Health Education Research, 2025, **40(2)**, cyaf009
DOI: https://doi.org/10.1093/her/cyaf009
Original Article

Correction approaches and hashtag framing in addressing Mpox misinformation on Instagram

Kelly Y.L. Ku[1], Jiarui Li[2], Yueming Luo[1] and Yunya Song[3,*]

[1]Department of Education and Psychology, Hong Kong Baptist University, Hong Kong SAR, China, [2]School of Journalism and Communication, Guangxi University, Nanning, China and [3]Division of Emerging Interdisciplinary Areas, Hong Kong University of Science and Technology, Hong Kong SAR, China

*Corresponding author. Division of Emerging Interdisciplinary Areas, The Hong Kong University of Science and Technology, Hong Kong SAR, China. E-mail: yunyasong@ust.hk

Abstract

The rapid spread of health misinformation on social media poses significant challenges to public health crisis. Mpox misinformation has portrayed it as exclusively a sexually transmitted infection, resulting in misperceptions about infection risk and stigmatization of affected groups. This study aimed to evaluate the effectiveness of different correction approaches and message framing in reducing misperception and shaping disease-related attitudes, both immediately after exposure and after a 1-day delay. We employed a 2 × 2 design with a control group to test correction approaches (fact-based vs. logic-based) combined with hashtag framing (health literacy vs. inclusivity) through an experiment ($N = 274$). Findings showed that all corrections reduced misperception both immediately and after 1 day and increased the likelihood of sharing corrective messages. Only corrections with inclusivity hashtags promoted more positive attitudes towards Mpox immediately after exposure. Stereotypes played a significant moderating role where participants with stronger stereotypes showed a greater reduction in misperception when exposed to corrections with inclusivity hashtags but were less likely to share logic-based corrective message. These findings contributed to understanding effective health communication by highlighting the role of social media hashtags in message framing, promoting user sharing of corrective information, and addressing stereotypes when designing interventions against health misinformation.

Introduction

The proliferation of health misinformation on social media can greatly distort public understanding, shape attitudes towards diseases, and hinder effective responses to health crises [1]. Misinformation, defined as false or inaccurate information that contradicts available evidence [2], poses significant challenges to public health. Its rapid spread impacts public understanding, health behaviours, and trust in institutions [3]. The global Mpox outbreak in 2022 exemplifies this, with rapidly spreading misinformation leading to misperceptions about infection risk, reduced protective measures, and stigmatization of affected groups [4]. The urgency of addressing health misinformation has become even more apparent with the World Health Organization's (WHO) recent declaration of an international public health emergency in August 2024 due to the circulation of Clade I Mpox cases [5]. This variant, associated with more severe illness and death than the previously dominant Clade II [6], poses an elevated threat that demands strategies for promoting accurate public health communication.

While social media can be a source of misinformation, it offers platforms for combating false beliefs through observational correction, which occurs when users witness the debunking of misinformation during their regular platform use [7]. Research has demonstrated the impact of observational correction across health contexts. For instance, corrective graphics shared by the WHO have successfully reduced

Received on 3 September 2024; accepted on 4 March 2025

COVID-19 misbeliefs [8], and exposure to corrective comments on vaccine safety reduced viewers' autism-vaccine misperceptions [9]. However, research suggests that correcting misbeliefs may not automatically translate into attitudinal changes—defined as summary evaluations ranging from positive to negative [10]. While some studies show that debunking false claims about vaccines fosters more positive attitudes [11], corrections of mental illness misinformation have not consistently reduced stigma [12], and negative attitudes can persist even after factual information is provided [13]. The effectiveness of observational correction also depends heavily on message amplification across social networks [14]. Given that false information often outpaces accurate content on social media [15], corrections must not only reduce misperceptions but also motivate sharing. Mihailidis and Viotty [16] argue that users should be empowered to become 'spreadable' agents of change, actively sharing corrective content. This user-driven approach to observational correction extends the reach of health communications while utilizing the trust inherent in peer-to-peer interactions [17]. Critical to this sharing-based strategy is understanding how corrections perform over time. Studies examining correction duration have yielded mixed results. Some researchers have observed a 'backfire effect' where misbeliefs intensify following a delay [18], while others demonstrate that corrections can maintain their effectiveness over time [19], highlighting the importance of designing interventions that sustain their impact as they spread through social networks.

