

6-20-2022

PM Me the Truth? The Conditional Effectiveness of Fact-Checks Across Social Media Sites

Austin Horng-En Wang
University of Nevada, Las Vegas, austin.wang@unlv.edu

Follow this and additional works at: https://digitalscholarship.unlv.edu/political_science_articles



Part of the [Communication Technology and New Media Commons](#), [Mass Communication Commons](#),
and the [Social Media Commons](#)

Repository Citation

Wang, A. H. (2022). PM Me the Truth? The Conditional Effectiveness of Fact-Checks Across Social Media Sites. *Social Media + Society*, 8(2), 1-15.

<http://dx.doi.org/10.1177/20563051221098347>

This Article is protected by copyright and/or related rights. It has been brought to you by Digital Scholarship@UNLV with permission from the rights-holder(s). You are free to use this Article in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself.

This Article has been accepted for inclusion in Political Science Faculty Publications by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact digitalscholarship@unlv.edu.

PM Me the Truth? The Conditional Effectiveness of Fact-Checks Across Social Media Sites

Social Media + Society
April-June 2022: 1–15
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20563051221098347
journals.sagepub.com/home/sms


Austin Horng-En Wang 

Abstract

People use multiple social media daily. Some platforms feature public interactions like Facebook, others emphasize private communications such as Line. Although misinformation is rampant on all platforms, literature on fact-checks (FC) focuses primarily on public ones. This article provides an integrated psychological model and argues that FC is less effective on private platforms. People expect to encounter “unwelcome” FCs (incongruent with their beliefs) on public platforms, but selectively approach the “welcome” FC on private platforms. An experiment ($n=601$) and a national survey ($n=1060$) were implemented to test these hypotheses in the 2020 Taiwan Presidential Election. The experiment shows that respondents prefer FC on Line, which helps their party, but prefer FC on Facebook which disadvantages their party. The survey shows that consuming FC with more private platform usage has lower media literacy, while is the opposite on public platforms. Future work should focus on both FC and how it is consumed.

Keywords

fact-check, private messaging apps, experimental design, Taiwan politics, media literacy

Introduction

In the era of misinformation and disinformation, significant efforts are made to combat the circulation of fake news on social media (e.g., Clayton et al., 2020; Lazer et al., 2018; Nekmat, 2020). Facebook, for example, has launched news guards and warning systems to crack down on rumors on social networks. The majority of US citizens also support the idea of fact-checking on social media sites (albeit with partisan bias, see Rich et al., 2020). Numerous social media sites, including Facebook, Twitter, YouTube, and Pinterest, also implement their fact-check (FC) policies, and the number of FCs increased by 900% within 3 months after the coronavirus disease 2019 outbreak (Brennen et al., 2020).

Meanwhile, social messaging apps also have become widely popular worldwide. They are designed in culturally specific ways to promote usage and take advantage of third-party workarounds to add extra layers of utility, creating new avenues of potential abuse in the process. However, a recent spate of news stories worldwide pointed to a different and possibly more urgent problem—the spread of misinformation on popular messaging services, such as WhatsApp, WeChat, and Line. Studies have found that rumors spread quickly via person-to-person interactions (Carlson, 2019).

Private messengers can even aggravate the spread of rumors between users due to their convenience and privacy in communications. The nature of private apps makes it difficult to curb fake news because these messages are protected by encryption.

The swirling false information on private messengers has led to serious consequences in countries where they are highly popular. In Mexico, two people were beaten and burned to death in August 2018 after rumors spread on WhatsApp about kidnapping children and harvesting organs. In India, a video edited to look like a kidnapping event that spread on WhatsApp inspired mobs to kill two dozen people.¹ In Taiwan, elections were overwhelmed by misinformation spread on the LINE app. Programmers who run a fact-checking bot on the app reported that at least 20% of the misinformation came from China.²

University of Nevada, Las Vegas, USA

Corresponding Author:

Austin Horng-En Wang, Department of Political Science, University of Nevada, Las Vegas, 4505 S. Maryland Pkwy. Las Vegas, NV 89154-9900, USA

Email: austin.wang@unlv.edu



Fact-checking is one of the many tools to respond to the misinformation on social media. However, most existing studies focus on the FC on the public social media sites like Facebook (e.g., Brandtzaeg et al., 2018), Twitter (e.g., Zhang et al., 2021), or did not specify (e.g., Freiling et al., 2021). We know very little about its effectiveness across different social media platforms. There is little to no literature investigating how fact-checking works in correcting news perceptions on private messengers.

Given the different interfaces and environment settings between the open and private social media platforms, we argue that the effectiveness of the fact-check, especially for the FC information conflicting with people's prior beliefs, would be conditional to the characteristic of the platforms. Specifically, we hypothesize that people prefer the "welcome" FCs (the fact-checking information that is consistent with one's political affiliation) in the private messaging apps but are fine with the unwelcome FCs in the open social media platform. This conditional FC effect is driven by people's expectations of social media sites. When people consume the FC information selectively in the private messaging apps, the biased consumption will hurt people's ability to discern the misinformation in the long run. Meanwhile, when people expected to encounter, and indeed received, the unwelcome FC information from the open forums which debunked their previous beliefs, their media literacy will rise by learning more unwelcome facts. We will elaborate more in the theory section.

We test our argument using one randomized survey experiment and one representative survey in Taiwan right before the 2020 Presidential election. Taiwan is a unique but generalizable case for studying the effectiveness of FCs across social media sites for three reasons. First, in recent years, Taiwan has been heavily inflicted by fake news attacks and is now ranked the top among all countries that are routinely influenced by false information from foreign countries (Mechkova et al., 2020). The severity of fake news infiltration raised the attention of the Taiwanese government, which started to sponsor programs to combat fake news on social media. Misinformation also became a salient issue right before the 2020 election.

Second, according to a telephone survey conducted in 2019, 85.6% of Taiwanese people regularly use the Internet. Furthermore, more than 90% of them use Facebook and Line, and there is no difference across generations. They also registered on other social media sites such as Instagram, YouTube, PTT, and Twitter.³ Third, right before the 2020 election, both Facebook and Line noticed the spread of misinformation and actively collaborated with the Taiwan government and civil groups to help design fact-checking chatbots to dissipate false information.⁴ The FC services appeared on both the public and private social media sites.

Hence, Taiwanese people use multiple social media sites on a daily basis, are familiar with the FC on both the private and public platforms, and are also seriously influenced by

misinformation. The wide usage of both platforms and false information plagued Taiwanese social media make the country a fruitful setting to test our hypotheses.

The Integrated Model of Information Seeking, FC, and Social Media Sites

Whether FC can change people's perceptions of false information remains debated. Some studies found that FC increases awareness of misinformation and corrects people's beliefs. Well-designed news alerts are found to be effective in countering false information on social media (Clayton et al., 2020; Fridkin et al., 2015). However, others found that it depends. The effectiveness of FC is related to the format of the message (Young et al., 2018), the context of the message (Garrett et al., 2013), the strength of the prior belief (Nyhan & Reifler, 2010), and the perception of the fact-checker (Nieminen & Rapeli, 2019).

However, before we examine the effectiveness of FC on people's media literacy, we should take one step back and ask: What motivates people to seek and consume the fact check information on social media sites in the first place?

One important feature of this social media era is that people control what kind of information they want to consume. When a piece of "unwelcome" FC information is presented which is inconsistent with an individual's prior political belief, they may not accept the information (Flynn et al., 2017) or simply ban, unlike, or unfollow the information source. When the incoming FC conflicts with one's existing belief, such a conflict may even cause the arousal of anxiety (e.g., Marcus & MacKuen, 1993) because one's worldview is challenged.

Nevertheless, the classic approach-avoidance approach in psychology renders opposite predictions on the effect of anxiety (Roth & Cohen, 1986): one may apply the avoidance strategy and avoid the source of threat, but one may also apply the approach strategy and ". . . noticing and taking advantage of changes in a situation that might make it more controllable." As a result, Valentino et al. (2008) show that the anxiety may increase both the quality and quantity of information seeking during the campaign when an individual's preferred candidate is under threat (i.e., losing in the election).

Meanwhile, following the approach-avoidance approach, positive emotions also usually link to information seeking. People tend to seek information that may bring pleasure to them, and such a tendency may be much more salient when people can use the Internet to choose their information source (e.g., De los Santos & Nabi, 2019). Similarly, both the in-group enthusiasm and the out-group anxiety play an important role in the political polarization and information seeking in the political domain (Lyons & Sokhey, 2014; McLaughlin et al., 2020).

