Fighting misinformation on social media: effects of evidence type and presentation mode

Yunya Song, Sai Wang and Qian Xu

1Department of Journalism, School of Communication, Hong Kong Baptist University, Lee Shau Kee Communication and Visual Arts Building, 5 Hereford Road, Kowloon, Hong Kong SAR, China, 2Department of Communication Studies, School of Communication, Hong Kong Baptist University, Lee Shau Kee Communication and Visual Arts Building, 5 Hereford Road, Kowloon, Hong Kong SAR, China and 3Department of Strategic Communications, School of Communications, Elon University, 2850 Campus Box, Elon, NC 27244, USA.

*Correspondence to: S. Wang. E-mail: saiwang@hkbu.edu.hk

Received on 20 September 2021; editorial decision on 20 April 2022; accepted on 23 April 2022

Abstract

Designing corrective messages to debunk misinformation online is an important practice toward ending the coronavirus disease (COVID-19) pandemic as health-related misinformation has proliferated on social media misguiding disease prevention measures. Despite research on the use of statistical evidence and message modality in persuasion, the effects of evidence type (assertions with versus without statistical evidence) and presentation mode (text-only versus image-only versus text-plus-image) have been understudied. This study examined the impact of evidence type and presentation mode on individuals’ responses to corrective messages about COVID-19 on social media. The results showed that the presence of statistical evidence in assertions reduced message elaboration, which in turn reduced the effects of the message in correcting misperceptions, decreased perceived message believability and lowered social media users’ intentions to further engage with and disseminate the corrective message. Compared to the text-only modality and the text-plus-image modality, the image-only modality triggered significantly lower levels of message elaboration, which subsequently heightened message believability and increased user engagement intentions. The theoretical and practical implications are discussed.

The spread of misinformation during the coronavirus disease (COVID-19) pandemic is a serious public health threat to society. As the Director-General of the World Health Organization (WHO), Dr Ghebreyesus [1], pointed out: ‘We’re not just fighting an epidemic; we’re fighting an infodemic.’ Misinformation is defined as ‘incorrect or misleading statements’ that are ‘presented as facts’, and exposure to misinformation may lead to false beliefs [2]. The coronavirus-related misinformation circulating on social media takes on different forms, ranging from misleading claims that drinking methanol or bleach prevents COVID-19 to baseless assertions that the coronavirus is a human-made bioweapon. Misinformation may lead people to reject scientific consensus and engage in misguided behaviors, thereby reducing the effectiveness of disease prevention and control practices. It is thus imperative to examine factors that could mitigate misperceptions and to develop communication strategies to counter the tide of misinformation.

Although a burgeoning body of research has examined the effectiveness of various corrective message strategies [3, 4, 5], how different types of factual evidence and varying presentation modes affect the effectiveness of corrective information has been understudied. Past research suggests that providing evidence to support arguments might be an effective way in enhancing the persuasiveness of health-related messages [6, 7].
However, little attention has been given to the effectiveness of using different types of evidence in support of corrective messages in debunking misinformation. Moreover, given that health organizations (e.g. WHO) have designed and disseminated infographics to debunk coronavirus falsehoods [4], testing the efficacy of different presentation formats of corrective messages (e.g. text versus image) is also essential to the current practice of combating misinformation.

Drawing upon the elaboration likelihood model (ELM) [8], this study investigated the effects of evidence type (assertions with versus without statistical evidence) and presentation mode (text-only versus image-only versus text-plus-image) on cognitive and behavioral intention toward corrective information posted on social media. According to the ELM, the effects of evidence type and presentation mode on misinformation correction may be relayed through how much cognitive resources people allocate to processing corrective information (i.e. message elaboration) [9]. As such, this study tested the mediating effects of message elaboration, focusing on how evidence type and presentation mode may influence the extent to which people cognitively process the corrective message and how the information processing in turn affects their issue attitudes and behavioral intentions. Findings from this study not only contribute to our understanding of the cognitive processes underlying misinformation correction but also provide practical insights for the design of effective corrective messages to combat misinformation in a multimedia environment.

## Literature review

### Effects of evidence type

In health communication, providing evidence to support an advocated position is an important strategy to increase message persuasiveness [6, 7]. Here, evidence refers to factual statements that are provided to support a speaker’s claims [10]. Specifically, evidence may take the form of simple qualitative assertions or the assertions may include statistics presenting empirical numerical information. Thus, for this study, we focus on testing the effectiveness of assertions with versus without statistical evidence (i.e. an empirical quantitative description of objects or phenomena) [11, 12] in correcting health-related misinformation.

According to the ELM [8], there are two different routes for information processing that lead to persuasion: a central route through which individuals carefully evaluate the content of persuasive messages when making judgments and a peripheral route through which individuals form judgments by relying on situational factors or heuristic cues (e.g. message length or source credibility) rather than the message content itself. When persuasion occurs through the central route, individuals scrutinize the arguments presented in the message and relate the claims to what they have already known about the topic. More importantly, attitude change via the central route of information processing is more likely to ‘persist over time, resist change, and have an impact on other judgments and behavior’ than attitude change through the peripheral route [13, p. 234]. Therefore, if people attend carefully to the corrective information presented, they are more likely to become resistant to misinformation they encounter later.