Previous studies have investigated two main correction approaches. Fact-based corrections, which directly counter false claims with accurate information, have demonstrated efficacy across politics, health, and science communication [20]. The evolving complexity of misinformation has prompted the investigation of logic-based corrections, which address reasoning flaws underlying misinformation [21, 22]. For instance, Cook et al. [23] found that explaining flawed argumentation techniques was more effective than providing facts alone in countering climate change misinformation. Research examining these approaches has yielded mixed results, with fact-based corrections perceived as more credible in some studies [24], while others suggest that logic-based corrections offer benefits across various scenarios [22].

Beyond correction approaches, social media platforms offer unique features that can enhance message effectiveness. Particularly noteworthy is the evolution of hashtags from simple content organizers to sophisticated tools for message framing [25]. Hashtags have become integral features of social media, shaping how information is presented and perceived [26]. For instance, during the COVID-19 pandemic, the use of hashtags like '#chinesevirus' was associated with increased anti-Asian sentiment compared with more neutral tags like '#covid19' [27]. This demonstrates how hashtag choice can significantly impact public perception and response to health information.

Framing, as conceptualized by Entman [28], involves emphasizing certain aspects of a message to promote particular interpretations. Wang and Xu [29] highlighted that social media affordances allow for such framing through hashtags, providing additional context and guiding perceptions. Health influencers on Instagram use an average of six hashtags per post to promote specific health ideals [30], suggesting their strategic importance in health communication. The power of hashtags to frame discussions is evident in campaigns like #MeToo for social activism [31] and #VaccinesWork for public health advocacy [32]. In health communication specifically, hashtags can be strategically employed to enhance health literacy or promote inclusivity. Health literacy hashtags align with research indicating that improved health literacy can enhance understanding of health information [33], while inclusivity hashtags draw from studies showing that fostering a sense of community can increase engagement with health messages [34].

Another crucial consideration is how stereotypes influence information processing. Stereotypes, defined as cognitive structures containing beliefs and expectations about social groups [35], can act as filters leading to selective attention, interpretation, and recall of new information [36]. During the COVID-19 pandemic, individuals who endorsed greater stereotypes associating certain racial groups with spreading the virus were less likely to seek testing [37]. Similarly, beliefs in vaccine–autism connections made it challenging to correct related vaccine myths [38].

Our study examined the interplay between correction approaches and hashtag framing in addressing Mpox misinformation on Instagram, a platform that has received limited attention in misinformation research [24]. We targeted a pervasive misperception portraying Mpox as exclusively a sexually transmitted infection [39]. This misperception, stemming from disproportionate case numbers among men who have sex with men (MSM), has led to widespread risk underestimation, compromised vaccination willingness, and stigmatization [40, 41].

Specifically, we compared the effectiveness of fact-based and logic-based correction approaches, each paired with hashtags framed to promote either health literacy or inclusivity, which align with public health communication goals during health crises [21, 33]. We assessed how these combined corrections influence misperceptions about Mpox, attitudes, and sharing of corrective messages. We also examined how

		Hashtag framing		
Correction approach		Health literacy	Inclusivity	
	Fact-based	1 (N = 55)	2 (N = 55)	
	Logic-based	3 (N = 56)	4 (N = 53)	
				Control group (N = 55)

Fig. 1. The 2 × 2 experimental design with a control group.

stereotypes moderate correction effectiveness. Findings may inform public health communication during outbreaks by highlighting the role of social media hashtags in message framing, promoting user sharing of corrective information, and addressing individuals' stereotypes when designing interventions against health misinformation. Our specific research questions (RQs) and hypotheses are detailed in Supplementary Appendix A.

Methods

Experimental design

We employed a 2-day on-site experiment in a government-funded university in Hong Kong, China. Ethical approval was obtained from the affiliated institution. We utilized a five-cell experimental design with reference to Vraga et al. [24], comprising one control condition and four treatment conditions arranged in a 2 × 2 factorial structure (Fig. 1). The factors included correction approach (fact-based vs. logic-based) and hashtag framing (health literacy vs. inclusivity).

A priori power analysis based on an omnibus *F*-test in fixed effect analysis of variance (ANOVA) ($f = 0.225$, $\alpha = 0.05$, power = 0.80) indicated a sample size of 245. Participants were informed that their data would remain anonymous, all responses would be confidential and destroyed within 2 years post-study, participation was voluntary with the right to withdraw at any time without consequences, and they would receive monetary compensation of USD 13 upon completion. Participants were randomly assigned to one of the five conditions. On the first day, participants viewed the experimental stimuli and completed a questionnaire immediately after exposure. On the second day, participants completed a follow-up questionnaire to assess the durability of the effects.

Experimental conditions and stimuli

The experimental protocol involved exposing participants to a series of four Instagram posts across all conditions (Supplementary Appendix B). Participants in the experimental conditions were presented with an ordered sequence comprising a neutral post, a misinformation post, a correction post, and a neutral post. The control group had the correction post replaced by one additional neutral post.