Following the discussion above, this article argues that the level of *publicity* of different social media sites may moderate what kind of FCs people are willing to consume through the

platforms. If a social network site is designed for its users to communicate with others privately, such as WhatsApp, Line, or WeChat, people are likely to follow the pleasure principle and seek the information that brings a positive emotion. As a result, he or she may prefer the FC which is congruent with his or her prior belief in these private messaging platforms. They will keep in touch with the FC service on these private platforms only if the private platforms offer the fact which reaffirms their prior understanding of the world.

On the other hand, if the social network site is designed for its users to debate the political issues publicly and the debate may be easily observed by others, such as Facebook and Twitter, people would expect to meet the counterargument and unwelcome FCs when they start to use this platform every morning. As a result, people may make the best use of this environment and collect the unwelcome FCs through the public social media platforms. In the long run, this strategy may enhance one's quality of decision-making. For example, Lu and Lee (2019) find that Facebook usage positively correlates with encountering counter-attitudinal political information.

Even though the perception of publicity is hard to operationalize, previous studies suggested that people indeed behave differently on different social media platforms; the behavioral difference can be properly explained by the people's perceived publicity of these platforms. For example, Lottridge and Bentley (2018) collect 1,000 Facebook users' online behaviors and found that they are twice as likely to share political content publicly as a post than to share through private messaging to friends. In contrast, people are twice as likely to share funny content or local news with friends through private messaging than through public posts.

Besides, a focus-group study conducted by the Pew Research Center shows that young people already notice the difference between public posts and private messaging (Madden et al., 2013). They tend to share inside jokes with their close friends privately but not post to their friends publicly. They would rather not to upload many selfies on their online account to maintain a clean and neat profile; instead, they prefer to share selfies via emails.

Another study on Facebook also shows an interesting pattern: Wise et al. (2010) shows that people felt happier when they read their friends' Facebook profile, but they still spend most of their time on social browsing, that is, surfing the newsfeed. Last but not least, one study comparing Facebook and Pinterest usage in Taiwan shows that Facebook users were less happy than the Pinterest users, and people behave on the two platforms differently (Lin et al., 2017).

These four studies based on Facebook show a similar pattern: people are aware of the consequence of their public online behaviors and tend to pursue pleasure in private messaging apps. Hence, people may pay more attention to the information they could collect on Facebook instead of following the pleasure principle.

If people have a different tendency to expose themselves to the FCs on different social media sites, their source of information consumption may moderate the effectiveness of FC in the long run. Specifically, if one mainly relies on the public social media sites (SNS hereafter) to consume the FCs, since many of the FCs may be incongruent with his or her prior belief, he will be less biased and would have a higher level of media literacy in the long run. In contrast, if one mainly relies on the private messaging apps to consume the FCs, following the pleasure principle, he or she may select the FC source which is congruent with his or her prior belief; in the long run, he or she would have a lower level of media literacy because of the unbalanced and biased FC consumption.

Figure 1 summarizes the theoretical framework and the empirical strategy of this article. The theory starts with the assumption that people perceive different levels of publicity on different social network platforms. Since the characteristic of the SNSs is externally given, people adjust their information-seeking behavior and expectations when using different SNSs. This adjustment influences people's interaction with the fact-checking services provided on these platforms and, therefore, influences the effectiveness of fact-checking. Two hypotheses are, therefore, suggested and will be examined in this article:

Hypothesis 1 (H1): People prefer the FC, which aligns with their political preference on private social media platforms than the public ones, and vice versa.

Hypothesis 2 (H2): Selective exposure of FCs in private messaging apps lowers people's level of media literacy, and vice versa.

In H1, the two independent variables are the publicity of the SNSs and whether the FC is welcome or not, and the key dependent variable is the information-seeking behavior after receiving the FC. Hence, a survey experiment is designed in Study 1 to manipulate the treatments while controlling other covariates. In H2, the key variable is the interaction between the FC consumption and the usage of the social media sites. Since H2 is established on the result of Study 1, Study 2 will exploit a national survey to examine the external validity of Study 1.

It is worth noticing that this integrated model focuses on the difference between social media sites. Hence, it is a supply-side approach to investigate how the features of different platforms may systematically influence how users consume FCs. In comparison, recent literature also focuses on the demand-side approach of FCs, indicating that users themselves may have different proposes for using the social network sites, which influences their willingness to consume FCs (e.g., Celik, 2020; Celik et al., 2021). The model and the studies in this article focus on the supply side to bridge the gap in the literature.

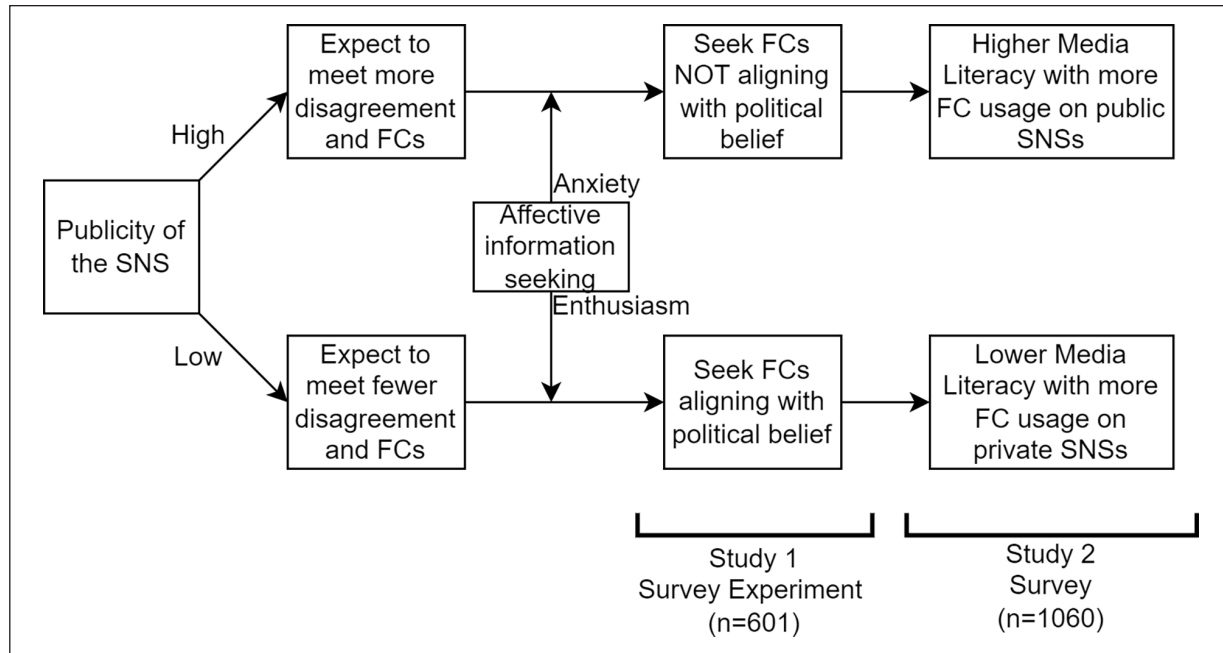


Figure 1. The integrated model linking publicity of SNSs, FC, and media literacy. FC = fact-checks; SNS = social media sites.

To our knowledge, Rossini et al. (2021) is the first study comparing the impact of misinformation and correction across different social media sites. Through a nationally representative (and mainly) Internet survey, they find an interestingly positive correlation between social correction and sharing misinformation on Facebook, but not on WhatsApp. To be specific, in their research question 4, their result shows that “Having witnessed someone else being corrected is also positively associated with having shared misinformation on Facebook, but not on WhatsApp.”

However, Rossini and colleagues’ article suffers from three methodological issues. First, they rely on self-reported misinformation sharing. However, those who suffered from the misinformation the most may not know they were sharing the misinformation; it is the respondent who had been corrected before that knows he or she had shared the misinformation before. Second, sharing misinformation is a rare phenomenon (Guess et al., 2019). Therefore, it may not be a good measure to examine the effect of fact-checking by self-reports. Third, the study relies on a cross-sectional survey, so the findings are at most correlational.

