

Misinformation during the COVID-19 outbreak in China: cultural, social and political entanglements

Yan Leng^{*^x}, Yujia Zhai^{†^x}, Shaojing Sun[‡], Yifei Wu[§], Jordan Selzer[¶], Sharon Strover^{||}, Hezhao Zhang^{**}, Anfan Chen^{††}, Ying Ding^{‡‡}

^{*}McCombs school of business, University of Texas at Austin
yan.leng@mccombs.utexas.edu

[†]School of Management, Tianjin Normal University
School of Information Management, Wuhan University
yjzhai@tjnu.edu.cn

[‡]School of Journalism, Fudan University
ssun@fudan.edu.cn

[§]Department of Computer Science and Technology, Tsinghua University
wu-yf17@mails.tsinghua.edu.cn

[¶]Department of Emergency Medicine, George Washington University
jordan.selzer@gmail.com

^{||}School of Journalism, University of Texas at Austin
sharon.strover@austin.utexas.edu

^{**}School of Management, Tianjin Normal University
zhanghezhao794@gmail.com

^{††}University of Science and Technology of China
caf16@ustc.edu.cn

^{‡‡}School of Information, University of Texas at Austin
ying.ding@ischool.utexas.edu

^x Equal contribution

Abstract—Not only did COVID-19 give rise to a global pandemic, but also it resulted in an infodemic comprising misinformation, rumor, and propaganda. The consequences of this infodemic can erode public trust, impede the containment of the virus, and outlive the pandemic itself. The evolving and fragmented media landscape, particularly the extensive use of social media, is a crucial driver of the spread of misinformation. Focusing on the Chinese social media Weibo, we collected four million tweets, from December 9, 2019, to April 4, 2020, examining misinformation identified by the fact-checking platform Tencent—a leading Chinese tech giant. Our results show that the evolution of misinformation follows an issue-attention cycle pertaining to topics such as city lockdown, cures and preventive measures, school reopening, and foreign countries. Sensational and emotionally reassuring misinformation characterizes the whole issue-attention cycle, with misinformation on cures and prevention flooding social media. We also study the evolution of sentiment and observe that positive sentiment dominated over the course of Covid, which may be due to the unique characteristic of "positive energy" on Chinese social media. Lastly, we study the media landscape during Covid via a case study on a controversial unproven cure known as Shuanghuanglian, which testifies to the importance of scientific communication in a plague. Our findings shed light on the

distinct characteristics of misinformation and its cultural, social, and political implications, during the COVID-19 pandemic. The study also offers insights into combating misinformation in China and across the world at large.

Index Terms—Misinformation; Covid-19; Social media; Evolution; Scientific communication.

1 INTRODUCTION

While the world is struggling to fight the COVID-19 pandemic, another more insidious epidemic—an infodemic, the term coined by WHO Director-General Tedros Adhanom Ghebreyesus to describe the prevalence of misinformation on the coronavirus—is threatening the health and safety of the public. Misinformation has existed throughout history; however, with the rise of social media platforms—Twitter, Facebook, Reddit, Weibo, microblogs, among others—its spread has transcended borders and picked up its speed [1].

Understanding the flow and spread of misinformation is a crucial prerequisite to mitigating its negative impacts. Just as COVID-19 has now spread to millions worldwide, misinformation, ranging from pseudoscience to conspiracy, also has spread online at an alarming rate, stirring panic and causing confusion. In China, such panic and confusion have led to massive scrambling for fictitious "cures," such as garlic and Shuanghuanglian, a traditional Chinese medicine.

As a complex and elusive concept, misinformation is hard to define rigorously despite the growing volume and scope of research on this topic. Briefly, it refers to "information considered incorrect based on the best available evidence from relevant experts at the time" [2]. In the past decade, the explosion of social media has significantly increased dissemination of misinformation pertaining to numerous topics, including health, politics, and entertainment. As a result, identifying and combating health-related misinformation on social media has become a top concern of public health [3]. During the 2018-19 Ebola outbreak, more than 25 percent of people in DR Congo believed misinformation about the outbreak, which significantly impeded public adoption of preventive behavior [4]. Factors affecting one's response to misinformation are complex and varied. Social norms, as well as the compatibility between misinformation and one's belief system, have a significant influence on one's acceptance of misinformation [5].

To date, the majority of research on misinformation has focused on Western media platforms; however, Chinese social media constitutes a very significant market. Weibo hosts more than 556 million registered users, with 313 million being active monthly, compared to 319 million on Twitter [6]. There are some differences when comparing Weibo and Twitter, however. Past research has found that Weibo demonstrates a different pattern of information dissemination than Twitter, though the two platforms are often viewed as comparable [7]. Twitter is more open to contestation and plural views, whereas Weibo is under stricter control and draconian censorship by the Chinese government [8]. Besides, compared to Twitter, the network connections on Weibo are more hierarchical because users tend to follow others at a similar or higher social level [6]. Considering the consequential impact of misinformation, it is vital to investigate the spread of misinformation in the ongoing COVID-19 pandemic. China—the first country undergoing the massive infection and lockdown—provides a rich cultural context in which to investigate COVID-19 related misinformation.

2 DATA

We collected Weibo posts identified as misinformation or gossip from Tencent, the Jiao Zhen Platform¹, from January 18, 2020 to March 29, 2020. There were 531 COVID-related posts identified as misinformation or gossip in total. The Jiao Zhen Platform is a crowdsourcing platform where users can dispel posts on Weibo as rumor or identify as inaccurate. The platform requests identification of credentials, such as education and affiliated institution, as well as provision of evidence such as citing the original source of information or

1. This platform can be accessed through <https://vp.fact.qq.com>.

a research article. The platform employs a full staff to verify the information once a record has been uploaded².

We collected 4.047 million Weibo posts from December 7, 2019, to April 4, 2020, by using the advanced search function of the Weibo keyword indexing. Our search used a series of keywords such as COVID-19, novel coronavirus, corona, epidemics, novel pneumonia, pneumonia in Wuhan. The data were crawled in two phases, with the first crawler collecting Weibo posts from January 18, 2020 to February 26, 2020, and the second crawler collecting posts from December 7, 2020 to April 4, 2020. We demonstrate the validity of our data collection in Appendix A.

The crawler program automatically entered one of the keywords into the query box and set the query time range to be a specific hour. For each query, the time range increased by one hour and all the new posts within that hour were crawled. Each query returned a maximum of 50 pages, with each page containing about 20 posts. As such, our data do not include those posts outside of the limits of 50 pages for each query. We have made the Weibo data publicly available, which can be accessed through <https://github.com/yleng/COVID-Weibo>.

3 EVOLUTION OF MISINFORMATION

Using the data collected from the Jiao Zhen platform, we examine how misinformation (rumors and pseudoscience) and gossip (inconclusive information) evolved throughout different stages of the COVID-19 epidemic in China, and how its dissemination correlates with significant news events and government policies. It should be noted that the lockdown was around the time of the Spring Festival—the most widely celebrated holiday in China. During this time, local residents and migrant workers return home for family reunions, leading to the most significant human migration on the planet. Due to the large-scale migration, it is also a time when the government is on high alert to safeguarding social order. The timing of the lunar-year festival may have amplified the abruptness of the lockdown measure, arousing more concern and anxiety about the severity of infection. With this timeline in mind, we propose the following exploratory research question:

RQ 1: How did different topics evolve during the course of COVID-19 in China?

3.1 Event Driven Misinformation

In Figure 1a, we present the temporal trends of misinformation and gossip that illuminate three waves of misinformation topics: city lockdown, virus transmission and cure, and school reopening. The first and most prominent peak emerged shortly after the lockdown was announced on January 23, 2020 in Wuhan. In response to this unprecedented policy, there was a surge of sensational information on severe disruptions in Wuhan (e.g., gas stations within Wuhan had stopped operating) and similar policies in other cities in China (e.g., a lockdown in Xinyang in Henan

2. More information regarding the platform can be accessed through the following <https://cloud.tencent.com/developer/news/17550> in Chinese.

province). Prompting the second wave, on February 2, Nanshan Zhong, a renowned pulmonologist and advisor to the Chinese government, declared that COVID-19 might spread via fecal transmission [9]. Following this announcement a spike of misinformation occurred regarding possible modes of transmission, cures, and prevention methods. For example, on social media appeared a few false claims suggesting that the virus may spread through mosquitoes, flies, and the sewer. Finally, a third peak of misinformation occurred around March 9, when the epidemic transitioned into a containment stage. During this stage, the focus of misinformation shifted toward lifting the lockdowns, restarting schools, and COVID-19 events in foreign countries (e.g., Italy and the U.S.).

