

Influences of Fake News Type on Audiences’ Trust in News Source, News Forwarding and Commenting: A Case Study of Sina Weibo

Wenqian Luo¹, Ratanasuda Punnahitanond², Vimvirya Limkangvanmongkol³

¹*School of Communication Arts, Bangkok University, Thailand*

Abstract

This paper investigates the impact of exposure to various types of fake news on Chinese audiences in terms of trust and behavioral reactions. According to Wardle and Derakhshan's (2017) seven types of fake news – 1) satire/parody; 2) false connection; 3) misleading content; 4) false context; 5) imposter content; 6) manipulated content and 7) fabricated content, a survey study involving 300 Chinese Weibo users aged 18-35 years old with self-administered questionnaire was carried out through online and social media. The hypothesized effect of exposure to fake news type on trust, as well as on forwarding, and commenting behavior was tested by Multiple Regression and Linear Regression Analysis. Findings have shown that fabricated and manipulated content has a strong negative effect on trust whereas trust in news source is positively associated with news forwarding but not commenting behavior. Beyond its practical contributions, this study offers managerial implications for government offices, news media, and practitioners along with influencers to avert the spread of fake news, and also serves to inform the public about the significance of digital media literacy, which creates a barrier against the adverse effects of fake news.

Keywords: Fake news exposure, Sina Weibo, trust in news source, forwarding, commenting

Background and Statement of the Problem

The proliferation of fake news within the social media has turned into one of the biggest global issues of information ecosystems. Following the 2016 presidential elections in the United States, researchers documented that fake news was viewed more than 30 million times on Facebook, which was faced with massive concerns of how it may have affected the opinion and democracy (Allcott and Gentzkow, 2017). Such a study of false news dissemination was conducted by Vosoughi, Roy, and Aral (2018) on

Twitter, involving more than 126,000 news stories, who showed that fake news spread much further, in a shorter time, and more widely than factual information, and that fake political news is highly viral. In 2017, the fake news phrase was declared the Word of the Year in Collins Dictionary because it shows the unprecedented scale of such an event in the digital era. The breakdown of the institutional safeguards against misinformation in the age of the Internet needs multidisciplinary study in order to understand the individual vulnerability and develop countermeasures (Lazer, Baum, Benkler, Berinsky, Greenhill, Menczer, and Zittrain, 2018).

Sina Weibo is the major microblogging platform in China that has more than 500 million monthly active users and presents unique challenges of fake news spread in the country. Contrasting the Western platforms, Weibo is the only site where interpersonal, group, and mass communication can be mixed, providing complex information streams where false information can be distributed in several directions at the same time. The 140-character format of the platform does not focus on verifying content, but on timeliness and, given that around 80 percent of platform users are younger than 30, most of the user base is especially vulnerable to information disorder. Research on multi-domain false news on Weibo has revealed that political misinformation has the strongest diffusion capacity, while daily-life topics generate higher volumes of posts (Guo, Ding, Yao, Liang, & Yu, 2022). The Chinese social media studies have revealed that young people tend to blindly follow the information that is posted by their friends, forming chain-like patterns of transmission which enhance false information within networks (Li & Zhang, 2018). This tendency of behavior complicates the work of platforms that must preserve the quality of information during the dynamic cycles of content sharing and the absence of proper fact-checking mechanisms.

Although previous literature has addressed the dimension of fake news detection and trends of annualization, the study was scanty to measure the difference in the impact of the various forms of fake news on the audience credibility in the Chinese scenario. Although researchers like Pennycook and Rand (2021) have been able to summarize an increasing literature on the psychology of fake news in Western settings, limited literature has been conducted on the investigation of type-specific effects in the Chinese social media landscape. It is a key gap since it is necessary to learn about type-

specific effects to drive platform governance policies that allocate content moderation resources to the most harmful types of misinformation, as well as to develop evidence-based interventions that account for the specifics of Chinese social media platforms like Sina Weibo. This study addresses this gap through research on the impact of seven types of fake news according to the typology of Wardle and Derakhshan (2017) and how it affects the user trust towards Sina Weibo and subsequent forwarding and commenting behavior. In particular, this paper relies on the Agenda Setting Theory, Source Credibility Theory, and Uses and Gratifications Theory to hypothesize that the types of fake news have had different levels of influence on platform trust and that platform trust, in its turn, has a different impact on forwarding and commenting behavior of the users.

Objective

The study aims to realize three research objectives based on the analysis of literature and applicable theories: (1) to reveal whether exposure to fake news have impact on the trust; (2) to examine whether trust in news source on Sina Weibo serves as a significant predictor of audiences' forwarding and commenting behaviors in response to fake news, thereby exploring the potential linkage between fake news exposure, trust, and behavioral outcomes; and (3) to determine whether fabricated content and manipulated content have stronger effect on the platform credibility than other types.

Expected Benefits

The results of the research are likely to offer a number of contributions in various fields. Theoretically, the present study would be useful in the communication field since it combines Agenda setting Theory, Source credibility Theory and Uses and Gratifications Theory to a single concept under the aim of analyzing how the different forms of fake news produce varied effects of platform trust in the Chinese social media setting, thereby introducing Western-centric misinformation research into a culturally specific virtual space and broadening the theoretical scope of the type-specific effects of fake news. To govern the platforms, first, the results can inform Weibo and other platforms on which types of fake news most likely to damage trust should be detected and labeled, and thus, the resources of content moderation can be optimized. Second,

the findings of the type-specific impacts on trust can guide the development of specific regulations and socialization campaigns that can be used by the government offices and policymakers in response to the most detrimental types of misinformation. Third, the results underscore the pivotal significance of open sourcing and verification activities to the practitioners and influencers of news media to ensure audience credibility. Fourth, as an educator or media literacy activist, the study offers empirical support to designing culturally-sensitive digital media literacy campaigns that explicitly train young Chinese social media consumers to be aware of different forms of fake news. Lastly, to the general audience, the research highlights the importance of critical analysis of the material of online news and the importance of digital media literacy as one of the barriers to the negative impact of fake news.