The major methodological challenge in examining the two hypotheses and Figure 1 is manipulating the publicity of social media sites in Study 1. Since people would have already relied on different information and FC sources on different social media sites based on the characteristic of SNSs, a cross-sectional study cannot mitigate the problem of endogeneity in examining the first hypothesis.

Hence, I exploit the survey experimental design to simplify the treatment (private and public platforms) while controlling for other covariates, including the length of the message or other unrelated elements on the platforms. The survey experimental design uses the social network sites’ screenshots while editing the content to serve as the main treatments. In our case, the treatments are whether the misinformation and the FCs are sent to the respondent through a frame of Facebook or a frame of Line (see Study 1 for the details). Recent studies use similar designs such as Facebook or Twitter (Mena, 2020; Tully et al., 2020) to serve as the treatments. The other treatment is whether the FC is welcome or unwelcome. I exploit the highly-polarized Taiwan politics to create the FCs which would politically advantage or disadvantage a Taiwanese voter’s preferred or disliked party.

To improve Rossini and colleagues’ measurement of media literacy, reliable measurement of the respondent’s capacity to discern the misinformation is necessary. Hence, in Study 2, I followed Guess et al. (2020)’s guidelines to create a fake news battery for the 2020 Taiwan Presidential Election. The fake news battery is composed of facts and misinformation that the non-profit FC organization in Taiwan had verified before the survey started. The battery enables us to estimate the respondent’s media literacy directly. Therefore, we can use this estimation to examine whether FC services on private and public SNS platforms would influence people’s media literacy, especially when

misinformation was a salient issue right before the 2020 presidential election.

Study 1: People Prefer to Welcome FCs on Line but Unwelcome FCs on Facebook

The first study is designed to examine whether people respond to FCs differently on different SNSs. Since it is usually not legal to send a personalized message in the private messaging apps, I follow Tully et al. (2020) and Mena (2020) to create simulated Facebook and Line posts and FCs in the survey questionnaire. The simulated posts are designed to imitate the SNS posts that the respondents encounter in their daily life so that the measurements in the survey can capture how people respond to the posts in real life.

There are two treatments in this experiment for examining H1. The first is whether the FC is aligned with the respondent's political preference, while the second is in which SNS platform the respondent receives the FCs.

Research Design of Study 1

The survey experiment was conducted on 19–21 April 2019. Overall, 601 subjects were recruited by PollcracyLab at National Chengchi University, a top university in Taiwan, through a random selection from its pool (PL hereafter). PL builds and maintains the subject frame based on the Taiwan government's household registration records. Because PL is established under the top research university in Taiwan, it can access official household registration records for academic purposes. Therefore, all Taiwanese citizens have a non-zero probability of being invited for registration by PL, a crucial foundation for establishing the representativeness of any sampling procedure. Compared with other opt-in platforms such as Amazon MTurk or Survey Sampling International, PL recruits Taiwanese subjects from diverse and more representative backgrounds. Since all PL respondents are recruited through official household registration records, it may avoid the problem of fake accounts or robot accounts in other opted-in platforms (e.g., Storozuk et al., 2020).

The whole survey experimental design is pre-registered before PL's survey implementation. This survey is sponsored by the author's institution. It passed the IRB from the IRB committee at the University of Nevada, Las Vegas (number 1374707-2) and was pre-registered before implementation (<https://osf.io/trz9b/>). The socio-demographic background of the subjects can be found in Online Appendix Table A1. Male and 30–39 respondents were slightly overrepresented in this study.

In the invitation letter, subjects were asked to participate in a survey titled “*Survey of Public Opinion and Political*

Participation.” Subjects were informed that they could skip any question, and their answers will be kept anonymous. The anonymity of the subjects was ensured because PL was responsible for both recruiting respondents and sending the gift cards as compensation, and all identifiable information was cleaned before the dataset was sent to the researchers. All subjects received an NTD \$100 gift card (about US\$3) after completing this survey.

The questionnaire includes 40 items. Subjects were first asked to report their news consumption, their level of political interest, and the number of Line messages they received daily. Survey subjects were then randomly assigned to one of the 2 (Kuomintang [KMT] or Democratic Progressive Party [DPP] scandal) \times 2 (Line or Facebook) groups. Taiwan features a two-party system where KMT and DPP dominate the major political cleavage—unification versus independence (Achen & Wang, 2017; Wang, 2019). In each group, they were asked to imagine that they were reading a message from “one of your elementary school classmates,” and the post was simulated that it was sent through a Line chatroom or shared on the Facebook newsfeed. The post shows a news link describing—with a picture, a headline, and a short sentence to simulate the online environment—a previously unreported sex scandal involving several city councilors belonging to either KMT or DPP. Two sample messages are shown in Figure 2 (the translation is on the right side of Figure 2 and was not posted to the Taiwanese respondents). After reading the news message, each subject was asked whether they trusted the news or not (from 0 to 100).

After the message, all participants were then asked to read a message from an FC account that disapproves of the previous scandal they just read. The FC account is called “Internet News Helper” with a cute profile picture. The FC message was posted on the same platform as the fake news treatment. Two examples are shown in Figure 3. This short rebuttal format follows Ecker et al. (2020) which has been found effective in fact-checking.

The design of FC messages imitates a popular FC service offered by Cofacts. Cofacts is a non-profit organization, which builds FC chatbots on Line. A Line user can add the Cofacts chatbot as a friend and then forward any dubious message to the chatbot. One can also invite the chatbot into the chatroom and the chatbot will screen the content of discussions automatically. The chatbot will then search its database to see if any content sent to it or appeared in the chatroom has been debunked by other third-party FC organizations. The chatbot will then share the FC with the users through the chatroom. The details of the Cofacts chatbot can be found in Online Appendix Figure A1.

All participants then answered a series of follow-up questions. They were asked (a) How much they still believe in the original scandal (0–100)? (b) Whether they will add the Internet News Helper to their friend list on social media? and



Figure 2. Survey experiment: a simulated rumor against city councilors on social media (top: Line and KMT [國民黨] version. Bottom: Facebook and DPP (民進黨) version).

KMT = Kuomintang; DPP = Democratic Progressive Party.

(c) Whether they will seek more information about the original fake news?⁵ In the end, they were debriefed that the scandal and the fact-checking accounts are both designed by the experimenters.

It is worth noticing that this survey did not ask about the respondent's socio-demographic background or partisanship. Instead, these personal variables were obtained when PL recruited the respondents in the previous surveys. Hence, their socio-demographic background or political preference would not be framed before or after the treatments (Klar et al., 2020).

All 601 subjects completed the survey (100%). The zero-dropout rate implies that the length of the questionnaire did not create a substantially cognitive burden on the subjects, so

they could concentrate on the questions and the treatments and were not distracted. Moreover, the workers at PL informed us that they did not receive any complaints during this survey implementation. This report enhances our confidence in the validity of the experimental design. However, owing to budget and time limitations, all questions were only asked once. Hence, it would not be able to estimate the reliability of the key dependent variables and leave room for a larger variance.

Operationalization of Study 1

This experiment captures the respondents' responses to FCs by two behavioral variables. The first outcome is whether