The misinformation post read:

> If you are not part of the MSM community, your risk of contracting Mpox is almost nonexistent. WHO reports: In 2022-2023, 83% of Mpox cases occurred among "men who have sex with men" (MSM).

Correction posts were tailored to the assigned experimental condition. The fact-based correction read:

> Mpox is primarily transmitted from person to person through direct skin-to-skin contact or by inhaling respiratory droplets from an infected person. It can also spread through contact with an infected person's personal items or clothing. Mpox may be transmitted during sexual activity, but it is currently unclear how long the Mpox virus persists in semen and other bodily fluids.

The logic-based correction read:

> Here's an analogy to illustrate: Imagine that in a city, 83% of car accidents involve red cars. If you drive a car that isn't red, it would be illogical to assume that your likelihood of being involved in a car accident is lower. This fallacy equates the frequency of Mpox cases within a specific group with non-MSM people's likelihood of contracting Mpox, which is a hasty generalization fallacy.

The hashtag manipulation involved appending a set of three hashtags to each correction post based on the experimental condition. For the health literacy framing, the correction posts were accompanied by the hashtags #knowyourhealth, #healthknowledge, and #combatmisinfo. For the inclusivity framing, the posts featured #antistigma, #inclusivecommunity, and #combatmisinfo.

Measures

Measures of Day 1

Mpox misperception

We adopted the approach of Vraga et al. [24] to assess participants' Mpox misperception by asking participants to rate their agreement on a seven-point Likert scale ‘If you are not part of the MSM (men who have sex with men) community, your risk of contracting Mpox is almost nonexistent.’

Attitudes towards Mpox

Attitudes towards Mpox were assessed using an adapted checklist method [42]. Participants were instructed to reflect on their feelings about Mpox and select multiple descriptors from a list of 16 terms: ‘inclusiveness’, ‘compassion’, ‘supportiveness’, ‘empathy’, ‘nonthreatening’, ‘preventable’, ‘fear’, ‘blame’, ‘disgust’, ‘resistance’, ‘threatening’, ‘isolation’, ‘news report’, ‘scientific research’, ‘public health’, ‘and rare’. The initial term pool was derived from big-data analyses of ‘MonkeyPox2022Tweets’ [43]. These terms reflected three main orientations identified in prior literature: neutral (e.g. ‘public health’ and ‘news’), negative (e.g. ‘fear’ and ‘disgust’), and positive responses (e.g. ‘care’ and ‘empathy’). The checklist was further expanded to include terms drawn from commonly employed disease-related implicit association tests [44], incorporating both positive (e.g. ‘nonthreatening’) and negative (e.g. ‘threatening’) descriptors. All terms were randomly presented to participants to control for order effects. Through exploratory factor analysis, corroborated by hierarchical cluster analysis using Ward’s method, three distinct attitude clusters emerged: inclusive attitude (comprising terms ‘inclusiveness’, ‘compassion’, ‘supportiveness’, and ‘empathy’), aversion attitude (‘fear’, ‘blame’, ‘disgust’, and ‘resistance’), and neutral attitude (‘news report’, ‘scientific research’, and ‘public health’).

Correction sharing likelihood

We measured participants’ likelihood to share either the Mpox correction post (experimental conditions) or a Mpox-related neutral post (control condition). The item adapted from Su et al. [45] asked, ‘How likely are you to share this post?’ on a seven-point Likert scale.

Stereotype towards Mpox

Stereotypes towards Mpox (ST) of the participants were assessed using a two-item measure adapted from Zimmermann et al. [46] on a five-point Likert scale: ‘Mpox will be seen as yet another “gay disease” by the general public’ and ‘How much do you think Mpox will be used to blame gay men again?’

Measures of Day 2

On Day 2, participants completed the same Mpox misperception and attitude measures as on Day 1.

Additionally, we assessed participants’ information-seeking behaviours adapted from Basch et al. [43] related to Mpox following Day 1. We asked participants to report their post-Day 1 search behaviour with the item ‘I searched for Mpox-related information’ and post-Day 1 discussion behaviour with the item ‘I discussed Mpox-related information with family and friends’ on a five-point Likert scale.

Analysis plan

We used repeated-measures ANOVA to assess the primary effects of corrections on misperception, and chi-squared tests for attitudes and sharing likelihood. *Post hoc* comparisons were conducted for significant group differences. For variables showing primary effects, we performed additional regression analyses with covariates and interaction terms to explore the moderating effects of stereotypes and to check the robustness of primary effects.