Conceptual Framework

Founded in 2009, Weibo is now the leading story-aggregating social media platform in China that boasts 511 million monthly active users. In contrast to the traditional media, which uphold editorial gatekeeping, Weibo allows any user to post content to the potential millions of followers. More than 75 percent of users peruse trending topics and hot search features to consume news on a regular basis. Engagement-focused algorithms used to make the platform might unintentionally promote sensationalism, regardless of its truthfulness, and rumor-debunking systems tend to be massively outpaced by viral rumors (Zeng, 2012). The structural vulnerability is consistent with the general findings that social media platforms have a special challenge of controlling the quality of information because of low barriers to content creation and prioritizing engagement at the expense of accuracy (Lazer et al., 2018). Song, Kwon, Lu, Fan, and Li (2021) examined COVID-19 rumor dissemination on Weibo and found that source credibility played a decisive role in whether users retransmitted rumors or corrections. When analyzing the reactions of users to misinformation, Tandoc, Lim and Ling (2020) reported that the majority of social media users merely disregard the presence of fake news posts on the platforms. Additionally, Del Vicario, Bessi, Zollo, Petroni, Scala, Caldarelli, Stanley, and Quattrociocchi (2016) demonstrated that online

misinformation tends to spread through homogeneous clusters of like-minded users, creating echo chambers that further entrench false beliefs.

Wardle and Derakhshan (2017) have come up with a detailed model of distinguishing misinformation (false content disseminated without malicious intent), disinformation (deliberately false content), and malinformation (true content disseminated to harm). They distinguished seven categories: (1) satire/parody, which refers to humorous or exaggerated content not intended to cause harm but with the potential to mislead; (2) false connection, where headlines, visuals, or captions do not accurately represent the accompanying content; (3) misleading content, involving the selective use of information to frame an issue or individual in a biased manner; (4) false context, in which genuine content is shared with false or inaccurate contextual information; (5) imposter content, where genuine sources are impersonated to lend credibility to fabricated claims; (6) manipulated content, in which genuine information or imagery is altered or distorted to deceive and (7) fabricated content, which is entirely false content designed to deceive and cause harm. This typology is complementary to the academic classification conducted by Tandoc, Lim, and Ling (2018), who defined fake news on two dimensions, one facticity and the other deception, and found six types of operations which are satire, parody, fabrication, manipulation, advertising and propaganda.

Especially the analysis of Sina Weibo by the typology applies to the fact that dissimilar types of content generate diverse vectors of different manifestations of fake news. It is demonstrated that fake and distorted information is especially bad and uses cognitive biases to provoke more emotional responses compared to other content (Zhang and Ghorbani, 2020). The authors Martel, Pennycook, and Rand (2020) supplied causal support of the idea that making people more emotionally dependent makes them more vulnerable to accepting fake news. These are deliberate types of deception which are more challenging to uncover in habitual consumption and cause dire reputational and social harm after being discovered to be untrue. Moreover, Gelfert (2018) described fake news as a deliberate display of false or misleading statements as a news, with the focus on the intentional creation of the design to mislead. Fake news on social media have to be understood in the context of content features and social realities that propagate misinformation and perceive it as legitimate as emphasized by

Shu, Sliva, Wang, Tang, and Liu (2017). Zhou and Zafarani (2020) conducted a thorough overview of the fake news theories and detection approaches and stated that the key problem is the interdependence between the content features and user psychology.

This paper combines three theories that go hand in hand. Firstly, the Agenda Setting Theory (McCombs & Shaw, 1972) indicates that mass media determines what people regard as important as selective coverage and focus on. Within the environment of fake news, repeated exposure to untrue information can influence the perceived significance of some issue in the audiences, and, at the same time, influence the perception of the reliability of media sources. Vargo and Guo (2017) affirmed that fake news can be an active participant in intermedia agenda-setting processes and proved that the fabricated content could shape what the rest of media ecosystem can talk about.

Secondly, the Source Credibility Theory (Hovland & Weiss, 1951) provides that the acceptance of the message will be based on the perceived trustworthiness and expertise of the source(s). If viewers learn about fake news falsity, platform credibility is reduced, especially false content that is intentionally created. Park, Fisher, Flew, and Dulleck (2020) built upon this knowledge and proved that the use of social media per se is a contributor to the decrease in news trust across the world, and the features of the platform significantly mediate the way people trust news.

Thirdly, the Uses and Gratifications Theory or UGT (Katz, Blumler, & Gurevitch, 1974) describes the use of the media as a goal-oriented behavior that pursues a particular gratification. Upon the erosion of information-seeking needs by fake news, the users can alter the trust and engagement behavior. According to Pennycook, McPhetres, Zhang, Lu, and Rand (2020), individuals post misinformation because they do not evaluate the accuracy of the information; therefore, they believe that the desire to experience gratification may override critical appraisal mechanisms. Altay, Hacquin, and Mercier (2022) further demonstrated that reputation concerns serve as a significant deterrent to sharing false news, suggesting that social motivations interact with gratification-seeking in complex ways.