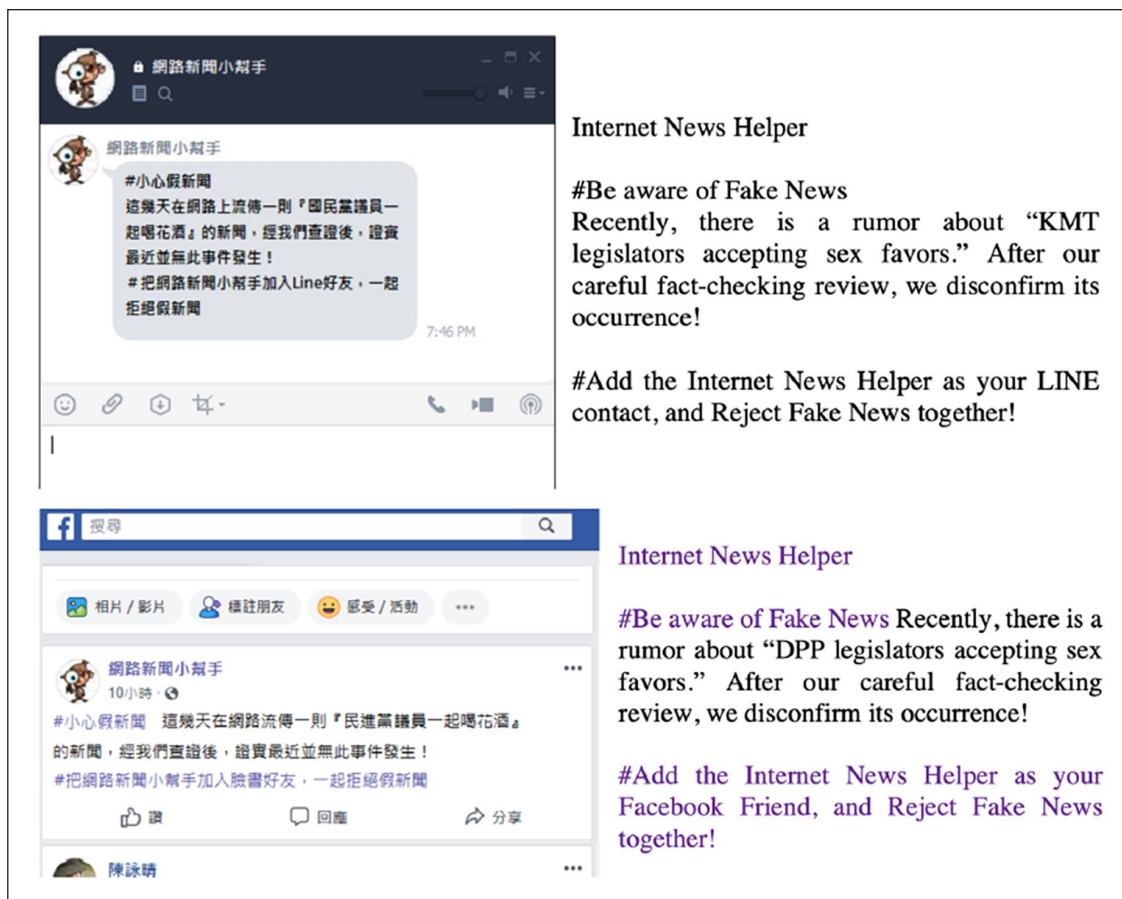


Figure 3. Survey experiment: rumor correction from a simulated news helper on social media (top: Line and KMT version. Bottom: Facebook and DPP version.)

KMT = Kuomintang; DPP = Democratic Progressive Party.

people plan to add this FC account Internet News Helper as their SNS friend (FRIEND). The second outcome is whether people will seek more information about the scandal (INFO). The respondent's response is coded as 1 if he or she answered "yes" or "probably," and 0 if "probably not" or "no" in both items.

Both outcome variables are designed to measure whether the respondent plans to seek more information in the future. If one tends to add the FC account as his or her SNS friend, it implies that they expect to receive other updates from the same FC account in the future, and expect to have more interactions with this FC account. Similarly, if the respondent plans to seek more information about the scandal after the face-check, he or she literally has a higher level of information-seeking tendency.

There are two treatments in this experiment. The first is the publicity of the SNS platforms. All respondents were randomly assigned to receive the fake news message and the follow-up FC through either the frame of Facebook or Line, as indicated in Figures 2 and 3.

The second treatment is whether the FC information is aligned with the respondent's preference or not. Given the

increasing level of political polarization in Taiwan (Wang, 2019), it is reasonable to assume that partisans prefer to see their preferred party's scandal being debunked. Hence, the treatment "WELCOME FC" is defined as a DPP voter reading a DPP's scandal being fact-checked, or a KMT voter reading a KMT's scandal being fact-checked. On the contrary, an FC is "unwelcome" if a DPP supporter reads an FC of KMT councilors' scandal or a KMT supporter reads a DPP's scandal being debunked. Since there is no theoretical guidance for the non-partisans and supporters of the small parties, they were dropped from this analysis.

After the operationalization, there are 89 respondents in [Welcome FC+Line] group, 83 in [Welcome FC+Facebook] group, 83 in [Unwelcome FC+Line] group, and 86 in [Unwelcome FC+Facebook] group.

Randomization and Manipulation Check of Study I

One-way ANOVA test and chi-square test shows that there is no significant difference between the four experimental groups in aspect of their age ($p=.968$), level of education ($p=.309$), gender ($p=.910$), ethnicity ($p=.207$), and national

Table 1. Cross-Table Explaining the Information-Seeking After FC.

	<i>n</i>	DV: Add FC as SNS friend = 1	DV: Seek more info about the scandal = 1
Line+Welcome FC	89	41.6%	43.8%
Line+Unwelcome FC	83	31.3%	34.9%
FB+Welcome FC	83	27.8%	31.3%
FB+Unwelcome FC	86	41.9%	51.2%

FC = fact-check; FB = Facebook; DV = dependent variable; SNS = social media sites.

identity ($p = .572$). The randomization check suggests that we can compare the mean value across the four groups directly.

The manipulation check examines whether the participants received the fake news and the FC, and whether their attitudes follow the theoretical assumptions made in this article. The first manipulation check is the effectiveness of the FC information. The average level of mistrust toward the scandal before FC is 55.72, while the mean of mistrust after FC is 66.33 (higher mistrust indicates that he or she did not believe in the fake news, which is the aim of FC). Paired t -test shows a significant increase in the level of mistrust toward the scandal ($p < .001$). In short, the political scandal provides a moderate level of mistrust to the respondents (given its short format and lack of details), while the FC can still successfully increase the readers' the level of mistrust, indicating that both the fake news message as well as the FC work in this experiment.

The second manipulation check is whether the Welcome/Unwelcome FC would induce partisan bias. When respondents read about the scandal, their level of mistrust of the scandal is 55.0 in the Unwelcome FC group and 59.7 in the Welcome FC group; the t -test shows that the difference is not significant ($p = .12$). After the FC, the average level of mistrust in the Unwelcome FC group is 63.7, while in the Welcome FC group is 69.01; the t -test suggests a significant difference at the $p = .1$ level ($p = .068$). Therefore, the FC that advantages the respondent's party is indeed more "welcome" by the respondents; it causes a higher increase of mistrust toward the original scandal. In contrast, the effect of the unwelcome FC is smaller than the welcome ones. Hence, the treatments indeed induce a certain level of partisan bias in the respondents, which is the goal of experimental design. It is worth noticing that the FC works for both welcome and unwelcome contexts in our experiment, possibly because the scandal is fabricated but not a well-known one (e.g., Nyhan & Reifler, 2010).

Result of Study 1

Table 1 shows the respondents' information-seeking tendency after reading the FC in the four experimental groups. Both dependent variables illustrate a similar pattern: When the FC was sent through Line, the respondents have a higher level of information seeking tendency only if the FC aligned with their political belief; meanwhile, the pattern is opposite if the FC was sent through Facebook.

Table 2 applies logit regression analysis to control covariates. Models 1 and 4 only include the two treatments and the interaction, Models 2 and 5 add the level of mistrust (0–100) to the fake news measured before and after the FC, and Models (3) and (6) include gender, sex, and the level of education. Among all six models, the interaction of the two treatments is statistically significant ($p < .05$). The result indicates that the two treatments jointly influence the respondent's information-seeking tendency after receiving the FC. The result holds even when the mistrust of the fake news as well as a series of personal background variables are controlled. In the end, the last row in Table 2 shows the maximum variance inflation factor for each regression. The values are all around 3, which suggests a modest level of collinearity given the interaction terms in the regression model (Brambor et al., 2006).

Figure 4 further shows the simulated likelihood of seeking more information after reading the FC, controlling all other covariates at the mean value. Basically, Figure 3 shows the same pattern as Table 1, suggesting that the effect of the welcome/unwelcome FC is opposite on different SNS platforms. When the FC is sent through Facebook, people are likely to add the FC account as a friend on Facebook if the FC sent unwelcome FC, but the trend reverses on Line.

Among the control variables, it is worth noticing that the pre-FC and post-FC mistrust have an opposite impact on the tendency of information seeking. If one did not believe in the fake news in the first place (high pre-FC mistrust), they might think that the FC did not provide additional information to their knowledge. Since they already knew what they need to know, no more information seeking is needed. In contrast, if one did believe in the fake news (low pre-FC mistrust), the FC may debunk their belief to find the FC valuable. Hence, it is reasonable to see the negative correlation between pre-FC mistrust and information-seeking tendency. The tendency reverses for post-FC mistrust. If one had a high post-FC mistrust, it means that their opinion is consistent with the FC information, or the FC updated his or her belief.