Results

Participants characteristics

The final sample included 274 participants (51 males, 223 females; age range: 19–38, $M = 22.8$, $SD = 3.08$) from 320 initially recruited university students, excluding those who failed attention checks or did not complete the second-day procedure. The experimental groups were comparable regarding demographic variables, with no significant differences in sex (χ^2 (4) = 2.22, $P = .70$, Cramer’s $V = 0.09$, $P = .70$), age (Welch’s F (4, 133.74) = 1.23, $P = .30$), or family income (Welch’s F (4, 134.21) = 0.66, $P = .62$). Descriptive statistics for all variables are presented in Table I. The measure of STs reported a Cronbach’s alpha of 0.76.

Primary effects of experimental groups

Mpox misperception

Table II summarizes the primary effects of corrections on dependent variables. A repeated-measures ANOVA revealed significant main effects of corrections (F (4, 269) = 12.93, $P < .001$, $\eta P^2 = .16$) and time (F (1, 269) = 4.06, $P = .045$, $\eta P^2 = .015$). The correction × time interaction was nonsignificant (F (4, 269) = 1.22, $P = .30$, $\eta P^2 = .018$). Standard ANOVAs showed significant main effects of corrections on both Day 1 (F (4, 269) = 11.95, $P < .001$, $\eta P^2 = .15$) and Day 2 (F (4, 269) = 9.10, $P < .001$, $\eta P^2 = .12$). Dunnett’s tests revealed significantly lower misperception in all treatment groups compared to the control group on both days (all $P < .001$).

Correction sharing likelihood

ANOVA revealed significant main effects of corrections (F (4, 269) = 3.65, $P = .007$, $\eta P^2 = .051$). Dunnett’s tests indicated significantly higher sharing likelihood in all treatment conditions compared to the control group (all $P < .05$).

Table I. Descriptive statistics of Mpox misperception, correction sharing likelihood, Mpox attitudes, STs, post-Day 1 information-seeking behaviours, and demographics

	Fact-based correction with inclusive hashtags ($n = 55$)		Fact-based correction with health literacy hashtags ($n = 55$)		Logic-based correction with inclusive hashtags ($n = 53$)		Logic-based correction with health literacy hashtags ($n = 56$)		Control group ($n = 55$)	
Measure	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD
Day 1										
Mpox misperception	2.13	1.29	2.58	1.56	2.34	1.29	2.27	1.15	3.78	1.80
Correction sharing likelihood	4.09	1.70	3.69	1.61	3.75	1.71	3.66	1.63	2.93	1.63
Inclusive attitude[a]	28 (50.9%)		22 (40.0%)		28 (52.8%)		22 (39.3%)		15 (27.3%)	
Aversion attitude[a]	31 (56.4%)		31 (56.4%)		30 (56.6%)		34 (60.7%)		36 (65.5%)	
Neutral attitude[a]	54 (98.2%)		52 (94.5%)		52 (98.1%)		55 (98.2%)		53 (96.4%)	
ST	1.95	0.87	2.15	0.98	2.17	1.01	2.20	0.89	2.41	0.98
Day 2										
Mpox misperception	2.24	1.23	2.58	1.40	2.79	1.51	2.52	1.22	3.76	1.77
Inclusive attitude[a]	30 (54.5%)		24 (43.6%)		31 (58.5%)		34 (60.7%)		24 (43.6%)	
Aversion attitude[a]	32 (58.2%)		35 (63.6%)		28 (52.8%)		34 (60.7%)		34 (61.8%)	
Neutral attitude[a]	52 (94.5%)		53 (96.4%)		51 (96.2%)		55 (98.2%)		54 (98.2%)	
Post-Day 1 search[b]	25 (45.5%)		29 (52.7%)		19 (35.8%)		31 (55.4%)		24 (43.6%)	
Post-Day 1 discussion[b]	12 (21.8%)		19 (34.5%)		14 (26.4%)		20 (35.7%)		15 (27.3%)	
Demographics										
Sex (female %)	48 (87.3%)		45 (81.8%)		43 (81.1%)		45 (80.4%)		42 (76.4%)	
Age	22.87	2.79	23.56	4.27	22.47	2.69	22.20	2.78	22.89	2.46
Family income[c]	2.53	0.90	2.53	1.14	2.53	1.01	2.57	1.17	2.29	0.99

[a]Percentages in parentheses represent the proportion of individuals in each group who selected at least one item within the corresponding attitude subcategory.
[b]Percentages in parentheses represent the proportion of individuals in each group who engaged in the corresponding information-seeking behaviour after the Day 1 survey.
[c]Family monthly income was measured on a five-point scale ($1 \leq$ HK\$15 000, 2 = HK\$15 001–29 999, 3 = HK\$30 000–69 999, 4 = HK\$70 000–99 999, $5 = \geq$ HK\$100 000).