Review of literature and relevant theories above leads to development of the conceptual framework of this study. The framework integrates Wardle and Derakhshan's (2017) seven-type fake news typology with three established communication theories --

Agenda Setting Theory, Source Credibility Theory, and UGT -- to examine how users’ exposure to different types of fake news affects trust in Sina Weibo, and how trust in news source in turn influences their forwarding and commenting behaviors. The independent variables are exposure to the seven types of fake news exposure (satire/parody, false connection, misleading content, false context, imposter content, manipulated content, and fabricated content). The mediating variable is trust in Sina Weibo as a news source. The dependent variables are forwarding behavior and commenting behavior. Based on the conceptual framework shown in Figure 1, four research hypotheses are proposed as follow:

HP#1: Chinese audiences’ exposure to different types of fake news on Sina Weibo influences their trust in Sina Weibo as news source.

HP#2: Chinese audiences’ trust in Sina Weibo as news source influences their forwarding fake news.

HP#3: Chinese audiences’ trust in Sina Weibo as news source influences their commenting on fake news.

HP#4: Among seven types of fake news, fabricated content and manipulated content have higher impact on Chinese audiences’ trust in Sina Weibo as news source than do other types.

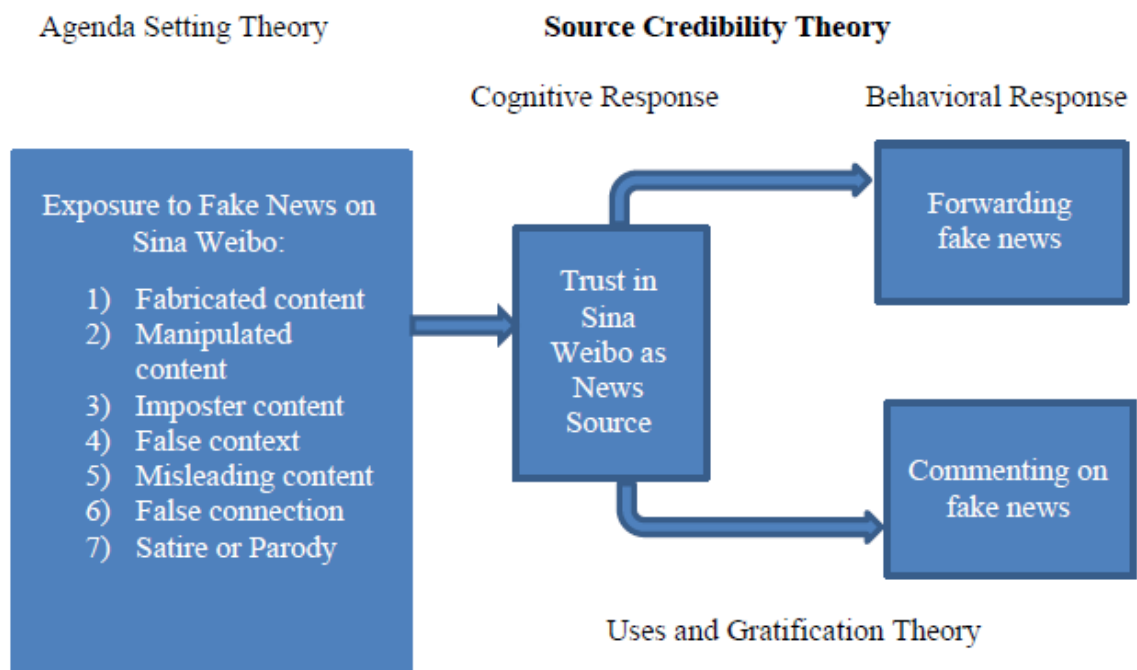


Figure 1. Conceptual Framework

Research Methodology

Online survey was adopted as a quantitative method to test the relationships of fake news exposure, trust and the behavioral outcomes. The target population was Chinese Weibo users aged 18-35, who also consume news content on the platform on a regular basis. The choice of this age group was due to the fact that users who are below 30 years old make up about 80% of the active users of Weibo. Online distribution channels were used to recruit 300 participants using stratified random sampling with G*Power analysis ($\alpha = .05$, power = .80, medium effect size). A filter question facilitated that the participants had received news information during three months before data collection that they eventually found to be untrue or deceptive.

The self-administered questionnaire had 6 sections: (A) demographics, which included gender, age, education, income; (B) exposure to each of the seven fake news types measured using 5-point scales (1 = Never/ 2 = Seldom / 3 = Moderate/ 4 = Often / 5 = Always); (C) trust in Sina Weibo as a news source (3 items measuring attractiveness, trustworthiness, expertise, $\alpha = .87$) and (D) forwarding behavior (3 items, $\alpha = .82$) (E) commenting behavior (3 items, $\alpha = .79$) and (F) engagement reasons basing on UGT dimensions. Attitudinal items were all Likert scales of 5 points. The criteria for interpreting the mean scores of each variable were established based on the equal interval method for a 5-point scale, where scores of 1.00 to 1.80 were classified as very low, 1.81 to 2.60 as low, 2.61 to 3.40 as moderate, 3.41 to 4.20 as high, and 4.21 to 5.00 as very high. The instrument was pretested with 30 respondents to establish initial reliability of the measurement scales, subjected to back-translation (English-Chinese-English) and pretested as well. Following the pretest, the questionnaire was administered to the full sample of 300 respondents for the main study. The collection of data was done using online and social media.