The extent to which people engage in effortful thinking might be a function of evidence type. Research has demonstrated a strong impact of statistical evidence on message elaboration: Kopfman et al. [6] found that participants who were exposed to statistical evidence-based messages generated more issue-relevant thoughts than those reading narrative messages depicting personal stories. They concluded that compared to narrative evidence, statistical evidence is more likely to engage people in the thorough process of persuasion messages.

In the context of misinformation correction, using statistical evidence to support arguments might be an effective strategy to improve people’s acceptance of corrective messages and to promote collective actions fighting against misinformation [14]. In an experiment, Van der Meer and Jin [14] showed that exposure to corrective information
Fighting misinformation on social media

with statistical evidence evoked greater intention to take preventive actions during a public health crisis than exposure to corrective information without statistical support, which points to the critical role of statistical evidence in enhancing the effectiveness of corrective information.

In this process, message elaboration may serve as an underlying mechanism that affects people’s responses to corrective information. ELM literature suggests that more effortful processing of issue-relevant arguments in a message tends to make attitude change last longer, have a greater influence on behavior and be more resistant to change [8]. If statistical evidence increases the likelihood of engaging in message-relevant thinking, heightened message elaboration should lead people to integrate evidence-based arguments against misinformation into their underlying belief structure, thereby reducing their misperceptions regarding COVID-19. The message elaboration activated by statistical evidence may also affect judgments of corrective information. Research shows that statistical evidence-based messages presenting reliable facts and numerical information are generally perceived as more persuasive and of better quality [11]. Given that individuals engaging in more thoughtful processing are inclined to rely on message quality to evaluate the information, enhanced elaboration can be expected to lead individuals to consider corrective messages with statistical evidence as more believable and persuasive.

Additionally, cognitive elaboration of corrective information in a social media post may also influence people’s intention to further engage with that post. Empirical evidence suggests that cognitive elaboration is a key catalyst that promotes various user engagement behaviors [15]. In an experiment, Nekmat et al. [15] examined the mediating role of cognitive elaboration in social media engagement (e.g. liking, sharing and commenting on messages). They found that the message from personal sources induced higher levels of cognitive elaboration than the message from impersonal sources, which subsequently promoted individuals to engage in interactive activities. These findings not only demonstrate the important role of message elaboration in stimulating user engagement behaviors on social media but also suggest that these user engagement behaviors demand more cognitive efforts than previously assumed. Following the same reasoning, if people tend to carefully process corrective information with statistical evidence, we may also expect that heightened message elaboration will increase their likelihood of further engaging with the social media post through commenting or sharing. Therefore, the following hypotheses are proposed:

**Hypothesis 1**: Exposure to a corrective message with statistical evidence will elicit greater message elaboration compared to the same corrective message without statistical evidence.

**Hypothesis 2**: Greater message elaboration trig-gered by statistical evidence will, in turn, (i) reduce misperceptions, (ii) increase message believability and (iii) heighten user engagement intention.

**Effects of presentation mode**

In addition to evidence type, presentation mode (or modality) of information is another important factor that may affect people’s processing of corrective information and their subsequent cognitive and behavioral outcomes. According to the cue summation theory [16], multimodality communication with additional relevant cues leads to greater learning than single-modality communication. When the cues in two channels are closely related and elicit compatible responses simultaneously, they are expected to ‘summate to yield increased effectiveness’ [16, p. 238]. Consistent with this notion, Lazard and Atkinson [17] showed that participants reading infographic messages reported higher levels of issue-relevant elaboration than those reading text-only or visual-only messages. Similarly, Lee and Kim [18] found that the multimodal presentation (i.e. text-plus-graphic) of a news article triggered more active processing of the news content than the single-modal presentation (i.e. text-only), but this effect only appeared among participants with a stronger interest in the topic. These findings
demonstrate the power of multimodal presentation in facilitating message elaboration. If multimodality promotes message elaboration, it can be expected that individuals in the hybrid mode condition (text-plus-image) are more likely to engage in thoughtful thinking of corrective information than those in either the text-only or image-only condition.

Message elaboration is likely to mediate the effects of presentation mode on cognitive and behavioral responses to corrective information. In the context of misinformation correction, Amazeen et al. [19] examined the effectiveness of different corrective message formats in reducing misperceptions. They found that the inclusion of a visual rating scale facilitated participants’ ability to process the corrective information and increased the efficacy of the corrective message in countering misinformation. As such, it seems plausible that the enhanced elaboration of corrective information triggered by multimodality plays a positive mediating role in correcting people’s false beliefs and enhancing their perceived believability of corrective messages. Moreover, as multimodality leads individuals to engage in a more elaborate cognitive process, careful thinking of issue-related arguments in the corrective post should also make individuals more likely to comment on the post or forward it to their offline and/or online social networks, given that cognitive elaboration promotes user engagement behaviors [15]. Accordingly, the following hypotheses are proposed:

**Hypothesis 3:** Exposure to a corrective message in a hybrid mode of presentation (i.e. text-plus-image) will elicit greater message elaboration compared with the same corrective message in a text-only mode or an image-only mode.

**Hypothesis 4:** Greater message elaboration triggered by a hybrid mode of presentation will, in turn, (i) reduce misperceptions, (ii) increase message believability and (iii) heighten user engagement intention.

Taken together, the hypothesized model is shown in Fig. 1.