To explore evolving topics, we present several word clouds by month in Figure 1b– 1d. Figure 1b covers the topics in January. The five most salient words are *prevention*, *disinfection*, *hospital*, *lockdown*, *treatment*. An interesting phenomenon during this period is misinformation that twists the original lockdown in Wuhan into lockdowns of other cities and other fabricated policies in Wuhan.

In February (as in Figure 1c), *prevention*, *disinfection*, *treatment* remained popular themes in misinformation. Meanwhile, new themes-related to patients, the U.S., and celebrities (e.g., Nanshan Zhong)–emerged. As the epidemic worsened and the testing capacity increased, the misinformation became more sensationalist in nature, for example, an outbreak that was happening in a particular city, or a confirmed patient who infected a large number of people. In the latter half of February, the U.S. attracted increasing media attention in China. Most of the misinformation was related to the confirmed cases and an upcoming outbreak in the U.S. At the same time, there were false claims and statements linked to a real statement by Nanshan Zhong [10]. Linking a credible/authoritative source to misinformation tends to increase the credibility of the untruth. For example, a piece of misinformation falsely stated that "Nanshan Zhong's team found that the infection rate of COVID is higher among smokers than non-smokers." Such a maneuver-exploiting a reputable source to establish legitimacy for a statement–may undermine the credibility of the actual experts and authoritative institutions.

In March, when the epidemic transitioned into a containment stage, *resume school*, *isolation*, *Italy*, *China*, *US* became the focus, as shown in Figure 1d. A noticeable shift in the focus is apparent compared with that in January and February, as misinformation about prevention, disinfection, and treatment became less frequent. Interestingly, there was a growing amount of misinformation on school-reopening and Italy.

We further inspected the temporal popularity of the top nine keywords in Figure 2. There are three types of temporal patterns: The first row, containing *isolation*, *treatment* and *prevention*, presents a lasting-popularity pattern; The second row exhibits a sudden-spike pattern, with an explosion of tweets related to patients in mid-January, followed by the spike of lockdown information; The third row depicts a slow-growth pattern, with certain topics (e.g., Italy, school-reopening) gradually becoming salient. This is also shown in Figure 1d.

3.2 Collective sensemaking and confirmation bias on cures and prevention

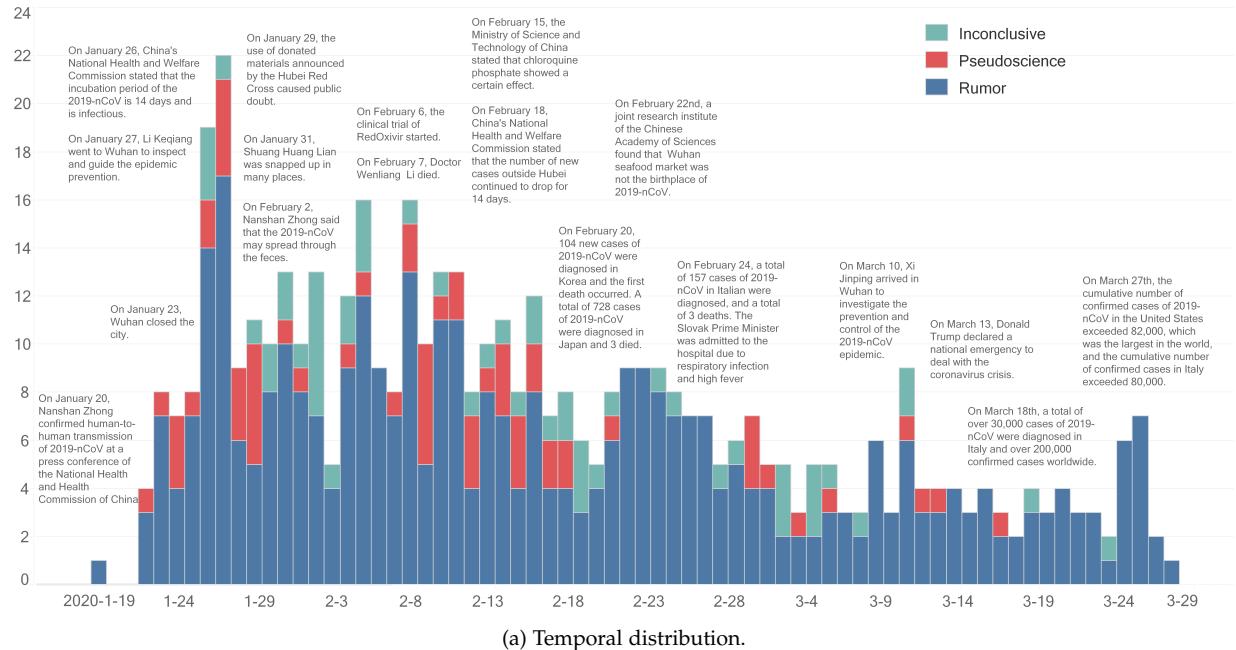
During a pandemic time rife with uncertainty and anxiety, people are more likely to encounter and subsequently become harmed from misinformation [11]. Confirmation bias may occur in such situations, referring to the tendency to fall back on what one already believes (or disbelieves) as one integrates new or unfamiliar information. As a major threat to public health, COVID-19 has triggered a surge of misinformation on treatment and prevention. Some of the misinformation may result in direct physical harm. Consider, for instance, the erroneous claims that firecrackers and drinking boiling water can cure the virus. In view of the multitude of misinformation, we ask the following questions:

RQ 2: Is the amount of distinct misinformation about COVID specifically also proportional to the number of posts about related topics (on cures and preventions) on Weibo?

RQ 3: Which particular cures and prevention measures attract most public attention on Weibo?

To answer RQ2, we first examine the correlation between Weibo posts and misinformation on Covid cures and prevention related topics (e.g., home remedies, Chinese medicines) by week. The Pearson's correlation coefficient is 0.663 ($p\text{-value} = 1e^{-9}$), suggesting a strong correlation between the two. We further analyze the evolution of topics on cures and prevention. In Figure 3-5, we examine the evolution of misinformation on cures and prevention. We observe a large number of rumors regarding home remedies at the beginning of the COVID-19 outbreak. This suggests that malicious accounts tailored misinformation to an audience demand that spiralled upward initially; with less incentive later, due to the reduction in consumption when the epidemic moved under control, malicious account presence diminished. Among the posts on Weibo about cures and prevention, home remedies attracted the most extensive attention, followed by disinfection supplies and traditional Chinese medicine, as shown on the upper panel of Figure ???. Specifically, the interests in home remedies took up more than 36.4% among all posts on misinformation throughout the observational period. Figure 5 presents the top 10 cures and prevention methods mentioned, based on the number of related posts in the Weibo. There is a constant interest in alcohol, tea, salt, disinfection supplies, and ginger. We also observed a significant spike in Shuanghuanglian.

A variety of factors drive the popularity of misinformation on home remedies in this pandemic context. First, the public's high familiarity with home remedies, connecting closely with day-to-day life experience, could result in intuitive thinking as compared to critical or cognitive thinking [12]. Many of these home remedies, rooted in Chinese culture, have been on the market for a long time (e.g., vinegar, alcohol, and garlic). Laypeople tend to rely on past experience and unconfirmed knowledge on home remedies to seek potential solutions. On a societal level, such a tendency can be viewed as collective sensemaking to resolve the collective anxiety caused by the ongoing health crisis [13]. Moreover, because of unwitting encounters in the past, people are often likely to trust and endorse the widespread, though unproven, use of home remedies [14],



(a) Temporal distribution.



Fig. 1: **Temporal variations of misinformation.** (a) Count of misinformation. The x-axis corresponds to the date and the y-axis is the number of posts for different types of misinformation, with the corresponding category shown in the legend. We label major events (e.g., policies or scientific news) with vertical lines. (b)-(d) Popular topics of misinformation. The size of each word is proportional to its frequency of occurrences in the titles of misinformation. For better demonstration, we removed some high-frequency words that were popular throughout the whole period, including novel coronavirus, infection, facial masks, Wuhan, epidemics.

evidencing why some untested home remedies become a constant focus of misinformation [15]. Second, pandemic profiteering and price gouging involving face-masks and fictitious cures emerge instantaneously. Some unscrupulous actors take advantage of the pandemic to sell their home remedies by making false or unsubstantiated claims. For example, some people attempted to gain profits over social media by selling mooncakes using Shuanghuanglian as an ingredient. Additionally, as home remedies become popular, some Weibo users tried to attract more followers by piggy-backing on the topic in order to benefit from an attention-driven economy [16].