Table II. Primary effects of corrections on key dependent variables

	Mpox misperception[a] (Day 1)	Mpox misperception[a] (Day 2)	Correction sharing likelihood[a] (Day 1)	Inclusive attitude[b] (Day 1)	Inclusive attitude[b] (Day 2)	Aversion Attitude[b] (Day 1)	Aversion attitude[b] (Day 2)	Neutral attitude[b] (Day 1)	Neutral attitude[b] (Day 2)	Post-Day 1 search[b]	Post-Day 1 discussion[b]
F (4, 269)/χ^2 (4)	11.95, $P<.001$	9.10, $P<.001$	3.65, $P=.007$	9.50, $P=.05$	5.82, $P=.21$	1.46, $P=.83$	1.57, $P=.82$	2.04, $P=.73$	1.65, $P=.80$	5.24, $P=.26$	3.66, $P=.45$

[a]ANOVA was performed.
[b]Chi-squared test was performed.

Attitudes towards Mpox

Chi-squared tests revealed a significant effect on inclusive attitudes on Day 1 (χ^2 (4) = 9.50, P = .05) but not Day 2 (χ^2 (4) = 5.82, P = .21). No significant differences were found for aversion or neutral attitudes on either day.

Information-seeking behaviours

Chi-squared tests showed no significant differences among groups in searching for (χ^2 (4) = 5.24, P = .26) or discussing (χ^2 (4) = 3.66, P = .45) Mpox-related information after Day 1.

Interaction effects

Table III presents regression analyses exploring interaction effects (correction × ST) for dependent variables showing primary effects, with covariates included.

Mpox misperception

Significant interactions were observed between corrections and STs (fact-based correction with inclusivity hashtags × ST: $b_{\text{Day 1}} = -0.56$, SE = 0.28, P = .042; $b_{\text{Day 2}} = -0.72$, SE = 0.27, P = .008; logic-based correction with inclusivity hashtags × ST: $b_{\text{Day 1}} = -0.72$, SE = 0.26, P = .006). This indicates that individuals with stronger STs experienced greater reductions in Mpox misconceptions when inclusive hashtags were paired with either logic-based or fact-based corrections. ST positively associated with misperception across both days ($b_{\text{Day 1}} = 0.76$, SE = 0.19, $P < .001$; $b_{\text{Day 2}} = 0.83$, SE = 0.18, $P < .001$).

Correction sharing likelihood

Significant interactions were found between corrections and STs (logic-based correction with health literacy hashtags × ST: $b = -0.63$, SE = 0.31, P = .041; logic-based correction with inclusivity hashtags × ST: $b = -0.61$, SE = 0.30, P = .041). This indicates that the effects of increasing individuals' correction sharing intention by exposure to the logic-based correction with either inclusive hashtags or health literacy hashtags were less pronounced among individuals with high levels of the ST compared to those with low levels of the ST.

Inclusive attitudes towards Mpox

No significant relationship was observed between STs and inclusive attitudes ($b_{\text{Day 1}} = -0.16$, SE = 0.31, P = .61; $b_{\text{Day 2}} = -0.40$, SE = 0.29, P = .18). Primary effects were observed on Day 1 for fact-based (b = 0.94, SE = 0.43, P = .028) and logic-based (b = 1.15, SE = 0.42, P = .007) corrections with inclusivity hashtags, but not for corrections with health literacy hashtags. These effects were not observed on Day 2.

Table III. Regression models with interaction terms predicting Mpox misperception, correction sharing likelihood, and inclusive attitudes ($N = 274$)