**Method**

**Study design**

This study was approved by the university’s Institutional Review Board. We conducted a web-based experiment with a 2 (evidence type: assertions with versus without statistical evidence) × 3 (presentation mode: text-only versus image-only versus text-plus-image) × 2 (misinformation topic: coronavirus origin versus face mask effectiveness) factorial design, with evidence type and presentation mode as between-subjects factors and misinformation topic as a within-subjects factor (repeated measures). Presentation mode describes the format of a Facebook post. We selected these two topics because they represent the types of misinformation that have widely spread during the COVID-19 pandemic [20], with the former involving conspiracy theories and the latter about virus transmission. The variation of misinformation topics will help enhance the generalizability of the findings. The information about the study procedure and message stimuli is detailed in Supplementary Material.

**Participants**

A total of 610 Hong Kong adults (301 males, 309 females) were recruited via Qualtrics, an online survey company that provides a diverse and representative sample. Study participation was restricted to those who have an active Facebook account. Given that the message stimuli used in this study were constructed following the layout of Facebook posts, participants need to have basic familiarity with Facebook to understand its primary user interactive features. The average age of the participants was 32.76 (SD = 9.69). Regarding the education level, 65.5% have obtained a bachelor’s degree or above. The majority of participants (80.2%) spent less than 3 h per day on Facebook.
Measures

The exact item wordings and descriptive statistics are shown in Table I. All items were measured on a 7-point Likert scale.

Message elaboration

Message elaboration was assessed using both the thought-listing technique and the self-report measure. Given that variations in the operationalization of the message elaboration construct may result in inconsistent findings, Shen and Seung [9] recommended research use both forms of message elaboration measures. Right after viewing each post, participants were asked to list any thoughts that came to their minds while reading the message. Two independent coders then coded the open-ended responses following the steps described in Shen and Seung [9]. The detailed coding procedure and the results were provided in Supplementary Material. A message elaboration index was then created by summing the number of favorable thoughts and unfavorable thoughts (coronavirus origin: $M = 1.22$, $SD = 1.15$, Min $= 0$, Max $= 6$; face mask effectiveness: $M = 1.24$, $SD = 1.03$, Min $= 0$, Max $= 9$). The higher the score, the greater the elaboration.

For the self-report measure, participants indicated the extent to which they agree with the five statements adapted from Kahlor et al. [21]. These items were averaged into an index for message elaboration, with higher values indicating greater elaboration (coronavirus origin: $M = 4.63$, $SD = 1.24$, Cronbach’s $\alpha = 0.89$; face mask effectiveness: $M = 5.14$, $SD = 1.17$, Cronbach’s $\alpha = 0.91$).

Perceptions about COVID-19

The approach to measure perceptions about COVID-19 was adapted from Vraga and Bode [22]. For each topic, participants rated their levels of agreement with three statements. The items were asked both before and after exposure to the corrective posts. A COVID-19 perception index was then created for each topic by averaging the three items. Subtracting the post-test value from the pre-test value indicates perception change. Responses to the face mask effectiveness questions were reverse-coded to ensure that the direction is consistent with responses to the coronavirus origin items. A negative value indicates a decrease in misperceptions, and vice versa (coronavirus origin: $M = -0.25$, $SD = 0.84$; mask effectiveness: $M = -0.05$, $SD = 0.78$).

Message believability

Adapted from Kim [23], participants were asked to indicate how well six pairs of adjectives describe
Table I. Item wordings and descriptive statistics of questionnaire items