The processing and cleaning of data was then followed by the evaluation of multicollinearity ($VIF < 5$). The effect of the exposure of the audiences on the trust (HP#1 and HP#4) was subjected to Multiple Regression Analyses as compared to the impact of the trust on the forwarding by the audiences (HP#2) and the commenting by the audiences (HP#3) using the Linear Regression Analysis at a significance level of .05 or below. It is also worth mentioning that, although the conceptual framework implies trust

as a variable between fake news exposure and behavioral outcomes, the current study does not utilize a formal statistical mediation analysis e.g. the Sobel test or bootstrapping techniques suggested by Hayes (2018). Rather, the analytical method considers each of the hypothetical relationships an independent test using distinct regression models and then tests, sequentially, the effect of exposure to fake news on trust (HP#1), the effect of trust on forwarding (HP#2) and commenting (HP#3) as independent predictors. It is a sequential regression approach that enables the research to be able to gauge the direction and significance of each path without being able to claim any mediation. Also, it is important to mention that the general regression equation of the HP#1 exploits a composite measure of total exposure to fake news as a single predictor, but the type-specific analysis in the lower section of Table 4 introduces all seven types of fake news as simultaneous predictors in a model of multiple regression, which is why this difference in the model specification is expected and status of the positive overall coefficient is the cause of the negative individual coefficients of fabricated and manipulated content.

Research Results

In terms of demographic characteristics, 51% of respondents were female, 47.3% of them were male and 1.7% of them responded not to say. The largest age group was 20-29 years old (45.7%), followed by 30-35 years old (28.0%), and 18-19 years old (15.7%). Majority of the respondents hold bachelor's degrees (47.3%), followed by those with master's degree or higher (29.7%), and those with high-school diploma or lower (14.0%). As for frequency and percentage of time spent using media to read news daily, majority of them spent 30-59 minutes (32.7%) or less than 30 minutes (21.3%). Detail characteristics of the respondents are presented in Table 1.

Table 1. Frequency and Percentage of Respondents’ Characteristics (N = 300)

Characteristics	Category	Frequency (n)	Percentage (%)
Gender	Female	153	51.0
	Male	142	47.3
	Prefer not to say	5	1.7
Age	18–19	47	15.7
	20–29	137	45.7
	30–35	84	28.0
	Other	32	10.7
Education	High school or below	42	14.0
	Bachelor’s degree	142	47.3
	Master’s or above	89	29.7
Time spent on using media to read news daily	Rarely	17	5.6
	Several times a week	41	13.7
	Less than 30 minutes	64	21.3
	30-59 minutes	98	32.7

Respondents’ exposure to fake news of all types is considered moderate level ($\bar{X} = 3.03$, S.D. = 1.04). The means of misleading content ($\bar{X} = 3.41$, S.D. = 1.05), and false context ($\bar{X} = 3.37$, S.D. = 1.09) were found to be the highest, which implies that these types of content were the most common issues encountered by the respondents, whereas fabricated content ($\bar{X} = 2.53$, S.D. = 1.22) and satire/parody ($\bar{X} = 2.64$, S.D. = 1.08) represented relatively low, which implies that these types of content were the least common issues exposed by the respondents.

The means of trust in Sina Weibo as a source of news was moderate ($\bar{X} = 3.28$, S.D. = 0.91), which indicated that the level of trust towards the usefulness of the platform as a news source was moderate. The average scores in forwarding behavior ($\bar{X} = 2.92$, S.D. = 1.16) and commenting behavior ($\bar{X} = 2.88$, S.D. = 1.21) were slightly lower, which means that active engagement with the content of fake news is not as prominent.

The Cronbach’s Alpha coefficients of all the measurement scales were ranging between .85 and .91 and this implies that all the constructs had sufficient measure reliability. It is worth mentioning that the reliability coefficients (Cronbach’s Alpha) measured in Table 2 were calculated using the entire sample of 300 respondents obtained in the course of the main research, and not the pretest sample of 30 respondents. Although the pretest was performed to ensure that the instrument was initially adequately good, the reliability values provided in this context will be indicative of the psychometric performance of the scales over the entire sample, which will give a

more robust and generalizable measurement of the internal consistency. Table 2 of the descriptive statistics such as means and standard deviations are also calculated on the entire sample of 300 respondents. Table 2 is a summary of the descriptive statistics and reliability coefficients of each measured variable.

Table 2. Mean, Standard Deviation, and Reliability of Measured Variables

Variable	N	Mean	S.D.	No. of Items	Cronbach's Alpha
Fake News Exposure	300	3.03	1.04	7	.90
– Satire/Parody	300	2.64	1.08	1	-
– False Connection	300	3.02	1.14	1	-
– Misleading Content	300	3.41	1.05	1	-
– False Context	300	3.37	1.09	1	-
– Imposter Content	300	2.85	1.12	1	-
– Manipulated Content	300	2.71	1.18	1	-
– Fabricated Content	300	2.53	1.22	1	-
Trust in Sina Weibo	300	3.28	0.91	3	.85
Forwarding Behavior	300	2.92	1.16	3	.88
Commenting Behavior	300	2.88	1.21	3	.91

Further analysis of the trust levels across different fake news types revealed important variations. As shown in Table 3, the exposed to satire or parody, respondents reported the highest level of trust in Sina Weibo ($\bar{X} = 3.47$, S.D. = 0.55), followed by false connection ($\bar{X} = 3.36$, S.D. = 0.60). In contrast, exposure to fabricated content resulted in the lowest trust level ($\bar{X} = 2.41$, S.D. = 0.72), closely followed by manipulated content ($\bar{X} = 2.53$, S.D. = 0.69). These results suggest that the more deliberately deceptive forms of fake news are associated with greater erosion of platform trust.