	Mpox misperception[a] (Day 1)		Mpox misperception[a] (Day 2)		Correction sharing likelihood[a] (Day 1)		Inclusive attitude[b] (Day 1)		Inclusive attitude[b] (Day 2)	
Coefficient	***b***	***P***	***b***	***P***	***b***	***P***	***b***	***P***	***b***	***P***
Sex (female = 0)	−0.06 (0.22)	.80	−0.001 (0.21)	.99	**−0.66 (0.25)**	.008	−0.44 (0.34)	.20	**−0.97 (0.34)**	.005
Age	0.09 (0.11)	.42	0.12 (0.10)	.26	**−0.36 (0.12)**	.003	−0.33 (0.18)	.07	−0.27 (0.17)	.11
Education	−0.07 (0.11)	.52	−0.10 (0.10)	.32	**0.54 (0.12)**	<.001	0.32 (0.17)	.06	0.30 (0.17)	.07
Family Income	−0.10 (0.08)	.24	−0.12 (0.08)	.14	−0.10 (0.10)	.32	−0.20 (0.13)	.12	−0.16 (0.13)	.21
ST	**0.76 (0.19)**	<0.001	**0.83 (0.18)**	<.001	0.22 (0.21)	.31	−0.16 (0.31)	.61	−0.40 (0.29)	.18
Group 1[c]	**−1.40 (0.27)**	<0.001	**−1.27 (0.26)**	<.001	**1.34 (0.31)**	<.001	**0.94 (0.43)**	.028	0.15 (0.42)	.72
Group 2[c]	**−1.00 (0.27)**	<0.001	**−0.96 (0.26)**	<.001	**0.90 (0.31)**	.004	0.63 (0.42)	.13	−0.06 (0.41)	.88
Group 3[c]	**−1.23 (0.27)**	<0.001	**−0.74 (0.26)**	.005	**0.99 (0.31)**	.002	**1.15 (0.42)**	.007	0.59 (0.42)	.15
Group 4[c]	**−1.30 (0.27)**	<.001	**−1.03 (0.26)**	<.001	**0.93 (0.31)**	.003	0.56 (0.42)	.18	0.68 (0.41)	.10
Group 1[c] × ST	**−0.56 (0.28)**	.042	**−0.72 (0.27)**	.008	0.14 (0.32)	.65	−0.33 (0.44)	.46	−0.44 (0.45)	.33
Group 2[c] × ST	−0.04 (0.26)	.89	−0.05 (0.25)	.83	−0.28 (0.30)	.35	0.14 (0.42)	.73	0.21 (0.41)	.61
Group 3[c] × ST	**−0.72 (0.26)**	.006	−0.38 (0.25)	.13	**−0.61 (0.30)**	.041	0.13 (0.41)	.75	−0.12 (0.41)	.77
Group 4[c] × ST	−0.50 (0.27)	.07	−0.32 (0.26)	.22	**−0.63 (0.31)**	.041	0.38 (0.43)	.38	0.43 (0.42)	.30
(Adjusted/Nagelkerke) R^2	0.22		0.25		0.14		0.09		.14	

Continuous variables were *z*-scored. SEs are in parentheses, and significant *b* values ($P < .05$) are marked in bold.

[a]OLS model was performed.

[b]Logistics regression model was performed.

[c]Reference group = control group. Group 1 = fact-based correction with inclusive hashtags; Group 2 = fact-based correction with health literacy hashtags; Group 3 = logic-based correction with inclusive hashtags; Group 4 = logic-based correction with health literacy hashtags.

Discussion

Our study investigated the effects of social media corrections (employing two correction approaches and two hashtag framing) on Mpox misperception, likelihood of sharing corrective information, and attitudes towards the disease. Results show that all corrections significantly reduced misperception both immediately and after a 1-day delay and increased immediate sharing likelihood. Corrections with inclusivity hashtags promoted more inclusive attitudes towards Mpox. The effectiveness of corrections varied based on individuals' STs.

Effectiveness of corrections on misperception

A single exposure of correction was found effective in reducing misperception about Mpox (RQ1), aligning with previous research on the correction effect of HPV vaccination [47] and the COVID-19 outbreak [7]. We demonstrated that simply witnessing a correction can significantly reduce misperception.

The correction effect persisted after a 1-day delay. This finding is consistent with studies that have observed sustained effects after a delay. Swire et al. [19] observed that belief corrections remained stable after a 30-min delay but noted a decline after 1 and 3 weeks. Our results extend previous research by suggesting a potential peak short-term effect following a single social correction.

Approximately half of the participants engaged in further searches for related Mpox information after Day 1, and about a third discussed the topic with their families or friends. Such behaviours suggest that the correction exposure may motivate ongoing engagement with the topic and bring the subject into their personal networks.

Our results demonstrated that both fact-based and logic-based correction approaches were effective in reducing Mpox misperception, with these effects persisting over a 1-day delay. We did not observe significant differences in effectiveness between the two approaches, nor when combined with health literacy and inclusivity hashtags (RQ3 and RQ4a). This suggests that both approaches may be viable options for addressing health misinformation, offering flexibility to health communicators in designing corrective messages.