| Item                                           | M                  | SD                  | M                  | SD                  |
|------------------------------------------------|--------------------|---------------------|--------------------|---------------------|
|                                                | Coronavirus origin | Face mask effectiveness | Coronavirus origin | Face mask effectiveness |
|                                                | 4.63 (4.15)        | 5.14 (5.65)         | 1.24 (1.69)        | 1.17 (1.02)         |
| message elaboration                            | 4.49 (4.35)        | 5.14 (5.65)         | 1.54 (1.82)        | 1.34 (1.74)         |
| 1. While reading the post, I thought about the actions I might take based on what I just read. | 4.51 (4.35)        | 4.95 (5.65)         | 1.53 (1.82)        | 1.44 (1.74)         |
| 2. While reading the post, I found myself making connections between the post content and the information I have read or heard elsewhere. | 4.84 (4.35)        | 5.21 (5.65)         | 1.40 (1.82)        | 1.32 (1.74)         |
| 3. While reading the post, I thought about how the content I just read could relate to the things I know. | 4.74 (4.35)        | 5.18 (5.65)         | 1.44 (1.82)        | 1.34 (1.74)         |
| 4. While reading the post, I tried to think of the practical application of what I just read. | 4.57 (4.35)        | 5.23 (5.65)         | 1.50 (1.82)        | 1.36 (1.74)         |
| 5. While reading the post, I tried to relate the ideas in the post to my own life. | 3.85 (3.16)        | 4.91 (3.97)         | 1.54 (1.73)        | 1.29 (1.66)         |
| message believability                          | 4.14 (3.55)        | 5.24 (4.34)         | 1.65 (1.82)        | 1.39 (1.74)         |
| 1. Unbelievable/believable                     | 3.83 (3.11)        | 4.98 (4.04)         | 1.68 (1.82)        | 1.46 (1.80)         |
| 2. Unconvincing/convincing                     | 3.88 (3.14)        | 5.02 (4.07)         | 1.66 (1.92)        | 1.44 (1.89)         |
| 3. Untrustworthy/trustworthy                   | 3.52 (3.06)        | 4.31 (3.83)         | 1.69 (1.93)        | 1.55 (1.88)         |
| 4. Uninformative/informative                   | 3.84 (3.14)        | 4.90 (3.59)         | 1.63 (1.87)        | 1.45 (1.82)         |
| 5. Inaccurate/accurate                         | 3.92 (4.00)        | 5.03 (3.68)         | 1.67 (1.87)        | 1.46 (1.82)         |
| 6. Non-credible/credible                       | 3.16 (3.55)        | 3.97 (4.34)         | 1.73 (1.82)        | 1.66 (1.74)         |
| user engagement intention                     | 3.55 (3.11)        | 4.34 (4.04)         | 1.82 (1.92)        | 1.74 (1.80)         |
| 1. This post is worth sharing with others.     | 3.14 (3.14)        | 4.07 (3.83)         | 1.92 (1.93)        | 1.89 (1.88)         |
| 2. I will recommend this post to others.       | 3.06 (3.06)        | 3.83 (3.59)         | 1.93 (1.87)        | 1.88 (1.82)         |
| 3. I will ‘Like’ this post on Facebook.        | 3.14 (4.40)        | 3.59 (5.65)         | 1.87 (1.69)        | 1.82 (1.02)         |
| 4. I will share this post on Facebook.         | 3.14 (4.15)        | 3.59 (5.70)         | 1.87 (1.69)        | 1.82 (1.02)         |
| 5. I will comment on this post on Facebook.    | 4.40 (4.15)        | 5.65 (5.70)         | 1.69 (1.81)        | 1.02 (1.08)         |
| perceptions about COVID-19                     | 4.15 (4.15)        | 5.70 (5.70)         | 1.81 (1.81)        | 1.08 (1.08)         |

(continued)
| Item                                                                 | M               | SD               |
|---------------------------------------------------------------------|-----------------|------------------|
| **Coronavirus origin**                                              |                 |                  |
| 1. Coronavirus is a manmade bioweapon.                              | 4.33 (pre-test) | 1.89 (pre-test)  |
|                                                                   | 4.14 (post-test)| 1.91 (post-test) |
| 2. Coronavirus is purposefully leaked from a laboratory.            | 4.32 (pre-test) | 1.79 (pre-test)  |
|                                                                   | 4.10 (post-test)| 1.83 (post-test) |
| 3. Coronavirus is a purposefully manipulated virus.                | 4.56 (pre-test) | 1.78 (pre-test)  |
|                                                                   | 4.22 (post-test)| 1.92 (post-test) |
| **Face mask effectiveness**                                         |                 |                  |
| 1. How likely do you think face masks will offer protection to the wearer against the COVID-19? | 5.62 (pre-test) | 1.11 (pre-test)  |
|                                                                   | 5.69 (post-test)| 1.14 (post-test) |
| 2. How likely do you think wearing face masks will be effective in reducing the transmission of the COVID-19? | 5.68 (pre-test) | 1.17 (pre-test)  |
|                                                                   | 5.68 (post-test)| 1.22 (post-test) |
| 3. How likely do you think wearing face masks will be useful in fighting against COVID-19? | 5.63 (pre-test) | 1.13 (pre-test)  |
|                                                                   | 5.64 (post-test)| 1.21 (post-test) |
| **Issue involvement**                                               |                 |                  |
| 1. Insignificant/significant                                        | 5.56            | 1.09             |
| 2. Unimportant/important                                            | 5.28            | 1.35             |
| 3. Irrelevant/relevant                                              | 5.71            | 1.23             |
| 4. Not concerning/concerning                                       | 5.64            | 1.27             |
| **COVID-19 risk perceptions**                                       |                 |                  |
| 1. The COVID-19 outbreak is having significant negative impacts on Hong Kong’s economy. | 6.11            | 1.07             |
| 2. The COVID-19 outbreak is a major threat to the health of the Hong Kong population. | 5.54            | 1.24             |
| 3. I am concerned about the spread of the COVID-19 on my health.    | 5.34            | 1.36             |
| 4. The COVID-19 outbreak is a serious threat to my personal financial situation. | 5.06            | 1.45             |
the corrective post they just read (coronavirus origin: $M = 3.85$, SD = 1.54, Cronbach’s $\alpha = 0.97$; mask effectiveness: $M = 4.91$, SD = 1.29, Cronbach’s $\alpha = 0.94$).

**User engagement intention**

Five items were adapted from Alhabash et al. [24] to measure participants’ intentions to further engage with or disseminate media content (coronavirus origin: $M = 3.16$, SD = 1.73, Cronbach’s $\alpha = 0.96$; mask effectiveness: $M = 3.97$, SD = 1.66, Cronbach’s $\alpha = 0.95$).

**Issue involvement**

Adapted from Zaichkowsky [25], participants indicated how well four pairs of adjectives describe the importance of the topic of COVID-19 to themselves ($M = 5.56$, SD = 1.09, Cronbach’s $\alpha = 0.88$).