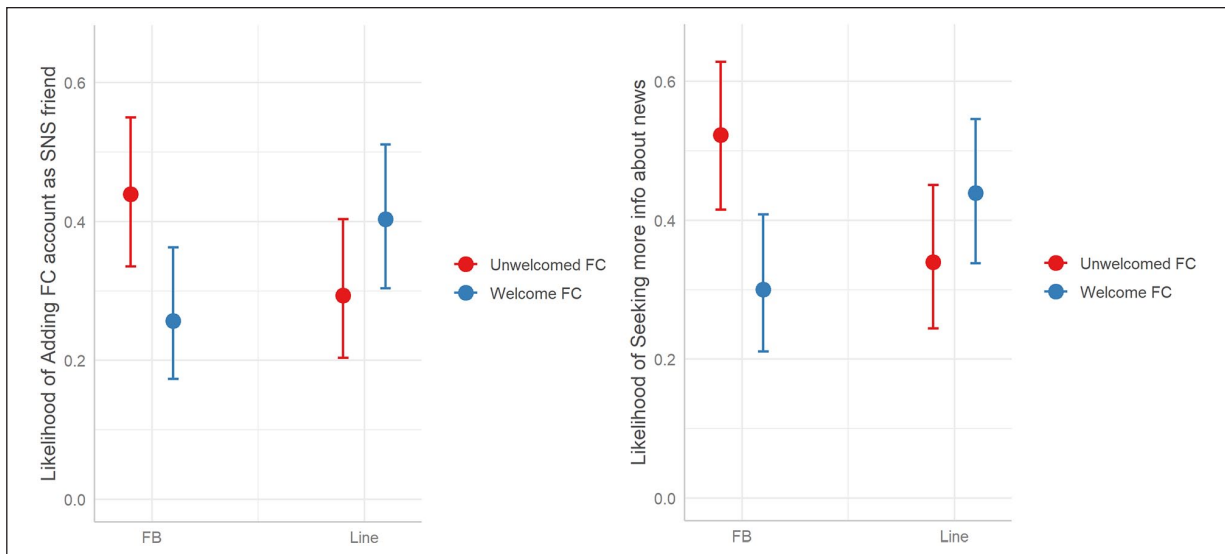
Discussion of Study 1

In Study 1, it is important to reiterate that the survey experiment had controlled all other covariates. The content of the fake news and the FC content are exactly the same on both

Table 2. Logit Regression Explaining the Information-Seeking After FC.

	DV: Add FC as SNS friend = 1			DV: Seek more info about the scandal = 1		
	(1)	(2)	(3)	(4)	(5)	(6)
Welcome FC = 1	-0.63 (0.33)	-0.81* (0.34)	-0.82* (0.34)	-0.83** (0.32)	-0.93** (0.33)	-0.94** (0.33)
Line = 1	-0.46 (0.32)	-0.61 (0.34)	-0.64 (0.34)	-0.67* (0.32)	-0.78* (0.32)	-0.76* (0.33)
Welcome FC × Line	1.08* (0.46)	1.29** (0.48)	1.31** (0.48)	1.20** (0.45)	1.37** (0.46)	1.36** (0.46)
Pre-FC mistrust		-0.02* (0.01)	-0.01* (0.005)		-0.02** (0.01)	-0.02** (0.01)
Post-FC mistrust		0.03** (0.01)	0.03** (0.01)		0.02** (0.1)	0.02** (0.01)
Age			-0.05 (0.11)			0.06 (0.11)
Level of Education			0.06 (0.20)			-0.11 (0.19)
Gender			0.26 (0.24)			-0.13 (0.23)
Intercept	-0.33 (0.22)	-1.22** (0.38)	-1.74 (1.23)	0.05 (0.22)	-0.14 (0.35)	0.40 (1.18)
N	341	341	341	341	341	341
AIC	466.94	432.8	437.2	459.8	452.5	457.3
Max(VIF)	3.27	3.34	3.36	3.20	3.27	3.27

FC = fact-checks; DV = dependent variable; SNS = social media sites; AIC = akaike information criterion; VIF = variance inflation factor.
* $p < .05$, ** $p < 0.01$.

**Figure 4.** Simulated likelihood of information seeking from Models 3 and 6 in Table 2.

simulated Facebook and Line treatments. The only difference that drives the difference in information-seeking behavior is the frame of Facebook and Line. The result in Study 1 shows how Taiwanese people behaved differently when they received the welcome/unwelcome FC from different platforms.

Study 2: FC Enhances the Media Literacy on Facebook but Worsens on Line

Since the behavioral pattern observed in Study 1 is robust and offers some level of internal validity, H2 can be further

examined that FC may cause opposite effects for people who rely on different SNSs for FC consumption. Specifically, if one heavily relies on the private messaging apps to consume the FC, following the tendency observed in Study 1, he or she may be more biased toward “welcome” FC in the long run; hence, he or she will suffer from a lower level of media literacy. In contrast, if one frequently uses the open social media platform to consume FCs, he or she can observe how these FCs were being scrutinized publicly, and some of the FCs clearly contradicted his political belief. Hopefully, he or she would have a higher level of media literacy ultimately because of this diversified information source.

The research design in Study 2 is straightforward. First, we create a Taiwanese media literacy scale including eight verified fake and real news. Second, Taiwanese respondents ($n=1,060$) were asked about their usage of the FC source, choice of information consumption, and took the media literacy scale. Third, the relation between FC source, information source, and their performance in the media literacy scale is examined. Study 1 focuses on internal validity, while Study 2 examines external validity.

Taiwan Media Literacy Scale 2020

We follow Guess et al.’s (2020) guidelines to create an eight-item fake news battery designed for the 2020 Taiwanese Presidential Election held on 11 January 2020. The detailed battery is reported in Online Appendix Table A2. Each item is composed of 60 to 74 Chinese characters. Items 1 and 2 are health-related and non-politics, 3 and 4 are about China, 5 and 6 are about the 2020 KMT presidential candidate Han Guo-yu, and 7 and 8 are about the 2020 DPP presidential candidate Tsai Ing-wen. All news pieces had been fact-checked by at least one of the major fact-checking sources in Taiwan before 24 September 2019.

Research Design of Study 2

Overall, 1,060 Taiwanese respondents were recruited by PL. The survey was conducted from 4 to 7 January 2020—1 week before Election Day. All subjects received NTD \$50 (about US\$1.3) after completing this survey with 40 items. This survey is sponsored by the author’s institution and American Psychological Association. It passed the IRB from the IRB committee at the University of Nevada, Las Vegas (number 1504367-1) and was pre-registered before implementation (<https://osf.io/8vumx/>).

The socio-demographic background of the subjects is similar to Study 1 and is presented in Table A3 in the Online Appendix. Compared with the population of Taiwan, the sample is younger, fewer non-partisans, and more males. Even though all Taiwanese people have a non-zero probability of being recruited by PL, those who are frequent Internet users, tend to answer the online surveys, and those who are interested in politics are more likely to accept our invitation.

The distribution is also similar to Online Appendix Table A1 in Study 1. The only major difference is the appearance of the Taiwan People’s Party, a new minor party in Taiwan, which was established before the 2020 election.

In this survey, all respondents were first asked about their media consumption across different platforms. They were asked how much time they spent consuming political information from newspapers, radio, television, news websites, Facebook, Line, and YouTube. The distribution of the respondents’ media consumption is in Online Appendix Table A4. More than 60% of the respondents use Facebook, Line, and YouTube to consume political information during the campaign. The percentage is higher than the newspaper but lower than the television and news website.

All respondents were then asked, “Have you ever used any Fact-checking website, such as Mygopen, Cofacts, or Taiwan Fact-check Center?” (yes=1, no=0). These three websites are the major fact-checking source during the 2020 Election, which officially cooperates and is endorsed by Facebook and Line.⁶ These FC sources were also widely shared on Facebook and on Line, including the Cofacts chatbot mentioned in Study 1. Overall, 27.3% of Taiwanese respondents in the survey used fact-checking sources before the presidential election (12% more than three times, and 15.3% one to two times). The percentage is close to Guess et al.’s (2018) study in the United States (25.3%) during the 2016 election.