4 SENSATIONALISM AND SENTIMENT EVOLUTION

Throughout COVID-19, one prominent characteristic of misinformation is sensationalism and scaremongering. Sources of misinformation manipulate true stories, forge ridiculous news, and use unverified sources to appeal to the crowds emotionally. Sensationalism in the media has long been a popular topic of fierce debate. Despite the absence of an authoritative definition of media sensationalism, researchers tend to agree that sensational messages are likely to provoke

senses and emotions. Furthermore, both content features (e.g., a topic about crime typically is more sensational than other topics) and formal elements (e.g., using a colorful font is more sensational than plain font) have proven to contribute to sensationalism [17]. From a linguistic angle, Molek-Kozakowska (2013) identified a range of pragmatic devices that may arouse audiences' sensational feelings [18]. Dubbed as sensational illocutions, the methods include exposing (e.g., denouncing a crime), speculating (suggesting future consequences), generalizing (extrapolating to a larger population), warning (generating anxiety) and extolling (extraordinary facts/events/individuals).

Integrating the perspectives mentioned above on content, form, and pragmatics about sensationalism, the present study incorporates those factors into a content analysis following the typical procedure. The authors classified the misinformation into five categories, with a final intercoder agreement of 0.94. In Figure 6, the categories of misinformation comprise: exposing the virus (e.g., the origin of the virus); extraordinary incidences (e.g., infected celebrity); warning of risk severity (e.g., relapse); conflict/conspiracy (e.g., biological warfare); and speculating about peoples'

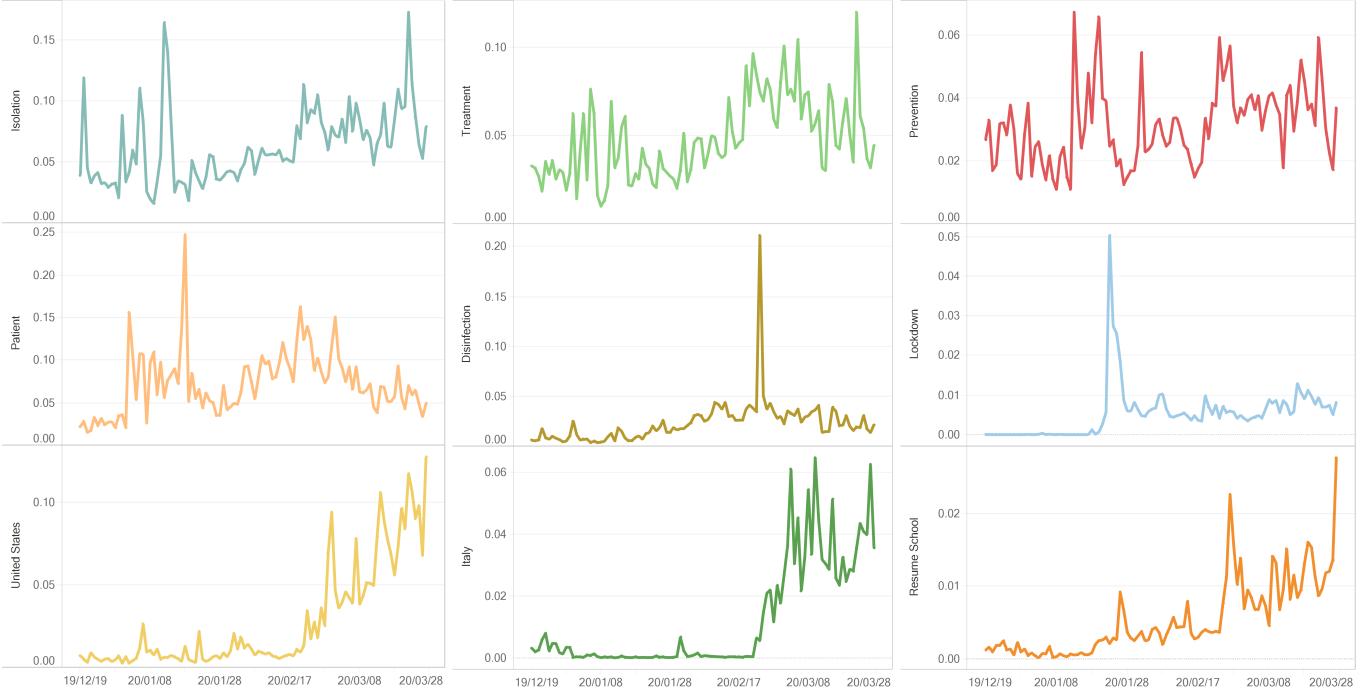
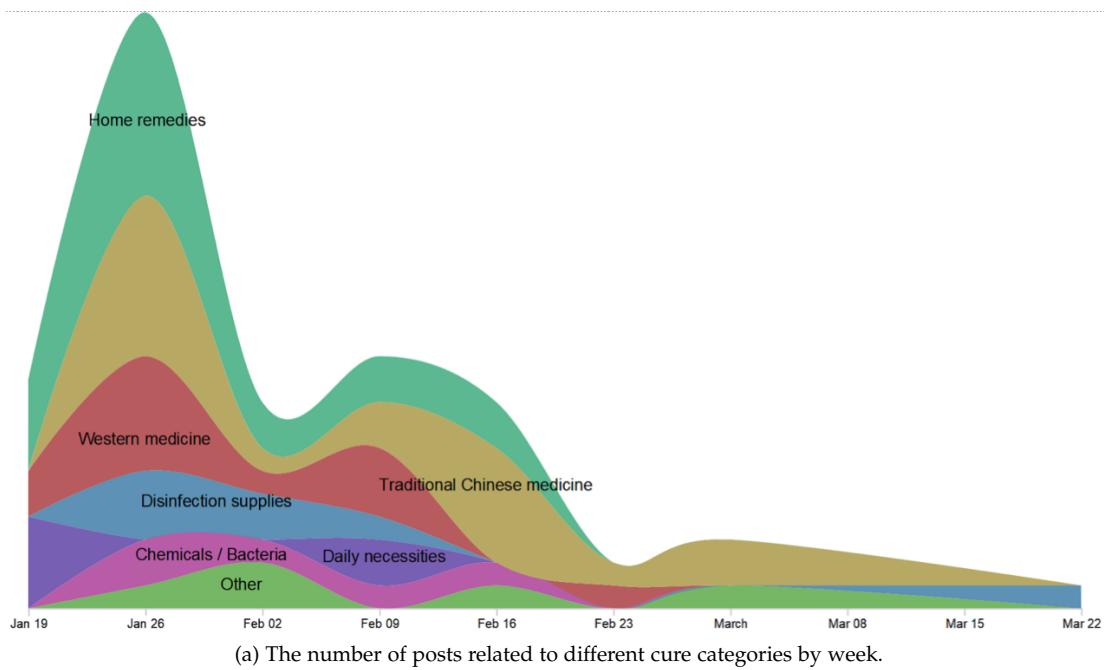


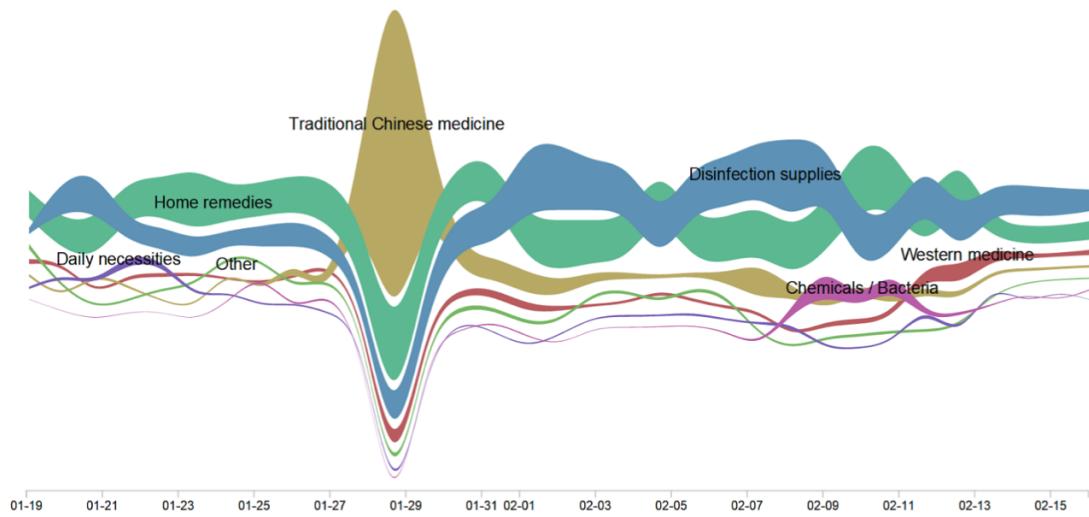
Fig. 2: Temporal popularity of the top keywords in the word clouds on Weibo. The keywords include, isolation, treatment, prevention, patient, disfection, lockdowns, United States, Italy and resume schools. The x-axis corresponds to the date and the y-axis corresponds to the percentage of weibo related to the topics of interest in our data.