**COVID-19 risk perceptions**

Participants’ risk perceptions about COVID-19 were measured with four items ($M = 5.51$, SD = 0.93, Cronbach’s $\alpha = 0.69$).

---

**Results**

**Effects of evidence type and presentation mode**

To test H1 and H3, mixed-effects analyses of covariance were conducted with the issue topic (coronavirus origin versus mask effectiveness) as the repeated measure and evidence type (assertions with versus without statistical evidence) and message modality (text-only versus image-only versus text-plus-image) as two between-subject factors. Two control variables—issue involvement and COVID-19 risk perceptions—were entered as covariates in the analyses. Table II presents descriptive statistics for the outcome variables by experimental conditions.

There was a significant main effect of evidence type on self-report message elaboration, $F(1, 602) = 6.79$, $p < 0.01$, $\eta^2 = 0.01$. Contrary to the prediction, assertions with statistical evidence elicited less elaboration ($M = 4.78$, SE = 0.06) than assertions without statistics ($M = 4.99$, SE = 0.06). For the thought-listing measure, there were no significant differences in message elaboration, $p = 0.13$. Thus, H1 was not supported.

The main effect of presentation mode on message elaboration for the thought-listing measure was also significant, $F(2, 602) = 4.72$, $p < 0.01$, $\eta^2 = 0.02$. Participants in the image-only condition generated fewer message-related thoughts ($M = 1.08$, SE = 0.06) than those in the text-only condition ($M = 1.36$, SE = 0.06), $p = 0.002$. Neither the difference between the text-plus-image condition and the text-only condition ($p = 0.22$) nor the difference between the text-plus-image condition and the image-only condition ($p = 0.07$) was significant. For self-reported measures, there were no significant differences in message elaboration between any conditions. Hence, H3 was not supported.

**Mediating effects of message elaboration**

To examine the mediating effects of message elaboration, Hayes’s [26] PROCESS macro (Model 4) was adopted. Indirect effects were estimated with 5000 bootstrap samples and 95% bias-corrected confidence intervals were reported (CIs). Evidence type and presentation mode were dummy coded with assertions without statistical evidence and image-only modality as respective baselines. The two forms of elaboration measures were entered as the mediating variables in the analyses, along with issue involvement and COVID-19 risk perceptions as covariates.

To address H2, for the coronavirus origin topic, the results revealed significant indirect effects of statistical evidence through self-report message elaboration on message believability [$B = −0.09$, SE = 0.04, 95% confidence interval (95% CI) = (−0.17, −0.01)] and user engagement intention [$B = −0.13$, SE = 0.06, 95% CI = (−0.25, −0.01)]. As shown in Fig. 2, the assertions with statistical evidence had a negative impact on message elaboration ($b = −0.20$, SE = 0.09,
To address H4, for the coronavirus origin topic, the results showed that the indirect effects of presentation mode on message believability via message elaboration (thought-listing) were significant [text-only modality versus image-only modality (reference): $B = -0.08$, SE = 0.03, 95% CI = ($-0.16$, $-0.03$); text-plus-image modality versus image-only modality (reference): $B = -0.05$, SE = 0.03, 95% CI = ($-0.12$, $-0.01$)]. Additionally, the results showed that message elaboration (thought-listing) significantly mediated the effects of presentation mode on user engagement intention [text-only modality versus image-only modality (reference): $B = -0.08$, SE = 0.03, 95% CI = ($-0.11$, $-0.01$)]. Specifically, compared to the image-only modality, the text-plus-image modality had a positive effect on message elaboration ($b = 0.26$, SE = 0.11, $p < 0.05$), and the text-only modality positively predicted message elaboration ($b = 0.41$, SE = 0.11, $p < 0.01$). The heightened message elaboration (thought-listing) was, however, associated with lower message believability ($b = -0.24$, SE = 0.05, $p < 0.05$).Message elaboration was positively associated with message believability ($b = 0.43$, SE = 0.05, $p < 0.001$) and user engagement intention ($b = 0.63$, SE = 0.05, $p < 0.001$). Other indirect effects failed to reach statistical significance (see Table III).

For the mask effectiveness topic, the results indicated that self-report message elaboration significantly mediated the effects of statistical evidence on misperceptions [$B = 0.02$, SE = 0.01, 95% CI = (0.003, 0.05)], message believability [$B = -0.08$, SE = 0.04, 95% CI = ($-0.16$, $-0.01$)] and user engagement intention [$B = -0.13$, SE = 0.06, 95% CI = ($-0.24$, $-0.02$)]. As shown in Fig. 3, The assertions with statistical evidence were negatively associated with message elaboration ($b = -0.21$, SE = 0.09, $p < 0.05$). Message elaboration negatively predicted misperceptions ($b = -0.12$, SE = 0.03, $p = 0.001$) and positively affected message believability ($b = 0.40$, SE = 0.04, $p < 0.001$) and user engagement intention ($b = 0.62$, SE = 0.06, $p < 0.001$). Other indirect effects of evidence type were statistically non-significant (see Table IV). Thus, H2 was not supported.