All subjects then took the 8-item Taiwan Media Literacy Scale (MLS2020). They were asked to judge whether each message is fake news or not. The order of the eight questions was randomized. On average, respondents got 5.69 correct in MLS2020. The correct answer percentage is significantly higher than 4 (one-tail t -test $p<.001$), indicating that the respondents do not randomly guess all the items. The distribution of the number of correct answers is shown in Online Appendix Figure A2. The distribution is single-peaked and very few of them got all correct or all wrong. Therefore, the MLS2020 scale did not underestimate the respondents’ ability owing to the upper limit. The distribution in Figure A2, albeit not a standard normal distribution, is still wide enough to capture the different levels of media literacy among Taiwanese respondents. Hence, we sum up the number of correct answers as each respondent’s level of media literacy level during the 2020 Taiwan Presidential Election.

After MLS2020, subjects were also asked to answer other questions beyond the scope of this article. Once again, their social-demographic background was recorded before this survey and is provided by PL directly.

Result of Study 2

Before we examine whether the FC has opposite effects on different SNS, it would be better to examine the effectiveness of the FC measure in the first place. T -test shows a significant difference in the number of correct answers between

Table 3. OLS Explaining the Level of Media Literacy ($n = 1,060$).

	DV: Number of correct answers in MLS2020 (0–8)			
	(1)	(2)	(3)	(4)
Facebook usage (1–5)	0.012*** (0.003)	0.053 (0.050)	–0.0008 (0.050)	–0.007 (0.053)
Line usage (1–5)	–0.185** (0.046)	–0.048 (0.058)	–0.011 (0.057)	–0.013 (0.059)
FC usage (0/1)	0.112 (0.080)	0.037 (0.024)	0.209 (0.236)	0.194 (0.238)
Facebook × FC Usage		0.182* (0.086)	0.183* (0.084)	0.195* (0.084)
Line × FC Usage		–0.364*** (0.095)	–0.323*** (0.094)	–0.323*** (0.095)
Age, gender, sex, party			YES	YES
Newspaper, radio, television, news websites, YouTube				YES
Intercept	5.741*** (0.121)	5.632*** (0.121)	5.595*** (0.341)	5.558*** (0.352)
<i>n</i>	1,060	1,060	1,060	1,060
Adjusted R^2	.019	.031	.065	.063

FC = fact-check; MLS = Media Literacy Scale; OLS = ordinary least square; DV = dependent variable.

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

the Taiwanese respondents who experienced FC service and those who did not (5.81–5.65, $p = .046$). Hence, it is reasonable to assume that MLS2020 and FC experience may capture the respondent's behavior and media literacy during the campaign.

Table 3 shows four ordinary least square (OLS) regression models explaining Taiwanese respondents' level of media literacy before the election. Model 1 includes Facebook usage, Line usage, and FC source usage before the campaign, while Model 2 includes the interaction between the FC usage and the two SNS platforms. Model 3 further includes the socio-demographic background of the respondents and their partisanship, and Model 4 includes all other information sources which were also asked in the survey. The full regression table, including the coefficients of covariates, can be found in Online Appendix Table A5.

In Model 1 in Table 3, it seems that Taiwanese respondents who used Facebook to consume political information have a higher level of media literacy, while those who used Line have a lower level of media literacy; the FC is not effective in this model. However, Model 2 clearly shows that the effect between FC and media literacy is actually moderated by the SNS platforms: If people use Facebook more, then the FC increases people's media literacy, but the effect turns to be negative if people use Line more. The main effect for Facebook, Line, and FC are all insignificant in Models 2 to 4, indicating that their relationship to media literacy cannot be explained without considering the interaction term. The interaction terms are statistically significant and robust in Models 2 to 4 after controlling for a series of covariates. The result in Table 3 supports H2.

To further illustrate the interactive effect between the SNS and FC on the level of media literacy, Figure 5 shows the simulation from Model 4 in Table 3, controlling all other covariates at their means. When an average respondent did not use the fact-checking source during the campaign, Model 4 in Table 3 suggests that the SNS platform will not influence his or her level of media literacy. However, when he or she used the FC source, more Facebook usage correlates with a higher level of media literacy, while more Line usage correlates with a lower level of media literacy. When a respondent used Line to consume political information for more than 120 min a day, using FC would drop his media literacy from 5.7 to 4.6 on the 0 to 8 MTS2020 scale. Figure 5 illustrates the key argument in H2.

Among the control variables (in Online Appendix Table A5), age negatively correlates with media literacy, while the level of education is positive. Previous studies in the US context also found that older people consume fake news more (Guess et al., 2018, 2020), while a higher level of education correlates with less fake news sharing in the Brazilian case (Rossini et al. 2021). Besides, all other sources of information do not have an impact on the respondents' level of media literacy, indicating that the conditional effect of FC may be a unique phenomenon on the new social media platforms. In the end, all part dummies are insignificant except for the KMT. The negative correlation indicates that the MLS2020 may bias against the KMT supporters in the 2020 election, or KMT supporters have a lower level of media literacy during the campaign than the nonpartisans. Nevertheless, the conditional effect between the SNS and FC is still robust after controlling for partisanship in Models 3 and 4 in Table 3.

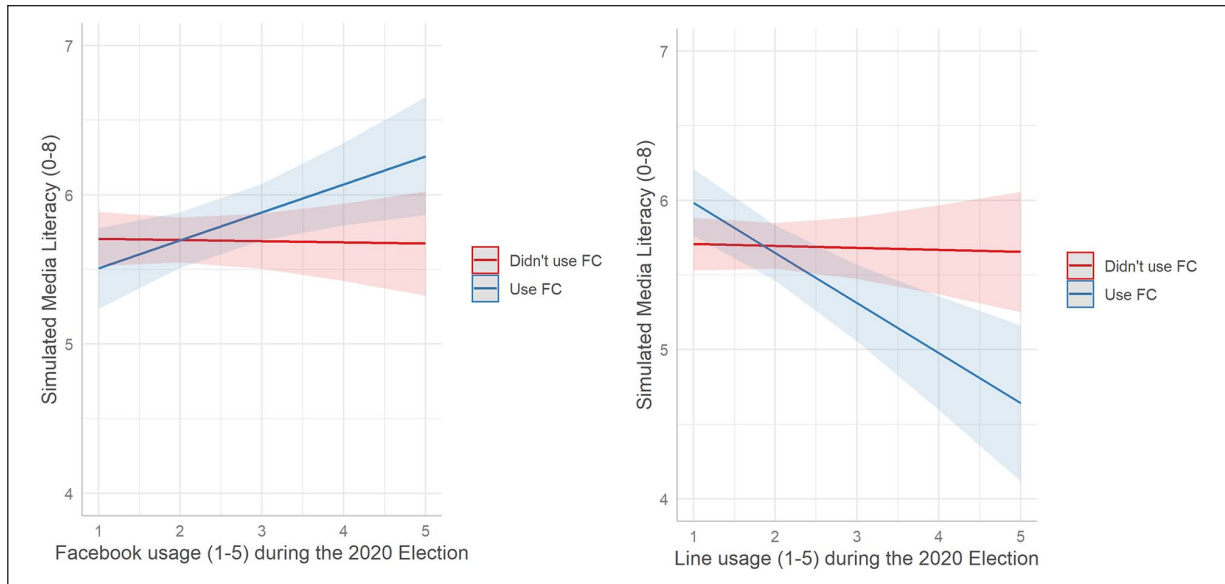


Figure 5. Simulated media literacy by Facebook/Line and FC usage from Model 4 in Table 3. FC = fact-check.

Table 4. OLS Explaining the Level of Media Literacy by Issue Domain ($n = 1060$).

	DV: Number of correct answers for health-related fake news (0–2)	DV: Number of correct answers for Taiwan politics fake news (0–4)
	(1)	(2)
Facebook usage (1–5)	–0.011 (0.025)	–0.0008 (0.050)
Line usage (1–5)	0.015 (0.027)	–0.011 (0.057)
FC usage (0/1)	–0.008 (0.011)	0.209 (0.236)
Facebook × FC usage	0.061 (0.040)	0.153* (0.064)
Line × FC usage	–0.086 (0.045)	–0.163* (0.070)
Age, gender, sex, party	Yes	Yes
Newspaper, radio, television, news websites, YouTube	Yes	Yes
Intercept	1.444** (0.165)	2.692** (0.261)
n	1,060	1,060
Adjusted R^2	.034	.115

FC = fact-check; OLS = ordinary least square; DV = dependent variable.
* $p < .05$, ** $p < .01$.