Fig. 3: Evolution of misinformation on cures and prevention on Jiaozhen platform and related posts on Weibo. We categorize different types of cures and prevention into food, Chinese medicine, Western medicine, disinfection suppliers, daily necessities, and chemicals/germs. In the early stage of the epidemic, there was a large amount of misinformation mentioning home remedies. As time progressed, there was a shift towards Chinese medicine, especially in late January to early February, with Shuanghuanglian being a representative example. The patriotism and cultural preservation advanced by the Chinese government amplified the ideas about the effectiveness of Chinese herb medicine for treating COVID-19. Throughout the epidemic, one continues encountering misinformation regarding Western medicine, both conventional medicines (e.g., aspirin and mucosolvan) and recently approved or studied medicine (e.g., oseltamivir and resveratrol). Interestingly, as the epidemic in China moved into the containment stage in mid-March, misinformation about cures and prevention almost disappeared completely.



(a) The number of posts related to different cure categories by week.



(b) The number of misinformation about different cure categories by day.

Fig. 4: Evolution of posts and misinformation regarding cures and prevention for COVID-19.

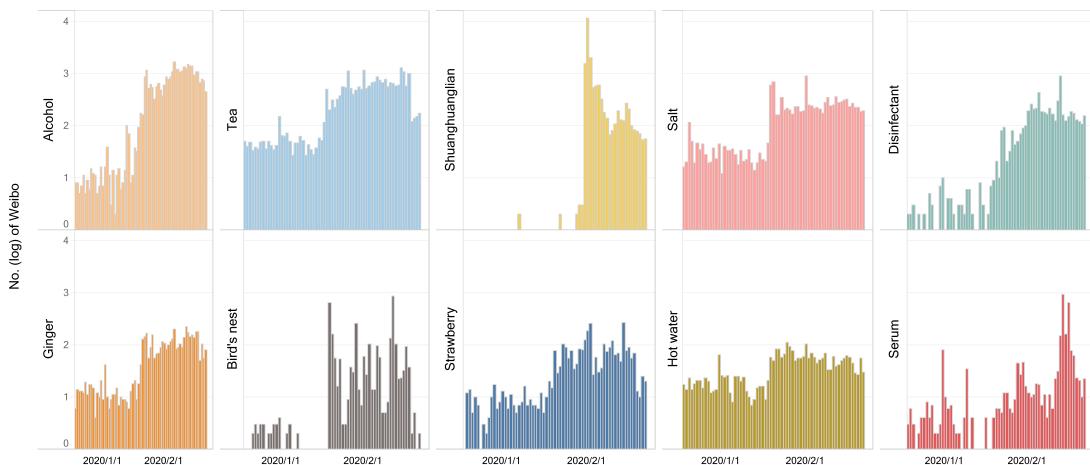


Fig. 5: Temporal popularity of the top 10 posts on cures on Weibo.

hard life (e.g., lockdowns). Take the category of exposing viruses as an example: several posts mentioned unknown virus sources, stories of diagnosed patients fleeing quarantine, or claims that one patient infected a large number of others. Anxiety and emotions also may be entangled with false claims of lockdowns of different cities, expressways, and even at gas stations.

We further analyze sentiment over the course of COVID-19 in China in Figure 7. We describe the method in obtaining the sentiment labels of each post in the Appendix B. We ask three questions:

RQ 4: Which sentiment dominates during COVID-19 on Chinese social media: negative sentiment or positive sentiment?

RQ 5: How does sentiment evolve across the time of the pandemic?

RQ 6: Does heterogeneity exist in sentiments over different topics?

To answer RQ4, we first use a one-sided t-test with the null hypothesis that positive posts dominate negative posts. We perform such tests for each week of the observation periods and we see that the hypothesis is rejected ($p\text{-val} < 0.05$) for twelve out of fifteen weeks (details to be found in Appendix C). We examine RQ4 by investigating the aggregated positive and negative sentiment in Figure 7a, and observe that there are much more positive posts than negative ones. This finding is somewhat surprising, but it demonstrates one characteristic of China's social media. Social media user tends to express "positive energy" on social media [19]. On August 2012, the term "positive energy" begin to appear in People's Daily, the China Communist Party's political mouthpiece, and then quickly spread out at the grassroots level. The rhetorical word is meant to disseminate hope and reduce negativity in the face of challenges and suffering [20]. This popularity of "positive energy" is particularly salient during the most uncertain and stressful time of the pandemic.

Figure 7b presents the evolution of the subsentiments in response to RQ 5. Anger increases in late February, mainly related to the source of the coronavirus and an article calling "Chinese virus" by American media and some news about whistle blowers in China. Fear, as another major emotion, presents two smooth peaks around January 19 at the beginning of the plague and in mid- to late February. In addition, there seems to be a decreasing trend with the emotion-happiness. We specifically study how news events generate anger sentiments. As an illustration, we chose three events related to President Donald Trump in the US. We compare the anger sentiment before and after the news events using a t-test in Table 1. We observe a stronger variation in sentiments in the early stages of the pandemics in China. More analysis about the change in sentiment in consecutive weeks can be found in Appendix D.

Figure 8 demonstrate the existence of heterogeneity in sentiments on different topics, in response to RQ 6. We manually code the COVID-related posts to eight categories, including information disclosure, donation support, everyday life, prosocial behavior, protective measures, negligence, medical treatment and epidemic prevention. We find that the sentiments around information disclosure if negative after mid-February, and the aggregate sentiment on neg-

ligence were negative before mid-January. Comparatively, the sentiments for donation support and epidemics control were overall positive, and the intensity increased after mid-January, and persisted until the end of the observational period.

5 MEDIA LANDSCAPE AND SCIENTIFIC COMMUNICATIONS

We conclude our analysis by commenting on the social media landscape during Covid. In order to mitigate harm to the public, Director-General Dr. Tedros Adhanom Ghebreyesus stated, "It is essential to fight infodemics with information" [21]. However, information itself is not enough to accomplish this goal. Due to a lack of comprehensive scientific training of the public and even some news media, misunderstandings or inaccurate interpretations of scientific studies may intensify fear and panic. The research question we ask is, RQ 7: What's the role of mainstream media and scientific accounts during this health crisis in China?

We answer this question using a case study on Shuanghuanglian (SHL). Specifically, we investigate the claim that SHL can inhibit Coronavirus. On January 31, 2020, the Shanghai Institute of Materia Medica (SIMM, under the Chinese Academy of Sciences) and the Wuhan Institute of Virology had discovered that the SHL herbal remedy could "inhibit" 2019-nCov [22]. The study was launched by laboratory in vitro studies and required further clinical studies to confirm its effectiveness on humans. However, this finding was commonly misinterpreted as 'SHL helps to prevent or cure coronavirus.' Misinterpreting the result in an environment of significant uncertainty contributed to a wave of panic-buying of SHL. The diffusion of the SHL topic is shown in Figure 9, which is especially interesting due to the polarized news and discussions, as revealed by the tags from the posts in Figure 10. We also present the Figure 11 to show the number of users, pro and against SHL, who engaged in SHL events by day. We describe the method to detect the pro and against Weibo posts in Appendix E. The upper panel presents the number of users engaged in SHL events, and the lower panel presents the number of pro and against-SHL posts. Specifically, the discussion revolves around the effectiveness and ineffectiveness of SHL.

Individuals are more likely to believe information with a logical narrative from a source they perceive to be credible [23]. Source credibility varies, and may even be related to what sort of media people use. A post that SHL was effective for *inhibition* of COVID-19 was endorsed by two credible news media outlets, both providing an abbreviated scientific explanation. However, due to a lack of understanding of science and the critical omission of a clear explanation by the media, many understood the report as claiming SHL to be an effective *treatment* for COVID-19. This highlights the importance of the media to be accurate and objective about scientific facts that are presented in scientific studies [24]. Otherwise, the media article may cause a misinterpretation by the public, which may deviate from the science. A similar phenomenon was observed in the U.S. with hydroxychloroquine, causing a shortage of the medication for lupus and rheumatoid arthritis patients, after claims of its (unproven)