### Table II. Means and standard errors for outcome variables by experimental condition (N = 610)

| Evidence type | Modality       | n   | VO (SE) | ME (SE) | VO (SE) | ME (SE) | VO (SE) | ME (SE) | VO (SE) | ME (SE) |
|---------------|----------------|-----|---------|---------|---------|---------|---------|---------|---------|---------|
| Assertions   | Text-plus-image| 99  | 4.49 (1.40) | 5.09 (1.26) | 1.36 (1.15) | 1.24 (1.07) | -0.24 (0.85) | -0.14 (0.71) | 3.96 (1.67) | 5.01 (1.37) | 3.27 (1.90) | 4.05 (1.76) |
| with          | Text-only      | 101 | 4.45 (1.31) | 5.17 (1.11) | 1.30 (1.15) | 1.28 (0.85) | -0.21 (0.75) | 0.02 (0.83) | 3.73 (1.53) | 4.92 (1.18) | 3.14 (1.78) | 4.15 (1.61) |
| statistical   | Image-only     | 105 | 4.63 (1.27) | 4.87 (1.13) | 0.87 (0.88) | 0.98 (0.76) | -0.27 (0.90) | 0.09 (0.96) | 3.86 (1.45) | 4.73 (1.35) | 3.12 (1.78) | 3.65 (1.77) |
| evidence      | Text-plus-image| 102 | 4.84 (1.12) | 5.31 (0.98) | 1.17 (1.14) | 1.25 (1.00) | -0.34 (0.89) | -0.15 (0.74) | 3.94 (1.49) | 4.90 (1.36) | 3.12 (1.60) | 3.87 (1.61) |
| without       | Text-only      | 101 | 4.69 (1.14) | 5.30 (1.06) | 1.11 (1.16) | 1.34 (0.98) | -0.20 (0.81) | -0.06 (0.70) | 3.70 (1.61) | 4.97 (1.29) | 3.06 (1.69) | 4.23 (1.61) |
| statistical   | Image-only     | 102 | 4.66 (1.14) | 5.12 (1.21) | 1.52 (1.30) | 1.35 (1.41) | -0.24 (0.85) | -0.06 (0.69) | 3.93 (1.51) | 4.95 (1.17) | 3.25 (1.64) | 3.90 (1.58) |

Note: VO = virus origin, ME = mask effectiveness.

For the image-only modality (reference): $B = -0.05$, SE = 0.03, 95% CI = ($-0.12$, $-0.01$)]. Additionally, the results showed that message elaboration (thought-listing) significantly mediated the effects of presentation mode on user engagement intention [text-only modality versus image-only modality (reference): $B = -0.08$, SE = 0.03, 95% CI = ($-0.11$, $-0.01$)]. Specifically, compared to the image-only modality, the text-plus-image modality had a positive effect on message elaboration ($b = 0.26$, SE = 0.11, $p < 0.05$), and the text-only modality positively predicted message elaboration ($b = 0.41$, SE = 0.11, $p < 0.01$). The heightened message elaboration (thought-listing) was, however, associated with lower message believability ($b = -0.24$, SE = 0.05, $p < 0.05$). Message elaboration was positively associated with message believability ($b = 0.43$, SE = 0.05, $p < 0.001$) and user engagement intention ($b = 0.63$, SE = 0.05, $p < 0.001$). Other indirect effects failed to reach statistical significance (see Table III).
Fig. 2. A mediation model with standardized β coefficients (coronavirus origin).

Table III. Indirect effects of evidence type and presentation mode on outcome variables through message elaboration (topic: coronavirus origin)

| Evidence type (1 = assertions with statistical evidence) | B     | SE    | LLCI  | ULCI  |
|---------------------------------------------------------|-------|-------|-------|-------|
| Assertions with statistical evidence → message elaboration (SR) → misperceptions | -0.01 | 0.01  | -0.02 | 0.004 |
| Assertions with statistical evidence → message elaboration (TL) → misperceptions | -0.003 | 0.004 | -0.01 | 0.01  |
| Assertions with statistical evidence → message elaboration (SR) → message believability | -0.09  | 0.04  | -0.17 | -0.006 |
| Assertions with statistical evidence → message elaboration (TL) → message believability | 0.02   | 0.02  | -0.02 | 0.06  |
| Assertions with statistical evidence → message elaboration (SR) → user engagement intention | -0.13  | 0.06  | -0.25 | -0.01 |
| Assertions with statistical evidence → message elaboration (TL) → user engagement intention | 0.02   | 0.02  | -0.02 | 0.06  |

| Presentation mode (1 = text-only) | B     | SE    | LLCI  | ULCI  |
|----------------------------------|-------|-------|-------|-------|
| Text-only → message elaboration (SR) → misperceptions | -0.004 | 0.01  | -0.02 | 0.01  |
| Text-only → message elaboration (TL) → misperceptions | 0.004  | 0.01  | -0.02 | 0.01  |
| Text-only → message elaboration (SR) → message believability | 0.004  | 0.01  | -0.02 | 0.01  |
| Text-only → message elaboration (TL) → message believability | -0.05  | 0.03  | -0.12 | -0.02 |
| Text-only → message elaboration (SR) → user engagement intention | -0.07  | 0.07  | -0.21 | -0.02 |
| Text-only → message elaboration (TL) → user engagement intention | -0.08  | 0.04  | -0.16 | -0.02 |

Notes: SR = self-report, TL = thought-listing, B = unstandardized path coefficient, SE = standard error, LLCI = lower-level confidence interval, ULCI = upper-level confidence interval. Significant indirect path coefficients are bolded.
Fighting misinformation on social media

Fig. 3. A mediation model with standardized $\beta$ coefficients (face mask effectiveness).