Impact on correction sharing likelihood

Participants exposed to correction demonstrated a greater likelihood to share the corrective message compared to the control group (H1). The willingness to share corrections is essential for combating the rapid spread of misinformation on social media [48]. Previous research has emphasized the potential of social media to facilitate user-driven dissemination of corrective information through social networks at a massive scale [19]. However, we did not observe significant differences in effectiveness between fact-based and logic-based correction approaches, nor between health literacy and inclusivity hashtags (RQ3 and RQ4b).

Influence on Mpox attitudes

Our findings for RQ2 revealed that exposure to any form of correction led participants to associate more inclusive adjectives with Mpox compared to the control group. This effect was evident immediately after exposure, but it did not sustain after a 1-day delay, in contrast to the persistent reduction in misperception observed. This discrepancy highlights that attitudinal changes can be more fleeting, influenced by affective responses that persist or bounce back, in our case, even after acknowledging corrective information [49].

Fact-based and logic-based correction approaches had similar effects on attitudes, with neither demonstrating a significant advantage (RQ3c), suggesting that the persuasive power of corrections on attitudes may lie less in the specific approach used.

Only corrections accompanied by inclusivity hashtags promoted more inclusive attitudes towards Mpox, though not sustained after a 1-day delay (RQ4c). While health literacy hashtags focus on individual health knowledge, inclusivity hashtags potentially triggered more empathetic responses by prompting participants to recognize stigma associated with the misinformation, fostering a sense of 'momentary connectedness' with broader anti-stigma conversations that take place in society [50].

Interaction with stereotypes

We found significant interactions between different corrections and STs (RQ5). Individuals with stronger STs, who at the same time exhibited higher levels of misperception compared to those with less strong stereotypes, benefited more in terms of reduction in misperception when exposed to corrections paired with inclusivity hashtags, regardless of whether the correction was fact-based or logic-based. One possible explanation is that when people encounter information inconsistent with their beliefs, they experience an unpleasant state of arousal that they are motivated to reduce [51]. In our study, while all individuals with stronger STs might experience this discomfort when exposed to any corrections, those exposed to inclusivity hashtags were offered a socially acceptable framing that eased them into adjusting their beliefs in a way that felt less threatening to their core values. For individuals with stronger stereotypes, exposure to language promoting acceptance for diversity may have prompted reflection on their broader views on social equity [26]. This could have motivated them to align more closely with these socially endorsed values.

We observed that individuals with higher STs were less likely to share corrective messages when the correction was logic-based. While the car accident analogy effectively reduces personal misperception by illustrating the probability fallacy, it lacks the Mpox-specific authoritative details that might be deemed more shareable. Those with high stereotypes may feel more comfortable disseminating factual, scientific-based information. Sharing scientific facts would allow them to engage in health discussions without directly confronting their own stereotypes.

Limitations

We noted several limitations. First, while our sample shows balanced demographic characteristics across experimental groups and reflects the typical Instagram user profile [52], where (I) users aged 18–34 years constitute the primary user base and (II) female users outnumber males across age groups—our sample, which primarily consists of university students, may not fully represent the broader population or all social media users. Younger adults, including university students, often exhibit higher digital literacy than users in other age groups [53], which may have enhanced their ability to process and evaluate corrective information. Furthermore, the gender imbalance (a high proportion of female participants) in our sample might have influenced the results of correction sharing, given documented gender differences in social media engagement patterns [54]. A more balanced gender ratio, and a more diverse age range and socioeconomic background could reveal different patterns in how corrections are processed and shared, particularly among groups with varying levels of health literacy and social media experience. Second, we only examined short-term effects, while research suggests that the impact of a single exposure to corrections may diminish after a week or longer [55]. Longer-term follow-ups and studies on repeated exposures could provide more insights into the sustainability of correction effects and potential attitude changes over time. Third, our use of nonverified accounts with neutral usernames, while aiming to mimic natural browsing, does not fully replicate real-world scenarios where information comes from various sources (e.g. experts, opinion leaders, and close social circles). Source credibility and social relationships may significantly influence message acceptance and sharing behaviour [56, 57]. Fourth, we focused on a specific piece of Mpox misinformation; future research should explore correction approaches across a broader range of health misinformation topics and contexts to establish their generalizability. Lastly, while our study focused primarily on textual information of corrections, social media communication increasingly relies on visual content such as images, videos, and interactive elements. The visual components of social media posts may significantly influence message processing, emotional responses, and information retention [58]. Future research should systematically examine how different visual elements interact with correction approaches to influence misperception reduction and attitude change.