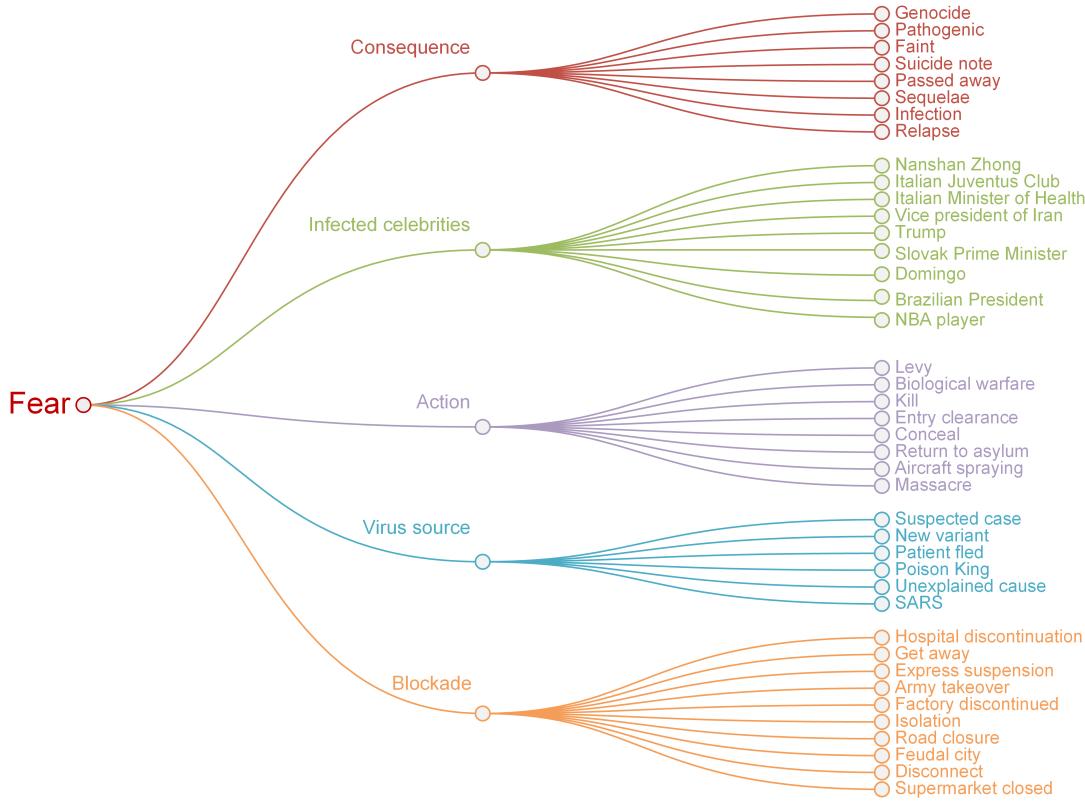


Fig. 6: **Summary of sensationalist misinformation.** We present a categorization over misinformation that may cause fear.

TABLE 1: Anger sentiment before and after the news events.

	Suspension of entry	Covid will disappear in April	Covid is going to disappear like a miracle
Mean (before and after)	0.0050 & 0.0068	0.0048 & 0.0075	0.0051 & 0.0082
Observation (before and after)	44 & 56	54 & 46	71 & 29
t-statistics (one-tail)	1.6628	1.6702	1.6883
p-val (one-tail)	0.00725	0.00055	0.0017

efficacy to treat COVID-19 by President Donald Trump and the repeated reported in many press media [25].

6 DISCUSSION

This study collected four million tweets on the Chinese social media Weibo, from December 9, 2019, to April 4, 2020. We investigated in the misinformation identified by the fact-checking platform Tencent—a leading Chinese tech giant. Our findings on the evolution of misinformation speak to how the attention cycle - a concept describing the complicated relationship between issue development and corresponding attention from media and the public - operates during an epidemic. Attention cycles interact with the ways that journalists frame issues, but social media communication patterns have unique trajectories that build on the emotions that the issue and the platforms may incite [26], [27]. Henry and Gordon (2001) argued that the public tends to pay more attention to an event when it is novel, and such attention will shift to other happenings or new aspects of the event as people become familiar with or bored with the incident [26]. The issue cycle is related to the influence of various factors such as journalistic practice, cultural values, and media systems, and how individuals

choose to use media and evaluate credible messages. The present study shows that discussion of misinformation follows a unique issue cycle, including city lockdown, cures and preventive measures, school reopening, and foreign countries. Sensational and emotionally reassuring misinformation characterizes the whole issue-attention cycle, with misinformation on cures and prevention flooding social media. We also study the evolution of sentiment and observe that positive sentiment dominated over the course of Covid, which may be due to the unique characteristic of “positive energy” on Chinese social media. We observe a stronger variation in sentiments in the early stages of the pandemics. Lastly, we study the media landscape during Covid via a case study on a controversial unproven cure known as Shuanghuanglian, which testifies to the importance of scientific communication in a plague. The example of how Chinese medicine and Western medicine play into unfounded beliefs illustrates the ways social and cultural belief systems clash with scientific sources. Past research suggests that social media and traditional media interact with each other, affecting public attention/discussion of an issue and leading to inter-media agenda-setting.

Our research findings provide insights into theories

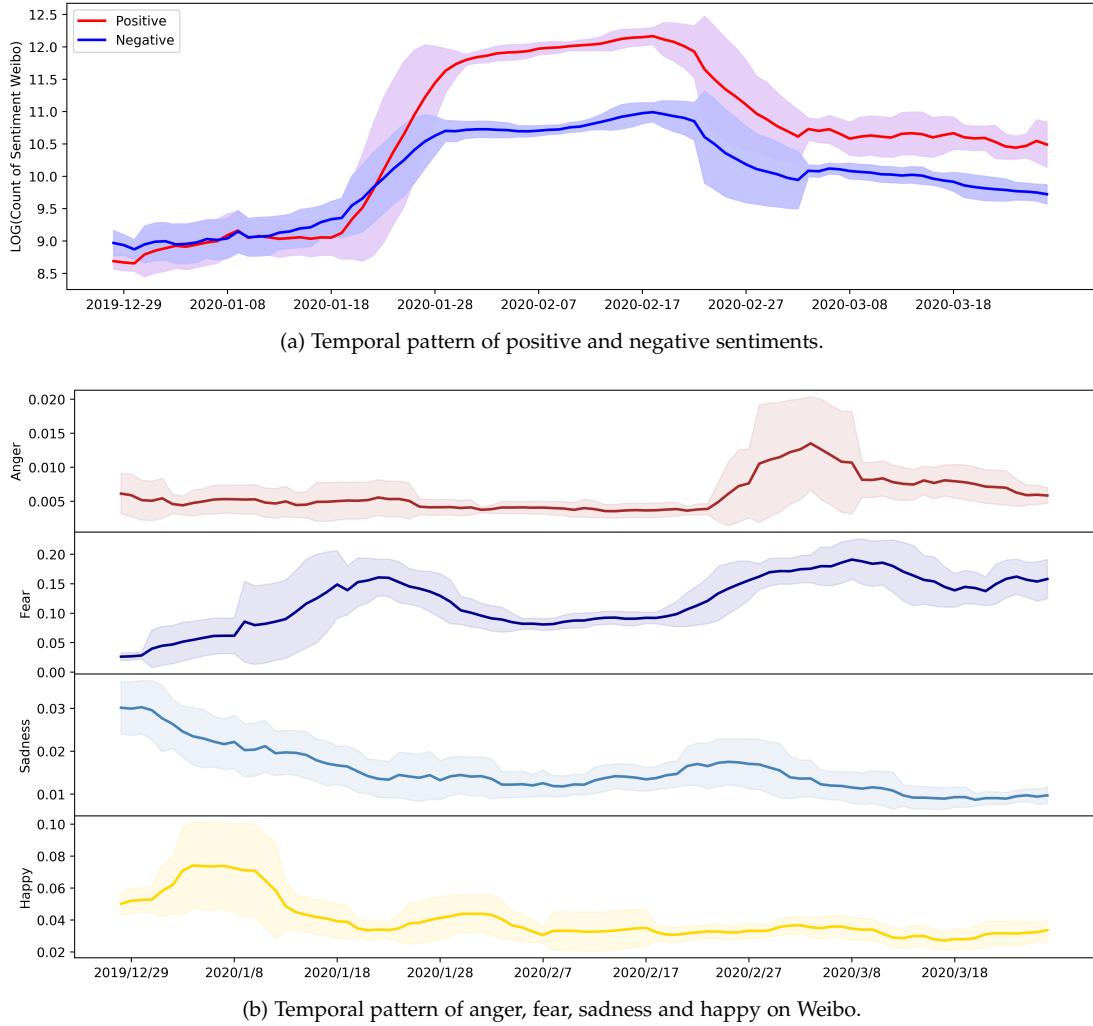


Fig. 7: Temporal pattern of positive and negative sentiments. (a) Positive and negative sentiments. The x-axis denotes the date and the y-axis denotes the count of weibo with positive or negative sentiments in log-scale. (b) Anger, fear, sadness and happy sentiments. The x-axis corresponds to the date. The y-axis corresponds to average scores of weibo with the specific sentiment of interest.

bearing on misinformation, emotion, values, and culture. According to the fuzzy-trace theory, it is not the verbatim message itself but rather the gist of the message—that is the essential or bottom-line meaning in context which determines what is remembered and shared through social media [28]. On this note, our study lent support to the scientific theory of gist communication. The gist of the vast majority of misinformation about Covid-19 on Weibo induces emotions, experiences and social values, be it related to Chinese traditional medicine or government policies. To better combat misinformation in social media, our study suggests that researchers and communication professionals need to take into account the roles of emotion, values and life experiences. Developing theories about gist communication within and across cultures will deepen the understanding of the diffusion and effect of misinformation in the era of social media.