$p < 0.001$) and lower user engagement intention ($b = -0.25$, SE = 0.06, $p < 0.01$). Other indirect effects failed to reach statistical significance (see Table III).

For the mask effectiveness topic, none of the indirect effects were statistically significant (see Table IV). Thus, H4 was not supported.

Discussion

This study examined how corrective messages with varying evidence types and presentation modes debunking COVID-19 misinformation on social media affected users’ cognitive and behavioral responses. The results indicated that corrective messages without statistical evidence enhanced message elaboration, which subsequently reduced misperceptions, increased perceived message believability and heightened intentions to share, like and comment on the Facebook post. Regarding the effects of presentation mode, the text-plus-image modality induced greater message elaboration than the image-only modality. The increased elaboration, however, resulted in less message believability and less user engagement intention.

While previous research found that statistical evidence prompted people to elaborate on message-relevant information [6], this study showed that participants in the non-statistical evidence condition processed the corrective post more actively and engaged in higher levels of message elaboration than those in the statistical evidence condition. One possible explanation is that our manipulation involved a very simple argument, such that participants might not need statistical evidence to make judgments about the message credibility or quality. As a result, they devoted less mental efforts to process the debunking information when it was backed by statistical evidence.

Combining a text-based message with an image made participants process corrective information more carefully and generate more message-relevant thoughts than an image-only message. It is consistent with previous findings that multimodal presentations enhanced message elaboration compared to single-modal presentations [17, 18], which echoes the cue summation theory [16], such that the cues
from different modalities (i.e. texts and images) conveying corrective information can complement each other and summate to produce a gain in learning.

Contrary to our expectations, message elaboration triggered by the multimodal presentation reduced participants’ perceptions of the believability of the corrective message and their intention to disseminate the corrective post. This finding might be explained by the limited capacity theory of information processing [27]. According to this theory, the message presented in multimodalities simultaneously would compete for users’ limited mental resources and overload the processing system, thereby negatively affecting the cognitive processing of the message [27, 28]. For example, Sundar [28] found that multimedia inhibited memory for news stories and evoked more negative perceptions of the website and its content. Similarly, in the present study, participants in the multimodal condition might exhaust their limited cognitive resources for attending to the information presented in the social media post and, thus, lead to more unfavorable evaluations of the corrective message and less intention to further engage with that post.

Furthermore, the effects of presentation mode on message elaboration were more prominent for the coronavirus origin topic but not for the mask effectiveness topic. The effects of presentation mode on user engagement intention were only held for the mask effectiveness topic but not for the coronavirus origin topic. These findings suggest that despite the same general topic (i.e. COVID-19), different focal aspects of the same topic can still produce differential results for misinformation corrective efforts. In other words, the effectiveness of presentation mode in corrective strategies may be more robust for certain types of falsehoods about COVID-19.
and weaker for the others. These differences could be explained by how mask wearing has become a norm in many Asian places, including Hong Kong [29]. Given that mask wearing has been considered as a socially acceptable (or even desirable) behavior, people might tend to have less elaboration on it regardless of how it is presented. But the graphic presentation could make the post more visually appealing and interesting, leading to more engagement intentions from social media users.

The current study also provides important practical guidelines for public health organizations to improve the effectiveness of corrective information on social media platforms. With an increasing reliance on social media to educate the public and debunk falsehoods, health organizations, public health professionals and campaign strategists can use the findings of this study to inform the design of corrective messages. For example, given that assertions without statistical evidence were shown to elicit greater misperception reduction, higher message believability and stronger user engagement intention than statistics-based assertions, it might be helpful to use more general statements (e.g. highly protective) as opposed to concrete numbers (e.g. 92% protection rate) in corrective messages. Additionally, considering that image-only mode allows users to process corrective information through the peripheral route and produces desirable persuasive outcomes, adopting graphic presentations of corrective information or adding visual elements to text-based corrective messages could increase the efficacy of misinformation correction and facilitate the spread of corrective information on social media.

However, our findings should be interpreted with caution for several reasons. First of all, the limited generalizability of the study’s findings merits attention. Participants in this study were limited to Hong Kong residents, and misinformation was limited to health-related issues. The conclusions of this study may not be applicable to populations with different cultural backgrounds or other types of misinformation (e.g. political topics). Future research could test these variables in different cultural contexts with different misinformation topics.

Furthermore, the current experiment was conducted during a global public health crisis, such that participants generally had a high level of involvement with the issue. As ELM suggests, highly involved individuals tend to carefully process the messages through the central route, while the peripheral route is likely to predominate for those with a low involvement in the topic [8]. Thus, individuals may process and evaluate corrective messages differently when the issue is less relevant or less important to oneself. Future research could take the timing into consideration and explore how people’s elaboration on misinformation correction varies during and after a public health crisis.

**Supplementary data**

Supplementary data are available at HEAL online.

**Conflict of interest statement**

All authors declare that they have no conflicts of interest.

**Funding**

This study is funded by the Public Policy Research Funding Scheme (Project Number: 2021.A2.047.21B) from the Policy Innovation and Co-ordination Office of the Government of the Hong Kong Special Administrative Region, and the Interdisciplinary Research Clusters Matching Scheme (IRCMS/19-20/D04) of Hong Kong Baptist University.