One may question whether the heterogenous FC effect on SNSs is driven by partisan bias or not, since the media literacy scale includes both political and non-political items. In Table 4, I separate the dependent variable into non-political and heal-related fake news (Q1 and Q2) and politics-related fake news (Q5, Q6, Q7, and Q8; see Online Appendix Table A2). Regression analysis shows that the same conditional effect is still significant in explaining the media literacy of politics-related fake news but not the health-related fake

news. The adjusted R-squared is also higher in Model 2 in Table 4 than any other model in Tables 3 and 4. Hence, we have more confidence that the pattern we observed in Study 2 is mainly driven by political motivations, which is also reflected in Study 1.

Overall, the cross-sectional result in Study 2 suggests the generalizability of Study 1. FC is not a panacea. The effectiveness of FCs depends on how people consume them. When people consume the FC through private messaging

apps, they tend to receive the FC congruent with their prior beliefs. In the long run, receiving more FCs through private messaging apps lowers the level of media literacy among Taiwanese respondents, especially in the political domain.

Conclusion and Discussion

This article provides an integrated model to explain the psychological mechanism of how the publicity of social media platforms may influence the consumption of FC (in Figure 1), and then provides two empirical studies to support the hypotheses derived from the model. Study 1 conducts a survey experiment to overcome the methodological challenge of endogeneity, and shows that Taiwanese respondents accept unwelcome FC on Facebook but welcome FC on Line. Study 2 further examines the external validity of Study 1 by a representative survey in Taiwan. The result shows an interactive effect between the FC and platform usages, reaffirming the main hypotheses in the model and the findings in Study 1.

The results of this article contribute to the existing literature on misinformation and fact-checking in several important ways. First, our findings provide a mixed view of fact-checking's effectiveness among the debate on its usefulness, showing that interaction with fact-checkers in the open platform helps increase the reader's ability to discern fake news and gain media literacy. Meanwhile, FC hampers the reader's media literacy through the private messaging apps.

Second, this study expands on prior literature that focused narrowly on open platforms on social media by demonstrating how fact-checking changes perceptions of information shared on private platforms. Using the novel survey experiment and the representative survey, this article specifies conditions and mechanisms through which corrective information is more welcomed and helps cultivate media literacy.

Third, the findings point to the potential hurdle in fighting political misinformation when corrections or alerts are unwelcome by citizens who hold a strong predisposition. It may suggest a scope condition of fact-checking effectiveness in terms of the platforms and the reviewed content. The effect of FC may depend on how the FC is spread. If people rely more on private messaging apps, the FC may only strengthen people's prior beliefs instead of shaking them.

One limitation of this article is that we did not examine, but only summarize, the role of emotions in information seeking. We theorize the information-seeking behavior in two different scenarios (SNSs with a higher/lower publicity) based on their expectations of the SNS, but the emotions steering these behaviors are not examined. Following Lyons and Sokhey's (2014) study in the US context, we expect that the partisan bias on the private messaging apps may also be driven by in-group enthusiasm, while the balanced information seeking on the public SNSs may be influenced by the out-group anxiety/fear. Future work may verify the role of emotion in FC consumption.

Another limitation of this article is that the experiment in Study 1 is not conducted directly on social media sites. The experimental design only uses the image to intimate how people use social media. The major benefit of the experimental design is to control other covariates such as other unrelated information which may distract the respondents, but the major drawback is that people may think that the treatment is not real enough. The manipulation check showed that people indeed have politicized perceptions toward the fake news message and improved their beliefs after reading the fact check. Hence, we believe that the respondents in Study 1 should have a certain level of desired responsiveness.

Besides, this article manipulates the level of publicity through Line and Facebook, but it may not capture how respondents proposed to use these platforms themselves (Celik et al., 2021). This article assumes the publicity is mainly decided by the design of the software, and both studies reveal the average effect of the platforms. In real life, the line between the public platforms and private messengers may be blurred. For example, many private messenger users may use the group features and interact with many people they do not know. Future work may focus on the interaction between the characteristics of the platform and its users' perceptions.

In the end, the conditional FC effect reveals some potential directions for future research. First, when people keep diversifying the channels of information consumption, the study of misinformation and FC should also extend to multiple platforms. When the young generation starts to mobilize through Instagram and TikTok, warning messages or red flags for combating misinformation may be no longer effective on these platforms with distinct interfaces and expected interactions. Second, the FC study should emphasize how people consume it, not only how the FCs are provided. The result of Study 2 in this article shows that FCs may be counterproductive when people use them selectively.

Acknowledgements

I appreciate the helpful comments from the two reviewers and the efficient review process facilitated by the editors. I also appreciate Dr. Howard Liu for the constructive comments on the literature review and research method. The early version of this project was presented at the 2019 American Political Science Association Annual Conference and received helpful comments from the discussants and panelists.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Both Study 1 and Study 2 are supported by the author's research grant in the Department of Political Science at the University of Nevada,

Las Vegas. Study 2 is also supported by the American Psychological Association Division 48 Small Grants Award (\$2000). The publication fees for this article were supported by the UNLV University Libraries Open Article Fund.

ORCID iD

Austin Horng-En Wang  <https://orcid.org/0000-0002-0256-2723>

Supplemental Material

Supplemental material for this article is available online.

Notes

1. <https://www.nbcnews.com/tech/tech-news/how-whatsapp-became-linked-mob-violence-fake-news-why-it-n929981> (accessed on 10 June 2021).
2. <https://www.stopfake.org/en/made-in-china-fake-news-overwhelms-taiwan/> (accessed on 10 June 2021).
3. https://report.twinc.tw/2019/assets/download/TWNIC_TaiwanInternetReport_2019_CH.pdf (accessed on 8 April 2020).
4. <https://www.inside.com.tw/article/16699-fake-news-self-regulatory> (accessed on 8 April 2020).
5. After this experiment, all subjects then went through other experiments that are beyond the scope of this article. This experiment was conducted first so there would be no spoiler problem.
6. <https://tfc-taiwan.org.tw/articles/1231> (accessed on 19 April 2020).

