |
| Followees | 0.550*** (0.024) | 0.389*** (0.022) | 0.207*** (0.015) |
| Emotions | ✓ | ✓ | ✓ |
| Language | ✓ | ✓ | ✓ |
| MonthYear | ✓ | ✓ | ✓ |
| Intercept | 0.013 (0.168) | 0.919*** (0.154) | -1.506*** (0.152) |
| News-level REs | ✓ | ✓ | ✓ |
| <i>N</i> | 70,904 | 70,904 | 70,904 |
| <i>R</i> ² | ✗ | ✗ | ✗ |

S9 Additional Analysis on Quote Count

Table S26 reports the full estimation results for quote count.

Table S26. Estimation results for quote count. Reported are coefficient estimates with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

| | (1) |
|------------------------------|----------------------|
| | Quote count |
| Falsehood | 0.102 (0.074) |
| Conspiracy | 0.585*** (0.084) |
| Diversity | 0.186*** (0.039) |
| Falsehood \times Diversity | -0.150*** (0.045) |
| Words | 0.414*** (0.021) |
| Media | 0.958*** (0.037) |
| Verified | 3.106*** (0.044) |
| AccountAge | 0.104*** (0.018) |
| Followers | 0.160*** (0.014) |
| Followees | 0.294*** (0.022) |
| Emotions | ✓ |
| Language | ✓ |
| MonthYear | ✓ |
| Intercept | -1.981*** (0.218) |
| News-level REs | ✓ |
| <i>N</i> | 70,904 |
| <i>R</i> ² | X |

Received July 2024; revised December 2024; accepted March 2025