Table 3. Mean and Standard Deviation of Trust in Sina Weibo Categorized by Fake News Type

Fake News Type	N	Mean	S.D.
Satire/Parody	300	3.47	0.55
False Connection	300	3.36	0.60
Misleading Content	300	3.24	0.58
False Context	300	3.12	0.61
Imposter Content	300	3.01	0.65
Manipulated Content	300	2.53	0.69
Fabricated Content	300	2.41	0.72

According to Table 4 that demonstrates the results of the detailed regression coefficients, according to Chinese audiences' exposure to various kinds of fake news on Sina Weibo has an impact on their trust in Sina Weibo as a news source, Multiple Regression Analysis results showed the support of HP#1 that exposure to fake news was a strong predictor of trust in Sina Weibo ($R^2 = .397$, $p < .001$). In this regard, the support of HP#1 was obtained ($\beta = .34$, $p < .001$). The overall positive value of the composite coefficient ($\beta = .34$) is to be viewed in terms of the aggregate measure of exposure that represents the overall frequency of exposure to fake news, as opposed to the trust-destroying impact of particular types of content. This positive correlation means that respondents who reported a higher overall exposure to fake news materials on Sina Weibo also reported a higher general awareness of the platform as a source of information, and this conceptual distinction is not similar to the type-specific effects of specific types of deceptive content that researchers are considering in the type-specific model below. The difference between this positive overall coefficient and the negative coefficients on fabricated and manipulated content in the type-specific model will indicate the difference in model specification since the overall model has one composite predictor, as compared to the type specific model in which the seven types are entered multitaskedly and the unique contribution of each type is obtained holding the rest constant.

When analyzing the impact of exposure to each type of fake news separately, the effects of fabricated content ($\beta = -.31$, $p < .001$) and manipulated content ($\beta = -.28$, $p < .001$) were significant and negative on trust. Interestingly, the effects of imposter content ($\beta = -.05$, $p > .05$), false context ($\beta = -.07$, $p > .05$), false connection ($\beta = -.03$, $p > .05$), misleading content ($\beta = -.09$, $p > .05$), and satire/parody ($\beta = .02$, $p > .05$) were insignificant.

Table 4. Multiple Regression Analysis Results on Predictors of Trust on Sina Weibo

Predictor	B	SE	Beta	t	P-Value	Result
Exposure to all types of fake news	0.55	0.04	.34	6.02	<.001***	Supported
<i>Note. R² = .397, F(1, 298) = 196.32, p < .001</i> <i>*p < .05, **p < .01, ***p < .001</i>						
Exposure to satire/parody	0.03	0.05	0.02	0.41	0.681	
Exposure to false connection	-0.07	0.05	-0.03	-0.62	0.537	
Exposure to misleading content	-0.06	0.06	-0.09	-1.67	0.096	
Exposure to false context	-0.10	0.05	-0.07	-1.32	0.187	
Exposure to imposter content	-0.12	0.05	-0.05	-0.91	0.364	
Exposure to manipulated content	-0.20	0.05	-0.28	-4.95	<.001***	Supported
Exposure to fabricated content	-0.22	0.05	-0.31	-5.52	<.001***	Supported

$R^2 = 0.39$, Adjusted $R^2 = 0.37$, $F(7, 292) = 26.20$, $p < 0.001$

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 5 presents the impact of trust in Sina Weibo on both behavioral outcomes, based on Linear Regression Analyses. In the case of HP#2: Chinese audiences’ trust in Sina Weibo as news source influences their forwarding fake news, it was found that trust in Sina Weibo was supported that trust in Sina Weibo significantly predicted forwarding behavior ($R^2 = .072$, $\beta = .269$, $p < .001$) meaning that the more trust people have in the platform, the more news content, including fake news, they will forward. HP#2 was supported. This positive relationship, albeit weak, indicates that platform credibility facilitates content dissemination, consistent with the inattention-based sharing mechanism described by Pennycook and Rand (2021).

As for HP#3: Chinese audiences’ trust in Sina Weibo as news source influences their commenting on fake news, was not supported as it did not significantly influence commenting behavior ($R^2 = .006$, $\beta = .076$, $p > .05$). HP#3 was thus not supported.

Table 5. Linear Regression Analysis on Trust as Predictor of Behavioral Outcomes

Behavioral Outcomes	R ²	B	Beta	t	P-Value	Result
Forwarding Behavior	.072	0.34	.269	4.81	<.001***	Supported
Commenting Behavior	.006	0.10	.076	1.31	.191	Not Supported

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

In the case of HP#4: Among seven types of fake news, fabricated content and manipulated content have higher impact on Chinese audiences’ trust in Sina Weibo as

news source than do other types, results of Regression Analysis confirmed that fabricated content ($\beta = -.31, p < .001$) and manipulated content ($\beta = -.28, p < .001$) had significantly stronger negative influences on trust than any other fake news type. However, as other five types of fake news had insignificant effect on trust, HP#4 was partially supported.