**References**

1. Adhanom Ghebreyesus T (2020). Munich Security Conference. World Health Organization. Available at: https://www.who.int/dg/speeches/detail/munich-security-conference. Accessed: 14 March 2021.
2. Green M, Donahue J. The effect of false information in news stories. In: Southwell BG, Thorson EA, Sheble L (eds). Misinformation and Mass Audiences. Austin, TX: University of Texas Press, 2018, 109–23.

3. Kim SC, Vraga EK, Cook J. An eye tracking approach to understanding misinformation and correction strategies on social media: the mediating role of attention and credibility to reduce HPV vaccine misperceptions. Health Commun 2021; 13: 1687–96.

4. Vraga EK, Bode L. Addressing COVID-19 misinformation on social media preemptively and responsively. Emerg Infect Dis 2021; 27: 396–403.

5. Vraga EK, Kim SC, Cook J et al. Testing the effectiveness of correction placement and type on Instagram. Int J Press/Polit 2020; 25: 632–52.

6. Kopfman JE, Smith SW, Ah Yun JK et al. Affective and cognitive reactions to narrative versus statistical evidence organ donation messages. J Appl Commun Res 1998; 26: 279–300.

7. Zebregs S, van den Putte B, Neijens P et al. The differential impact of statistical and narrative evidence on beliefs, attitude, and intention: a meta-analysis. Health Commun 2015; 30: 282–9.

8. Petty RE, Cacioppo JT. The elaboration likelihood model of persuasion. In: Petty RE, Cacioppo JT (eds). Communication and Persuasion. New York, NY: Springer, 1986, 1–24.

9. Shen L, Seung SY. On measures of message elaboration in narrative communication. Commun Q 2018; 66: 79–95.

10. McCroskey JC. A summary of experimental research on the effects of evidence in persuasive communication. Q J Speech 1969; 55: 169–76.

11. Baesler EJ, Burgoon JK. The temporal effects of story and statistical evidence on belief change. Commun Res 1994; 21: 582–602.

12. Church R, Wilbanks C. Values and Policies in Controversy: An Introduction to Argumentation and Debate. Scottsdale, AZ: Gorsuch-Scarisbrick, 1986.

13. Petty & Briñol. The elaboration likelihood model. In: Van Lange PA, Kruglanski AW, Higgins ET (eds). Handbook of Theories of Social Psychology, vol. 3. Los Angeles, CA: SAGE Publications, 2012, 224–45.

14. Van der Meer TG, Jin Y. Seeking formula for misinformation treatment in public health crises: the effects of corrective information type and source. Health Commun 2020; 35: 560–75.

15. Nekmat E, Gower KK, Zhou S et al. Connective-collective action on social media: moderated mediation of cognitive elaboration and perceived source credibility on personalness of source. Communc Res 2019; 46: 62–87.

16. Severin W. Another look at cue summation. Av Commun Rev 1967; 15: 233–45.

17. Lazard A, Atkinson L. Putting environmental infographics center stage: the role of visuals at the elaboration likelihood model’s critical point of persuasion. Sci Commun 2015; 37: 6–33.

18. Lee EJ, Kim YW. Effects of infographics on news elaboration, acquisition, and evaluation: prior knowledge and issue involvement as moderators. N Media Soc 2016; 18: 1579–98.

19. Amazeen MA, Thorson E, Muddiman A et al. Correcting political and consumer misperceptions: the effectiveness and effects of rating scale versus contextual correction formats. Journal Mass Commun Q 2018; 95: 28–48.

20. Brennen JS, Simon FM, Howard PN et al. (2020). Types, Sources, and Claims of COVID-19 Misinformation. Reuters Institute/University of Oxford. Available at: https://reutersinstitute.politics.ox.ac.uk/types-sources-and-claims-covid-19-misinformation. Accessed: 17 March 2021.

21. Kahlor L, Dunwoody S, Griffin RJ et al. Studying heuristic-systematic processing of risk communication. Risk Anal Int J 2003; 23: 355–68.

22. Vraga EK, Bode L. Using expert sources to correct health misinformation in social media. Sci Commun 2017; 39: 621–45.

23. Kim JW. Rumor has it: the effects of virality metrics on rumor believability and transmission on Twitter. N Media Soc 2018; 20: 4807–25.

24. Alhabash S, McAlister AR, Quilliam ET et al. Alcohol’s getting a bit more social: when alcohol marketing messages on Facebook increase young adults’ intentions to imbibe. Mass Commun Soc 2015; 18: 350–75.

25. Zaichkowsky JL. Measuring the involvement construct. J Consumer Res 1985; 12: 341–52.

26. Hayes AF. Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-based Approach. New York, NY: Guilford Publications, 2017.

27. Lang A. The limited capacity model of mediated message processing. J Commun 2000; 50: 46–70.

28. Sundar SS. Multimedia effects on processing and perception of online news: a study of picture, audio, and video downloads. Journal Mass Commun Q 2000; 77: 480–99.

29. Leung H (2020). Why wearing a face mask is encouraged in Asia, but shunned in the U.S. TIME. Available at: https://time.com/5799964/coronavirus-face-mask-asia-us/. Accessed: 17 March 2021.