Our findings show that traditional mainstream media—such as Xinhua News Agency and People's Daily, along with other popular market-driven or personal media accounts—play a role in the arena of misinformation. Such findings

echo the call for investigating the changing media landscape through a holistic and inter-media perspective. Our study also offers insight into combating misinformation in a massive public health crisis. As Lewandowsky et al. commented, debunking misinformation merely from a science communication perspective may not be sufficient to mitigate the adverse effects [29]. Instead, incorporating cultural, political, and social ingredients into the process of protecting the public from the threat of misinformation can yield superior outcomes. In a time of a plague, people are more likely to believe in unproven cures such as SHL. Social media amplifies emotional messages and responses; the platforms themselves contribute to fast or rushed responses, thus increasing the possibilities that misinformation will circulate widely. Seemingly naive beliefs, such as the efficacy of traditional herbal medicine, even without rigorous scientific evidence, may give residents imagined feelings of security and control. Sometimes, such misinformation about unproven cures can be perpetuated by governments and other stakeholders. As noted earlier, Donald Trump has

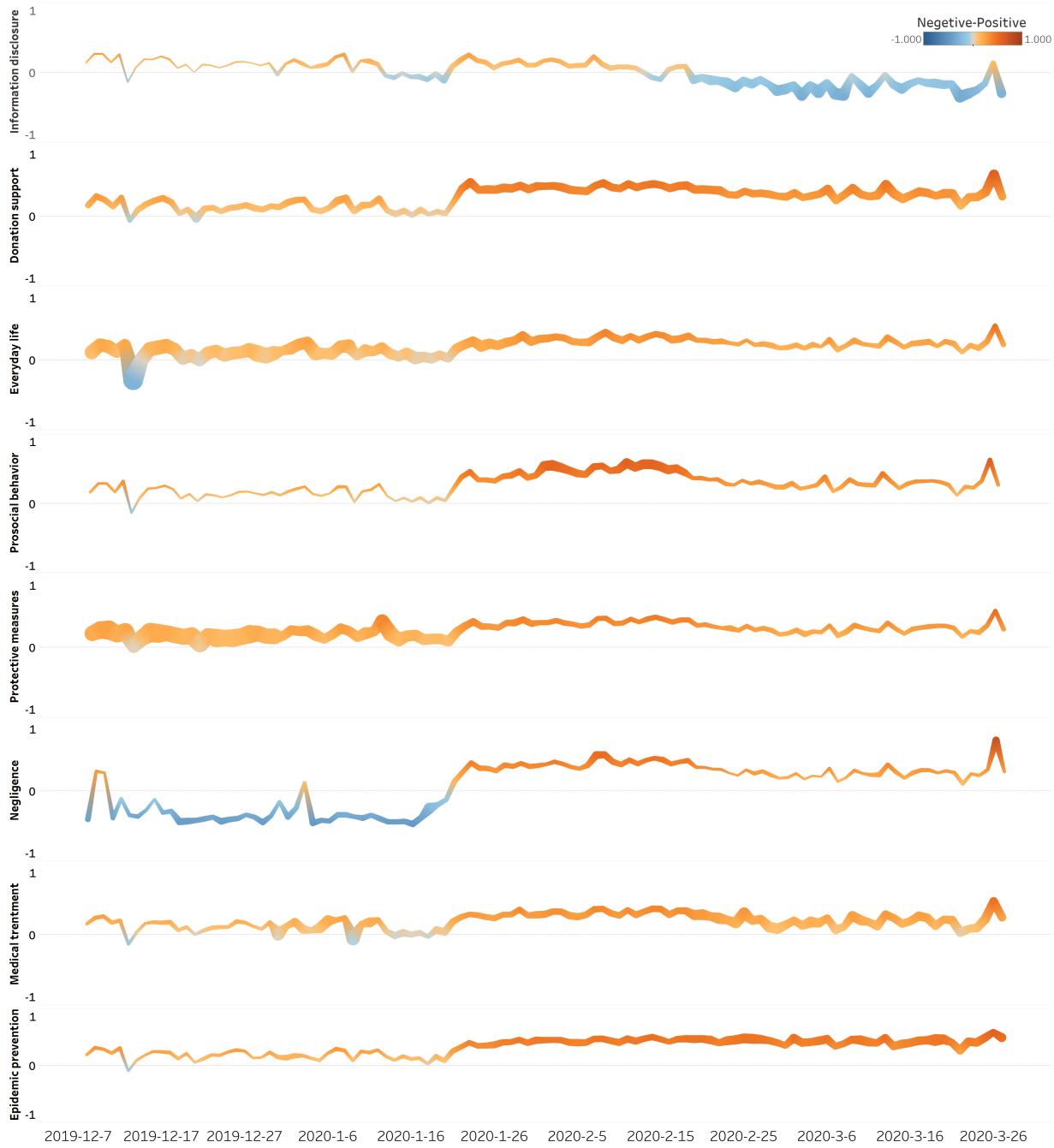


Fig. 8: **Sentiment evolution on topics on Weibo.** Sentiment evolution on eight topics. X-axis and y-axis correspond to the dates and the sentiment, respectively. Blue to red corresponds to negative to positive sentiment.

repeatedly recommended the use of hydroxychloroquine—a yet unproven treatment for COVID-19 which has well known potential risks. Meanwhile, the cacophony of different stakeholders may crowd out the authoritative opinions of legitimate experts. Therefore, only when the public, politicians, media, and other stakeholders are well informed and genuinely concerned with the public good, can they together push public discourses closer to the truth. The SHL case also speaks to the impact that a lack of consensus in the science community may have on the general public. When multiple sources report conflicting information, it can be difficult for a layperson to sort truth from fiction. When and

how to convey expert disputes to the public is a practical and meaningful question. Notably, during a pandemic crisis, the timing and amount of publicizing information about expert disputes warrant more attention and research.

Finally, our results also raise questions about the responsibilities of the different components in the media ecosystem in responding to infodemics. Ordinary citizens, the scientific community, governments and media platforms all have essential roles to play in combating disasters, rumors, and damaging effects of misinformation. It is known that ineffective scientific communication can be costly to society [30]. The case study on SHL highlights

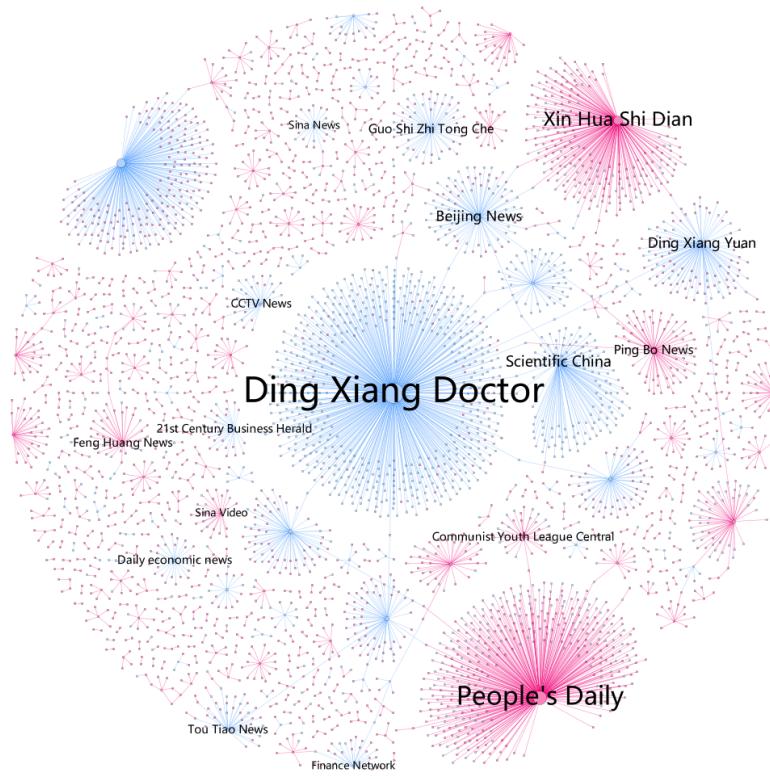


Fig. 9: **SHL-related Weibo.** Cascade of Weibo posts related to SHL. Users who are against the information (e.g., posted "fake, false, should not use, no, negative") are colored in blue. The remaining are colored in red. We removed users who are not connected to any other individuals. Accounts with a large number of posts are labeled by their account name. On January 31, 2020, Xin Hua Shi Dian, a party-owned publisher, cited this article and quoted: "SIMM and Wuhan Institute of Virology had discovered that SHL herbal remedy could "inhibit" 2019-nCov", and mentioned that SHL is currently under clinical trial in hospitals in Shanghai and Wuhan. Nine minutes later, People's Daily, with 1.1 billion followers, endorsed the same news. Four hours later, Ding Xiang Doctor, an online community for physicians, raised the concern that such news will cause a shortage, and therefore people should **not** use it for prevention. Four hours later, a personal account with 2.7 million followers shared an ironic post, which generated more reposts than many other accounts with more followers (e.g., Li Shi Pin, which has nine times more followers). This account held negative opinions about Chinese medicine and specifically satirized about a person's daily life with seven types of Chinese herb medicine. Six hours later, Scientific China, a science account with 3.5 million followers, shared the post with a catchy title, asking whether SHL can inhibit novel Coronavirus. SHL soon went viral on Weibo.