Summary of the Study

This online survey research examined the effects of seven types of fake news on audiences' trust in Sina Weibo and subsequent forwarding and commenting behaviors among 300 Chinese Weibo users aged 18–35 years old. Four research hypotheses were tested using Multiple Regression and Linear Regression Analyses. Results indicated that users' exposure to different fake news types significantly influenced their trust (HP#1 supported). Trust in Sina Weibo was found to positively predict forwarding behavior (HP#2 supported) but not commenting behavior (HP#3 not supported). HP#4 was partially supported, showing that fabricated content ($\beta = -.31, p < .001$) and manipulated content ($\beta = -.28, p < .001$) had stronger negative influences on trust compared to other types of fake news, whereas other types had no significant effect. Overall, the findings demonstrate that exposure to deliberately deceptive forms of fake news are most damaging to platform credibility, and that trust in news source facilitates content sharing, but does not drive commenting engagement.

The study also holds significant value to the misinformation research literature since it shows that the types of fake news have dissimilar levels of influence on platform trust in the Chinese social media setting, which has been underrepresented in the current literature primarily because of the influence of Western platforms. The combination of Agenda Setting Theory, the Source Credibility Theory, and Uses and Gratifications Theory enables the study to have a multi-theoretical approach, which adds to the comprehension of the functioning of fake news in culturally specific digital ecosystems. The observation that trust is a predictor of forwarding, but not commenting behavior is a particularly useful fact, as it indicates that there are two different ways of engagement, which are regulated by completely different psychological mechanisms. The informational gratifications and platform credibility judgment seem to drive

forwarding, whereas commenting is more directly related to the emotional arousal and social interaction motivations, which are independent of judgments based on trust, as proposed by the dual-process model presented by Pennycook and Rand (2021).

There are a few drawbacks in this study that must be mentioned. To begin with, the survey research design is cross-sectional, which does not allow drawing causal conclusions regarding the correlations between exposure to fake news, trust, and behavioral consequences. Second, the use of self-reported indicators on fake news exposure can be influenced by recall bias since respondents might not recall correctly the nature and frequency of fake news they came across. Third, the current study had sufficient statistical power due to the sample 300 Weibo users between the ages of 18 and 35, though further research should increase the sample size, which would improve the generalizability and acceptability of the research results, especially considering that the number of Weibo users in China is large and heterogeneous. Fourth, the trust-behavior models have relatively low values of R^2 ($R^2 = .072$ for forwarding, and $R^2 = .006$ for commenting), which implies that trust is not a major determinant of trust behavior outcome and other unmeasured factors, including emotional state, influence of the social network, and platform-specific characteristics are probably significant determinants of trust behavior outcome that should be investigated in further research.

Discussions

The results show that there are several theoretically and practically important patterns with regards to fake news and trust among Chinese social media. Such a high level of negative impact of fabricated and manipulated content on trust corresponds to the findings of previous research. Consistent with Tandoc and colleagues (2018), who argued that fabricated and manipulated content is widely perceived as more misleading because it mimics credible news forms, the current study confirmed that these two types of fake news significantly undermined trust in Sina Weibo more than other types. Zhang and Ghorbani (2020) further demonstrated that fabricated content on Chinese social media platforms exploits users' cognitive biases, making false information more likely to be initially accepted as true. The present study extends these findings by providing empirical evidence that exposure to fabricated content ($\beta = -.31$, $p < .001$) and

manipulated content ($\beta = -.28, p < .001$) produced the strongest negative effects on platform trust among all seven types.

It is also not surprising that satire and parody are rather unimportant in terms of trust, as it was also discovered by Tandoc and colleagues (2018) and Wardle and Derakhshan (2017), who both found that audiences tend to perceive satiric content as funny, as opposed to being misleading. This observation is also reflected by a study by Pennycook and Rand (2021) who claimed that a lack of critical thinking is what mainly motivates poor truth discernment and not the content itself, in that users can often be in the corrective position and determine that satire is non-serious.

The fairly low but meaningful correlation between trust and forwarding behavior ($R^2 = 0.072$) proves the paradox of the past literature: trust in a platform encourages sharing and does not guarantee checking of the content (Allcott and Gentzkow, 2017). The present observation is similar to Vosoughi, Roy, and Aral (2018) large-scale analysis, which established that the novelty of fake news, as opposed to intentionality to use it to deceive, was a key factor behind the sharing behavior. Similar words were used by Pennycook and colleagues (2020) who described this as inattention based sharing as the sharing among users is driven by the desire to share rather than consider the correctness of the content. Guess, Nagler, and Tucker (2019) also discovered that misinformation was infrequently shared on Facebook in general but was concentrated in several groups of demographic traits, which indicates that trust in the platform can selectively enhance forwarding in particular groups of users.

Particularly, it is interesting that the correlation between trust and commenting behaviour is insignificant. This finding contradicts the results described by Stieglitz and Dang-Xuan (2013), who indicated that the higher the engagement in comments, the more emotive content was being triggered. According to Martel and colleagues (2020), the use of emotion, instead of reason, facilitates the process of believing fake news, which can be the reason why commenting, as an emotion-sensitive behavior, does not depend on evaluations of trust. The conclusions made by Bakir and McStay (2018) were also based on the idea that fake news takes advantage of an economy of emotions, where user engagement is not based on the credibility of the news source but instead on affective reactions. The current research indicates that commenting behaviour in

Chinese Weibo scenario is more of a motivation to get emotional arousal and desire social interaction, as opposed to perceived credibility of the platform.