References

- Achen, C., & Wang, T. Y. (2017). *The Taiwan Voter*. University of Michigan Press. <https://doi.org/10.3998/mpub.9375036>
- Brambor, T., Clark, W. R., & Golder, M. (2006). Understanding interaction models: Improving empirical analyses. *Political Analysis*, 14(1), 63–82. <https://doi.org/10.1093/pan/mpi014>
- Brandtzaeg, P. B., Følstad, A., & Chaparro Domínguez, M. Á. (2018). How journalists and social media users perceive online fact-checking and verification services. *Journalism Practice*, 12(9), 1109–1129. <https://doi.org/10.1080/17512786.2017.1363657>
- Brennen, J. S., Simon, F., Howard, P. N., & Nielsen, R. K. (2020). *Types, sources, and claims of COVID-19 misinformation*. Reuters Institute. [https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2020-04/Brennen%20-%20COVID%2019%20Misinformation%20FINAL%20\(3\).pdf](https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2020-04/Brennen%20-%20COVID%2019%20Misinformation%20FINAL%20(3).pdf)
- Carlson, T. N. (2019). Through the grapevine: Informational consequences of interpersonal political communication. *American Political Science Review*, 113(2), 325–339. <https://doi.org/10.1017/S000305541900008X>
- Celik, I. (2020). Social media-specific epistemological beliefs: A scale development study. *Journal of Educational Computing Research*, 58(2), 478–501. <https://doi.org/10.1177/0735633119850708>
- Celik, I., Muukkonen, H., & Dogan, S. (2021). A model for understanding new media literacy: Epistemological beliefs and social media use. *Library & Information Science Research*, 43(4), Article 101125. <https://doi.org/10.1016/j.lisr.2021.101125>
- Clayton, K., Blair, S., Busam, J. A., Forstner, S., Glance, J., Green, G., & Nyhan, B. (2020). Real solutions for fake news? Measuring the effectiveness of general warnings and fact-check tags in reducing belief in false stories on social media. *Political Behavior*, 42(4), 1073–1095. <https://doi.org/10.1007/s11109-019-09533-0>
- De los Santos, T. M., & Nabi, R. L. (2019). Emotionally charged: Exploring the role of emotion in online news information seeking and processing. *Journal of Broadcasting & Electronic Media*, 63(1), 39–58. <https://doi.org/10.1080/08838151.2019.1566861>
- Ecker, U. K., O'Reilly, Z., Reid, J. S., & Chang, E. P. (2020). The effectiveness of short-format refutational fact-checks. *British Journal of Psychology*, 111(1), 36–54. <https://doi.org/10.1111/bjop.12383>
- Flynn, D. J., Nyhan, B., & Reifler, J. (2017). The nature and origins of misperceptions: Understanding false and unsupported beliefs about politics. *Political Psychology*, 38, 127–150. <https://doi.org/10.1111/pops.12394>
- Freiling, I., Krause, N. M., Scheufele, D. A., & Brossard, D. (2021, April). Believing and sharing misinformation, fact-checks, and accurate information on social media: The role of anxiety during COVID-19. *New Media & Society*. Advance online publication. <https://doi.org/10.1177/14614448211011451>
- Fridkin, K., Kenney, P. J., & Wintersieck, A. (2015). Liar, liar, pants on fire: How fact-checking influences citizens' reactions to negative advertising. *Political Communication*, 32(1), 127–151. <https://doi.org/10.1080/10584609.2014.914613>
- Garrett, R. K., Nisbet, E. C., & Lynch, E. K. (2013). Undermining the corrective effects of media-based political fact checking? The role of contextual cues and naïve theory. *Journal of Communication*, 63(4), 617–637. <https://doi.org/10.1111/jcom.12038>
- Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, 5(1), Article eaau4586. <https://doi.org/10.1126/sciadv.aau4586>
- Guess, A., Nyhan, B., & Reifler, J. (2018). *Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign*. European Research Council. <http://www.ask-force.org/web/Fundamentalists/Guess-Selective-Exposure-to-Misinformation-Evidence-Presidential-Campaign-2018.pdf>
- Guess, A. M., Lerner, M., Lyons, B., Montgomery, J. M., Nyhan, B., Reifler, J., & Sircar, N. (2020). A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proceedings of the National Academy of Sciences*, 117(27), 15536–15545. <https://doi.org/10.1073/pnas.1920498117>
- Klar, S., Leeper, T., & Robison, J. (2020). Studying identities with experiments: Weighing the risk of posttreatment bias against priming effects. *Journal of Experimental Political Science*, 7(1), 56–60. <https://doi.org/10.1017/XPS.2019.26>
- Lazer, D. M., 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>
- Lin, J. S., Lee, Y. I., Jin, Y., & Gilbreath, B. (2017). Personality traits, motivations, and emotional consequences of social media usage. *Cyberpsychology, Behavior, and Social Networking*, 20(10), 615–623. <https://doi.org/10.1089/cyber.2017.0043>
- Lottridge, D., & Bentley, F. R. (2018, April). *Let's hate together: How people share news in messaging, social, and public networks* [Proceedings]. 2018 CHI Conference on Human Factors

- in Computing Systems, 21–27 April, 2018, Montreal, Canada. <https://doi.org/10.1145/3173574.3173634>
- Lu, Y., & Lee, J. K. (2019). Stumbling upon the other side: Incidental learning of counter-attitudinal political information on Facebook. *New Media & Society*, 21(1), 248–265. <https://doi.org/10.1177/1461444818793421>
- Lyons, J., & Sokhey, A. (2014). Emotion, motivation, and social information seeking about politics. *Political Communication*, 31(2), 237–258. <https://doi.org/10.1080/10584609.2013.828138>
- Madden, M., Lenhart, A., Cortesi, S., & Gasser, U. (2013). *Teens and mobile apps privacy*. Pew Internet and American Life Project. https://www.pewinternet.org/wp-content/uploads/sites/9/media/Files/Reports/2013/PIP_Teens-and-Mobile-Apps-Privacy.pdf
- Marcus, G. E., & MacKuen, M. B. (1993). Anxiety, enthusiasm, and the vote: The emotional underpinnings of learning and involvement during presidential campaigns. *American Political Science Review*, 87(3), 672–685. <https://doi.org/10.2307/2938743>
- McLaughlin, B., Holland, D., Thompson, B. A., & Koenig, A. (2020). Emotions and affective polarization: How enthusiasm and anxiety about presidential candidates affect interparty attitudes. *American Politics Research*, 48(2), 308–316. <https://doi.org/10.1177/1532673X19891423>
- Mechkova, V., Pemstein, D., Seim, B., & Wilson, S. (2020). *Digital Society Project Dataset V2*. Database and Codebook. <http://digitalsocietyproject.org/data/>
- Mena, P. (2020). Cleaning up social media: The effect of warning labels on likelihood of sharing false news on Facebook. *Policy & Internet*, 12(2), 165–183. <https://doi.org/10.1002/poi3.214>
- Nekmat, E. (2020). Nudge effect of fact-check alerts: Source influence and media skepticism on sharing of news misinformation in social media. *Social Media+ Society*, 6(1), Article 2056305119897322. <https://doi.org/10.1177/2056305119897322>
- Nieminen, S., & Rapeli, L. (2019). Fighting misperceptions and doubting journalists' objectivity: A review of fact-checking literature. *Political Studies Review*, 17(3), 296–309. <https://doi.org/10.1177/1478929918786852>
- Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32(2), 303–330. <https://doi.org/10.1007/s11109-010-9112-2>
- Rich, T. S., Mildner, I., & Wagner, M. T. (2020). Research note: Does the public support fact-checking social media? It depends who and how you ask. *The Harvard Kennedy School Misinformation Review*. <https://misinfreview.hks.harvard.edu/article/research-note-does-the-public-support-fact-checking-social-media-it-depends-who-and-how-you-ask/>
- Rossini, P., Stromer-Galley, J., Baptista, E. A., & Veiga de Oliveira, V. (2021). Dysfunctional information sharing on WhatsApp and Facebook: The role of political talk, cross-cutting exposure and social corrections. *New Media & Society*, 23(8), 2430–2451. <https://doi.org/10.1177/1461444820928059>
- Roth, S., & Cohen, L. J. (1986). Approach, avoidance, and coping with stress. *American Psychologist*, 41(7), Article 813. <https://doi.org/10.1037/0003-066X.41.7.813>
- Storozuk, A., Ashley, M., Delage, V., & Maloney, E. A. (2020). Got bots? Practical recommendations to protect online survey data from bot attacks. *The Quantitative Methods for Psychology*, 16(5), 472–481. <https://doi.org/10.20982/tqmp.16.5.p472>
- Tully, M., Vraga, E. K., & Bode, L. (2020). Designing and testing news literacy messages for social media. *Mass Communication and Society*, 23(1), 22–46. <https://doi.org/10.1080/15205436.2019.1604970>
- Valentino, N. A., Hutchings, V. L., Banks, A. J., & Davis, A. K. (2008). Is a worried citizen a good citizen? Emotions, political information seeking, and learning via the Internet. *Political Psychology*, 29(2), 247–273. <https://doi.org/10.1111/j.1467-9221.2008.00625.x>
- Wang, A. H. E. (2019). The myth of polarization among Taiwanese voters: The missing middle. *Journal of East Asian Studies*, 19(3), 275–287. <https://doi.org/10.1017/jea.2019.25>
- Wise, K., Alhabash, S., & Park, H. (2010). Emotional responses during social information seeking on Facebook. *Cyberpsychology, Behavior, and Social Networking*, 13(5), 555–562. <https://doi.org/10.1089/cyber.2009.0365>
- Young, D. G., Jamieson, K. H., Poulsen, S., & Goldring, A. (2018). Fact-checking effectiveness as a function of format and tone: Evaluating FactCheck.org and FlackCheck.org. *Journalism & Mass Communication Quarterly*, 95(1), 49–75. <https://doi.org/10.1177/1077699017710453>
- Zhang, J., Featherstone, J. D., Calabrese, C., & Wojcieszak, M. (2021). Effects of fact-checking social media vaccine misinformation on attitudes toward vaccines. *Preventive Medicine*, 145, Article 106408. <https://doi.org/10.1016/j.ypmed.2020.106408>

Author Biography

Austin Horng-En Wang is an Assistant Professor of Political Science at the University of Nevada, Las Vegas. His research and teaching interests focus on public opinion, political psychology, data visualization, and East Asian politics.