Fig. 10: **Popular tags on Weibo related to SHL.** The size of the tag is proportional to the number of mentions.

the extraordinary significance of scientific communication in pandemics. It is especially crucial for the media to fact-check evidence and make immediate corrections as needed.

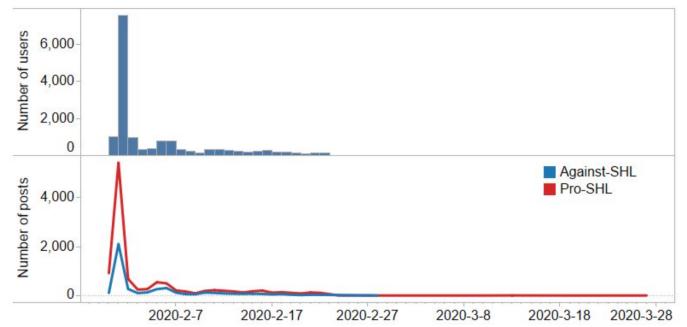


Fig. 11: **Number of users engaged in SHL events.** The y-axis correspond to number of users (upper panel) and number of posts (lower panel).

For scientists, understanding the significance of adequately communicated findings is paramount. Meanwhile, it would

be beneficial to advance science education on the public and improve their science literacy. It would be valuable for future research to investigate how any potential patterns or cycles of misinformation vary across disasters, especially pandemics of emerging infectious diseases. Wu2016mining analyzing Ebola-related rumors on microblogs in China, found that rumors tend to be related to real events associated with the virus. Specifically, rumors seem to concentrate on the source of the virus, fear about the virus, preventive measures, among others. Such a finding is, in general, consistent with those of the present study. Considering the nature of emerging infectious diseases, representative topics—such as a source of the pathogen, detection, and prevention—are more likely to generate misinformation. Future research should also look at how misinformation about pandemics differs from other crises such as earthquakes and political unrest. More concretely, it would be enlightening to compare types of rumors across different disasters or infodemics. Is there a typical lifecycle with what people attend to at different phases of a crisis? Further, comparative research across countries would be valuable too. Consider preventive measures in the U.S. where mask-wearing is more voluntary as compared to mandatory in China. Such a comparison reflects the role of the State, extending into more critical questions about democracy and the struggle with individual rights vs. collective good. Research along this line will inform worldwide joint efforts to combat a pandemic like Covid-19.

REFERENCES

- [1] J. Zarocostas, "How to fight an infodemic," *The Lancet*, vol. 395, no. 10225, p. 676, 2020.
- [2] E. K. Vraga and L. Bode, "Defining misinformation and understanding its bounded nature: Using expertise and evidence for describing misinformation," *Political Communication*, vol. 37, no. 1, pp. 136–144, 2020.
- [3] W.-Y. S. Chou, A. Oh, and W. M. Klein, "Addressing health-related misinformation on social media," *Jama*, vol. 320, no. 23, pp. 2417–2418, 2018.
- [4] P. Vinck, P. N. Pham, K. K. Bindu, J. Bedford, and E. J. Nilles, "Institutional trust and misinformation in the response to the 2018–19 ebola outbreak in north kivu, dr congo: a population-based survey," *The Lancet Infectious Diseases*, vol. 19, no. 5, pp. 529–536, 2019.
- [5] M. Del Vicario, A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi, "The spreading of misinformation online," *Proceedings of the National Academy of Sciences*, vol. 113, no. 3, pp. 554–559, 2016.
- [6] Z. Huang. (2017) China's twitter clone will soon have more users than twitter. quartz. [Online]. Available: <https://qz.com/918410/chinas-twitter-clone-sina-weibo-wb-willsoon-have-more-users-than-twitter-twtr/>
- [7] Z. Chen, P. Liu, X. Wang, and Y. Gu, "Follow whom? chinese users have different choice," *arXiv preprint arXiv:1212.0167*, 2012.
- [8] G. King, J. Pan, and M. E. Roberts, "How the chinese government fabricates social media posts for strategic distraction, not engaged argument," *American political science review*, vol. 111, no. 3, pp. 484–501, 2017.
- [9] S. L. S. C. Xiuting Li, Shaomin Wu, "Nanshan zhong: we need to be aware of fecal transmission."
- [10] A. R. Fragale and C. Heath, "Evolving informational credentials: The (mis) attribution of believable facts to credible sources," *Personality and social psychology bulletin*, vol. 30, no. 2, pp. 225–236, 2004.
- [11] T. Burkhardt, "Vaccine misinformation and social media," *The Lancet Digital Health*, vol. 1, no. 6, pp. e258–e259, 2019.
- [12] D. A. Effron and M. Raj, "Misinformation and morality: encountering fake-news headlines makes them seem less unethical to publish and share," *Psychological science*, p. 0956797619887896.
- [13] K. Starbird. (2020) How a crisis researcher makes sense of covid-19 misinformation.
- [14] A. J. Berinsky, "Rumors and health care reform: Experiments in political misinformation," *British journal of political science*, vol. 47, no. 2, pp. 241–262, 2017.
- [15] Tencent. (2017) 2017 tencent misinformation control report.
- [16] H. A. Simon, "Designing organizations for an information-rich world," *International Library of Critical Writings in Economics*, vol. 70, pp. 187–202, 1996.
- [17] M. E. Grabe, S. Zhou, and B. Barnett, "Explicating sensationalism in television news: Content and the bells and whistles of form," *Journal of Broadcasting & Electronic Media*, vol. 45, no. 4, pp. 635–655, 2001.
- [18] K. Molek-Kozakowska, "Towards a pragma-linguistic framework for the study of sensationalism in news headlines," *Discourse & Communication*, vol. 7, no. 2, pp. 173–197, 2013.
- [19] Y. Peidong and T. Lijun, "" positive energy": Hegemonic intervention and online media discourse in china's xi jinping era," *China: An International Journal*, vol. 16, no. 1, pp. 1–22, 2018.
- [20] K. Brown and U. A. Bērziņa-Čerenkova, "Ideology in the era of xi jinping," *Journal of Chinese Political Science*, vol. 23, no. 3, pp. 323–339, 2018.
- [21] L. The, "Covid-19: fighting panic with information." *Lancet (London, England)*, vol. 395, no. 10224, p. 537, 2020.
- [22] W. I. o. V. Shanghai Institute of Meteria Medica, "Shanghai institute of meteria medica, wuhan institute of virology preliminarily discovered that chinese medicine shuang huang lian can inhibit coronavirus," 2017. [Online]. Available: http://www.cas.cn/yw/202001/t20200131_4733137.shtml
- [23] D. A. Scheufele and N. M. Krause, "Science audiences, misinformation, and fake news," *Proceedings of the National Academy of Sciences*, vol. 116, no. 16, pp. 7662–7669, 2019.
- [24] A. Leshner, D. Scheufele, A. Boström, W. B. de Bruin, T. Dietz, W. Hallman, J. Henig, R. Hornik, A. Maynard, M. Nisbet et al., *Communicating Science Effectively: A Research Agenda*. The National Academies Press, 2017.
- [25] D. Jakhar and I. Kaur, "Potential of chloroquine and hydroxychloroquine to treat covid-19 causes fears of shortages among people with systemic lupus erythematosus," *Nature Medicine*, pp. 1–1, 2020.
- [26] G. T. Henry and C. S. Gordon, "Tracking issue attention: Specifying the dynamics of the public agenda," *Public Opinion Quarterly*, vol. 65, no. 2, pp. 157–177, 2001.
- [27] T.-J. Shih, R. Wijaya, and D. Grossard, "Media coverage of public health epidemics: Linking framing and issue attention cycle toward an integrated theory of print news coverage of epidemics," *Mass Communication & Society*, vol. 11, no. 2, pp. 141–160, 2008.
- [28] V. F. Reyna, "A scientific theory of gist communication and misinformation resistance, with implications for health, education, and policy," *Proceedings of the National Academy of Sciences*, 2020.
- [29] S. Lewandowsky, U. K. Ecker, and J. Cook, "Beyond misinformation: Understanding and coping with the "post-truth" era," *Journal of applied research in memory and cognition*, vol. 6, no. 4, pp. 353–369, 2017.
- [30] B. Fischhoff, "The sciences of science communication," *Proceedings of the National Academy of Sciences*, vol. 110, no. Supplement 3, pp. 14 033–14 039, 2013.
- [31] L.-h. Xu, H.-f. Lin, and J. Zhao, "Construction and analysis of emotional corpus," *Journal of Chinese information processing*, vol. 22, no. 1, pp. 116–122, 2008.