It is possible to draw the results in accordance with three developed theoretical frameworks. The Agenda setting theory (McCombs and Shaw, 1972) is useful in understanding how the cumulative exposure of fake news on Weibo is related to the perceived salience of the issue, and the perceived credibility of the site itself. Vargo and Guo (2017) have proven that fake news can be viewed as an active agent of the intermedia agenda-setting, and the present study has gone further to indicate that the most misleading forms of content have the highest determining power on how viewers will assess the credibility of the platform. The reason why fabricated and manipulated content had the largest negative impact on trust is well explained by the Source Credibility Theory (Hovland and Weiss, 1951) because in these instances of deliberate deception, the two elements of perceived trustworthiness and expertise are in the most direct conflict with one another. Appelman and Sundar (2016) have shown that message credibility is a unique construct than source credibility and that they are evaluated by viewers separately, which could be the reason why some forms of fake news harm platform trust more than others. The Uses and Gratifications Theory (Katz et al., 1974) can be used to explain the results of the conflict between the results of forwarding and commenting. The informational gratifications that determine forwarding and the non-significant ones that determine commenting indicate that trust is positively related to forwarding and does not impact commenting because commenting is driven by emotions and social interaction gratifications and does not rely on forms of credibility judgments (Stieglitz and Dang-Xuan, 2013). Altay, Hacquin, and Mercier (2022) discovered that reputational considerations are a major deterrent to the sharing of fake news, which is one possible reason why the trust-forwarding relationship is low.

Before doing this study, it was assumed that not every type of fake news would have the same impact on trust. It was predicted that intentional misrepresentative forms like fabricated and manipulated posts would cause the greatest harm to the platform credibility, and satire or parody would not have much influence. To some extent, the results supported these expectations. Nevertheless, the result of trust strongly predicting forwarding, and not commenting behaviour was partly surprising. The first hypothesis

assumed that trust would affect both types of interaction, but findings show a varying motivation driver can be used to comment. This surprising result complements the knowledge about the functioning of the audience attentions to fake news in terms of specific psychological processes, which are in line with the dual-process approach outlined by Pennycook and Rand (2021).

Recommendations

The present research not only gives managerial implications on government offices, news media, and practitioners, alongside the influencers to curb the spread of fake news, but also to the masses to realize the significance of digital media literacy that will protect against the increasing detrimental effects of fake news.

In the case of platform governance, Weibo will want to focus on the detection and labeling of fabricated and manipulated content since these forms of content have the most significant negative impact on user trust. Strategies to limit the dissemination of intentionally misleading information, by increasing the algorithmic transparency and formal liaisons with independent fact checking organizations, would help prevent the proliferation of the intentional misinformation, as Lazer and colleagues (2018) put, would be a vital step towards establishing new institutionalized protection against misinformation. Pennycook and colleagues (2020) demonstrated that simple accuracy prompts at the point of sharing can significantly reduce misinformation dissemination, a strategy that Weibo could implement to interrupt heuristic sharing patterns.

Media literacy education programs should be created to target the young Chinese users of social media and specifically learn to understand the types of fabricated and manipulated content. Leng, Zhai, Sun, Wu, Selzer, Strover, and Jiang (2022) found that COVID-19 misinformation on Weibo was driven by cultural and social entanglements, suggesting that media literacy programs should be culturally adapted to address the specific vulnerabilities of Chinese social media users. To news media practitioners and influencers, transparent sourcing and strict verification practices are the sole way to keep the audience trustful. As Aïmeur, Amri, and Brassard (2023) noted in their comprehensive review of social media misinformation, the convergence of fake news, disinformation, and malinformation requires multi-stakeholder responses that

combine technological, educational, and regulatory approaches. To conduct research in the future, cross-platform comparative research to include platforms like WeChat, TikTok, and similar platforms internationally would help us understand whether the type-trust patterns can be considered platform-specific or culturally generic. Causal inferences would be better supported by experimental methods involving exposure of the experimental participants to controlled fake news stimuli in natural real-time digital environments. Moreover, the fact that the level of R^2 in the trust-behavior regression models is relatively low ($R^2 = .072$ for forwarding, and $R^2 = .006$ for commenting) demonstrates that trust in news source is not the only factor that can explain the variance in audience engagement behavior. The future studies should thus include other predictor variables like the emotional arousal, perceived social influence, platform-specific algorithmic exposure and the individual media literacy level into the regression models to enhance the explanatory power and get better R^2 values. The application of Structural Equation Modeling or hierarchical regressions modeling that incorporate these other variables would give a better picture of the complex processes by which exposure to fake news is converted into behavioral results in social media.

Acknowledgement

The authors would like to express sincere gratitude to the School of Communication Arts, Bangkok University, for providing academic support and resources throughout this research. Special thanks are extended to all the survey participants who generously contributed their time and insights. The first author also highly appreciates the guidance and support of his thesis committee (Assoc. Prof. Dr. Ratanasuda Punnahitanond, Dr. Vimviriya Limkangvanmongkol, Asst. Prof. Dr. Duangkamol Chartprasert, and Dr. Yaninee Petcharanan), and the constructive feedback from peer reviewers that helped improve the quality of this paper.

Reference

Aïmeur, E., Amri, S., & Brassard, G. (2023). Fake news, disinformation and misinformation in social media: A review. *Social Network Analysis and Mining*, 13(1), 30. <https://doi.org/10.1007/s13278-023-01028-5>.

- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211-236. <https://doi.org/10.1257/jep.31.2.211>.
- Altay, S., Hacquin, A.-S., & Mercier, H. (2022). Why do so few people share fake news? It hurts their reputation. *New Media & Society*, 24(6), 1303-1324. <https://doi.org/10.1177/1461444820969893>.
- Appelman, A., & Sundar, S. S. (2016). Measuring message credibility: Construction and validation of an exclusive scale. *Journalism & Mass Communication Quarterly*, 93(1), 59-79. <https://doi.org/10.1177/1077699015606057>.
- Bakir, V., & McStay, A. (2018). Fake news and the economy of emotions: Problems, causes, solutions. *Digital Journalism*, 6(2), 154-175. <https://doi.org/10.1080/21670811.2017.1345645>.
- Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3), 554-559. <https://doi.org/10.1073/pnas.1517441113>.
- Gelfert, A. (2018). Fake news: A definition. *Informal Logic*, 38(1), 84-117. <https://doi.org/10.22329/il.v38i1.5068>.
- Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, 5(1), eaau4586. <https://doi.org/10.1126/sciadv.aau4586>.
- Guo, B., Ding, Y., Yao, L., Liang, Y., & Yu, Z. (2022). Characterizing multi-domain false news and underlying user effects on Chinese Weibo. *Information Processing & Management*, 59(4), 102968. <https://doi.org/10.1016/j.ipm.2022.102968>.
- Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. (2nd ed.). Guilford Press.
- Hovland, C. I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. *Public Opinion Quarterly*, 15(4), 635-650. <https://doi.org/10.1086/266350>.
- Katz, E., Blumler, J. G., & Gurevitch, M. (1974). Uses and gratifications research. *Public Opinion Quarterly*, 37(4), 509-523. <https://doi.org/10.1086/268109>.

- Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., & Zittrain, J. L. (2018). The science of fake news. *Science*, 359(6380), 1094-1096. <https://doi.org/10.1126/science.aao2998>.
- Leng, Y., Zhai, Y., Sun, S., Wu, Y., Selzer, J., Strover, S., Messer, K., & Jiang, Y. (2022). Misinformation during the COVID-19 outbreak in China: Cultural, social and political entanglements. *IEEE Transactions on Big Data*, 8(2), 422-437. <https://doi.org/10.1109/TBDATA.2021.3055758>.
- Li, W. J., & Zhang, X. (2018). How acquaintance sharing drives chain transmission of misinformation on Chinese social media. *Chinese Journal of Communication*, 11(1), 78-96. <https://doi.org/10.1080/17544750.2017.1372081>.
- Martel, C., Pennycook, G., & Rand, D. G. (2020). Reliance on emotion promotes belief in fake news. *Cognitive Research: Principles and Implications*, 5(1), 47. <https://doi.org/10.1186/s41235-020-00252-3>.
- McCombs, M. E., & Shaw, D. L. (1972). The agenda-setting function of mass media. *Public Opinion Quarterly*, 36(2), 176-187. <https://doi.org/10.1086/267990>.
- Park, S., Fisher, C., Flew, T., & Dulleck, U. (2020). Global mistrust in news: The impact of social media on trust. *The International Journal on Media Management*, 22(2), 83-96. <https://doi.org/10.1080/14241277.2020.1799794>.
- Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science*, 31(7), 770-780. <https://doi.org/10.1177/0956797620939054>.
- Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, 25(5), 388-402. <https://doi.org/10.1016/j.tics.2021.02.007>.
- Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, 19(1), 22-36. <https://doi.org/10.1145/3137597.3137600>.
- Song, Y., Kwon, K. H., Lu, Y., Fan, Y., & Li, B. (2021). The “parallel pandemic” in the context of China: The spread of rumors and rumor-corrections during COVID-19 in Chinese social media. *American Behavioral Scientist*, 65(14), 2014-2036. <https://doi.org/10.1177/00027642211003153>.

- Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and information diffusion in social media: Sentiment of microblogs and sharing behavior. *Journal of Management Information Systems*, 29(4), 217-248. <https://doi.org/10.2753/MIS0742-1222290408>.
- Tandoc, E. C., Lim, Z. W., & Ling, R. (2018). Defining “fake news”: A typology of scholarly definitions. *Digital Journalism*, 6(2), 137-153. <https://doi.org/10.1080/21670811.2017.1360143>.
- Tandoc, E. C., Lim, D., & Ling, R. (2020). Diffusion of disinformation: How social media users respond to fake news and why. *Journalism*, 21(3), 381-398. <https://doi.org/10.1177/1464884919868325>.
- Vargo, C. J., & Guo, L. (2017). Networks, big data, and intermedia agenda setting: An analysis of traditional, partisan, and emerging online U.S. news. *Journalism and Mass Communication Quarterly*, 94(4), 1031-1055. <https://doi.org/10.1177/1077699016679976>.
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151. <https://doi.org/10.1126/science.aap9559>.
- Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking*. Council of Europe.
- Zeng, Y. (2012). On the spread of fake news on Weibo. *Modern Economy*, 3(4), 100-101. <https://doi.org/10.4236/me.2012.34051>.
- Zhang, X., & Ghorbani, A. A. (2020). An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management*, 57(2), 102025. <https://doi.org/10.1016/j.ipm.2019.03.004>.
- Zhou, X., & Zafarani, R. (2020). A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys*, 53(5), 1-40. <https://doi.org/10.1145/3395046>.