APPENDIX

A. The validity of data collection

In order to demonstrate the validity of the data collection method, we collected additional data from three sources, including COVID cases (confirmed and suspicious), Baidu search index (similar to Google search index) Baidu index shows which search terms users entered and reveals the trending search behavior., and Toutiao³. Baidu search index demonstrates the searching frequency of COVID from January to August 2020, as shown in Figure 12. Toutiao built a popularity model and assigned the popularity score based on reading, sharing, and comments based on the events, articles, and keywords. We present the variations of the clinical data and media popularity in the following figure. Comparing across all statistics, we confirm that the social media data collected with our method effectively captures the daily variations.

B. Method to obtain sentiment labels

We used the emotional corpus developed in [31] to compute the sentiment score for each article. We count the number of sentiment words as the sentiment score for each article. In this emotional corpus, there are seven emotional words, including joy, good, anger, sadness, fear, disgust and surprise. There are 27,466 sentiment words, and the intensity of emotion are graded as 1, 3, 5, 7, 9, where 9 is the strongest and 1 is the weakest. Each word is assigned to one of the sentiment categories, with an associated intensity.

Each word is also associated with polarity. Specifically, 1, -1, and 0 correspond to positive, negative, and otherwise. The sentiment analysis on Weibo post is performed based on the emotional intensity and polarity based on the words. We use a simple and widely-adopted method to compute the emotional score (W), in which $W = \text{intensity} * \text{polarity}$. Finally, the emotional intensity of each Weibo post is assigned to the categories with the largest score.

Identifying sarcasm posts from languages is a challenging task and active research field in natural language processing, some of which are also difficult for humans. To better understand how this issue may affect our research (especially the controversial posts on SHL), we manually label one thousand Weibos about general covid-related posts and specifically about the controversial topic (e.g., SHL). We find that 30 out of 1000 Covid-related Weibo posts and 83 out of 1000 SHL-related Weibo posts are related to irony and sarcasm. We then apply the sentiment analysis on the irony and sarcasm Weibos. We find that 36% Weibos can be correctly labeled, according to human labeling. Even though the performance is not high, due to the small number of irony posts, only 2.3% of the Weibos may be mislabeled due to irony and sarcasm for overall Covid-related weibo, and 5.3% may be mislabeled for the controversial issue of SHL.

C. Statistical analysis on RQ4

To test the hypothesis on RQ4, we perform one-sided t-tests on (posts with negative sentiments - posts with positive sentiments) by week. We present the result in Table 2.

3. Toutiao is a Chinese news and information content platform, a core product of the Beijing-based company ByteDance.

TABLE 2: One-sided t-test on (posts with negative sentiments - posts with positive sentiments) by week.

	t-value	p-value
week 1	-2.66	<0.05
week 2	-6.62	<0.001
week 3	-16.31	<0.001
week 4	-2.37	<0.05
week 5	-2.78	<0.05
week 6	-0.68	>0.1
week 7	-6.84	<0.001
week 8	-19.62	<0.001
week 9	-17.99	<0.001
week 10	-12.76	<0.001
week 11	-1.96	<0.1
week 12	-5.73	<0.05
week 13	-2.21	<0.1
week 14	-4.13	<0.01
week 15	-5.72	<0.001

TABLE 3: Correlation coefficients (and the p-value) between the average intensity of sentiments of Weibo posts between consecutive weeks.

	correlation coefficient	p-value
week1-week2	0.58	<0.001
week2-week3	0.42	<0.05
week3-week4	0.31	<0.1
week4-week5	0.81	<0.001
week5-week6	0.90	<0.001
week6-week7	0.90	<0.001
week7-week8	0.94	<0.001
week8-week9	0.94	<0.001
week9-week10	0.93	<0.001
week10-week11	0.96	<0.001
week11-week12	0.94	<0.001
week12-week13	0.92	<0.001
week13-week14	0.98	<0.001

D. Statistical analysis on RQ5: Correlations between the average intensity of sentiments between consecutive weeks

We study the correlations between the average intensity of sentiments between consecutive weeks in Figure 13 and Table 2, partly to address RQ5. We see strong dynamics in the sentiment variations in the beginning three weeks. The variations slow down afterward.

E. Determining the pro and against SHL posts

We use the following process to determine whether the post is pro or against the Shuanghuanglian (SHL). First, we filter out the Weibo accounts whose post about SHL has been widely reposted. Examples of these accounts include Ding Xiang Doctor, People’s Daily, Xinhua Daily, and Scientific China. Then, we manually read through all the posts to label whether the posts are pro or against SHL. We then perform a similar process on the reposts based on these labeled SHL posts. For example, if the Weibo accounts repost the original pro-SHL post without changing the content or repost without any comments, we label these as pro-SHL. A similar process is performed on the against-SHL posts. To label the remaining Weibo posts, we manually labeled 500 posts, then used Support Vector Machine to train a classifier, and then predicted the labels of all the SHL-related posts.

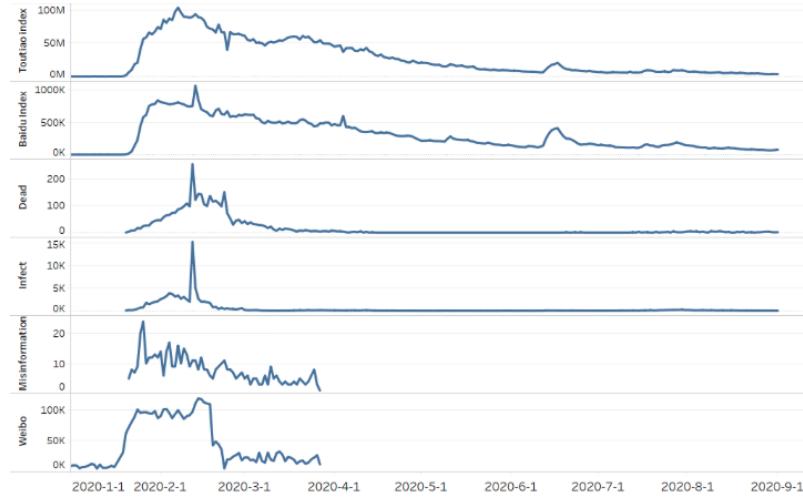


Fig. 12: Popularity of Covid according to different social media index (upper two), clinical cases (mid two), and the misinformation and Weibo data collected in our study (lower two).

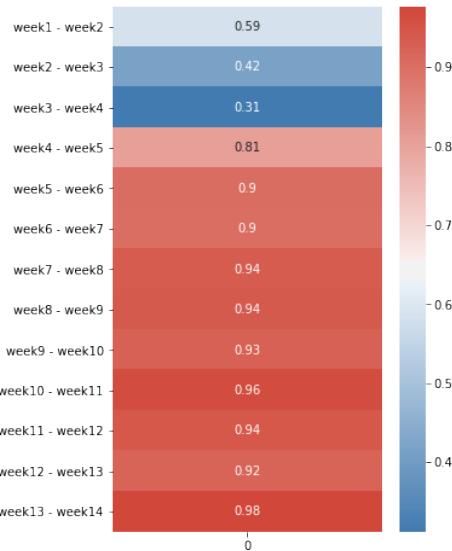


Fig. 13: Correlations between the average intensity of sentiments of Weibo posts between consecutive weeks.

We further demonstrate the effectiveness of the performance of our method in labeling pro- and against-SHL posts in Table 4.

TABLE 4: Correlation coefficients (and the p-value) between the average intensity of sentiments of Weibo posts between consecutive weeks.

	precision	recall	f1-score
Pro-SHL	0.8385	0.8084	0.8232
Against-SHL	0.7576	0.7937	0.7752
Macro average (both classes)	0.7980	0.8010	0.7992
Weighted average (both classes)	0.8037	0.8020	0.8025
Accuracy (both classes)	/	/	0.8020

The precision is computed as the true positive cases among all positive cases. The recall is computed as the true positive cases among all correctly identified cases

(i.e., true positive cases and false negative cases). F1-score is the harmonic mean of the precision and the recall. Macro average takes the average of the precision/recall/f1-score/support on the two classes. Specifically, the macro average of precision is computed as the average precision of the pro-SHL and against SHL. The weighted average is the precision/recall/f1-score/support on the two classes accounting for the proportion of that class among both classes. Accuracy measures the proportion of correct predictions among the total number of cases examined.