Conclusion

Our findings offered important insights for addressing misinformation during public health outbreaks. Health communication could adopt a multifaceted approach to debunking misinformation with fact-based and logic-based approaches, as well as context-sensitive message framing. Strategic use of social media hashtags can effectively frame health messages to foster more positive attitudes towards affected groups, particularly when addressing stigma-prone diseases. It is also important to address commonly held stereotypes in health communication to improve the reception of health messages. Encouraging user engagement is crucial, as facilitating the sharing of corrective information by users can amplify its reach and impact. The content of corrective messages should be designed with characteristics that make health-related information shareable to increase dissemination. Lastly, correction efforts should consider both immediate and longer-term effects to maintain sustained impacts on public beliefs and behaviours.

Acknowledgements

During the preparation of this work, the authors used Claude 3.5 Sonnet to perform a grammar check. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Supplementary data

Supplementary data is available at *HEAL* online.

Conflict of interest

None declared.

Funding

This work was supported by the General Research Fund of the Hong Kong Research Grant Council (grant number 12615922).

Data availability

The data that support the findings of this study are available from the authors upon reasonable request and with the permission of the funding organization.

References

1. Featherstone JD, Zhang J. Feeling angry: the effects of vaccine misinformation and refutational messages on negative emotions and vaccination attitude. *J Health Commun* 2020;**25**:692–702.
2. Vraga EK, Bode L. Defining misinformation and understanding its bounded nature: using expertise and evidence for describing misinformation. *Polit Commun* 2020;**37**:136–44.
3. Caceres MMF, Sosa JP, Lawrence JA *et al*. The impact of misinformation on the COVID-19 pandemic. *AIMS Public Health* 2022;**9**:262–77.
4. Dalton ME, Duffy R, Quinn E *et al*. A qualitative review of social media sharing and the 2022 monkeypox outbreak: did early labelling help to curb misinformation or fuel the fire? *Sexual Health* 2024;**21**.
5. World Health Organization (WHO). *WHO Director-General Declares Mpox Outbreak a Public Health Emergency of International Concern*. https://www.who.int/news/item/14-08-2024-who-director-general-declares-mpox-outbreak-a-public-health-emergency-of-international-concern. Accessed: 1 June 2024.
6. Hoffmann C. Mpox—is there a more dangerous new clade? *Lancet Infect Dis* 2024;**24**:e667.
7. Bode L, Vraga EK. Correction experiences on social media during COVID-19. *Soc Media + Soc* 2021;**7**:205630512110088.
8. Vraga EK, Bode L. Addressing COVID-19 misinformation on social media preemptively and responsively. *Emerg Infect Dis* 2021;**27**:396–403.
9. van der Linden SL, Clarke CE, Maibach EW. Highlighting consensus among medical scientists increases public support for vaccines: evidence from a randomized experiment. *BMC Public Health* 2015;**15**:1207.
10. Briñol P, Petty RE, Stavraki M. Structure and function of attitudes. *Oxford Res Encycl Psychol* 2019.
11. Zhang J, Featherstone JD, Calabrese C *et al*. Effects of fact-checking social media vaccine misinformation on attitudes toward vaccines. *Preventive Med* 2021;**145**:106408.
12. Dobson KS, Rose S. "Myths and facts" campaigns are at best ineffective and may increase mental illness stigma. *Stigma and Health* 2022;**7**:27–34.
13. Swire-Thompson B, Lazer D. Public health and online misinformation: challenges and recommendations. *Ann Rev Public Health* 2020;**41**:433–51.
14. Bode L, Vraga EK. See something, say something: correction of global health misinformation on social media. *Health Commun* 2017;**33**:1131–40.
15. Vosoughi S, Roy D, Aral S. The spread of true and false news online. *Science* 2018;**359**:1146–51.
16. Mihailidis P, Viotty S. Spreadable spectacle in digital culture: civic expression, fake news, and the role of media literacies in "POST-Fact" society. *Am Behav Sci* 2017;**61**:441–54.
17. Vraga EK, Bode L. Using expert sources to correct health misinformation in social media. *Sci Commun* 2017;**39**:621–45.
18. Swire-Thompson B, Miklaucic N, Wihbey JP *et al*. The backfire effect after correcting misinformation is strongly associated with reliability. *J Exp Psychol Gen* 2022;**151**:1655–65.
19. Swire B, Ecker UKH, Lewandowsky S. The role of familiarity in correcting inaccurate information. *J Exp Psychol* 2017;**43**:1948–61.
20. Walter N, Murphy ST. How to unring the bell: a meta-analytic approach to correction of misinformation. *Commun Monogr* 2018;**85**:423–41.
21. Lewandowsky S, Ecker UKH, Seifert CM *et al*. Misinformation and its correction. *Psychol Sci Public Interes* 2012;**13**:106–31.
22. Schmid P, Betsch C. Effective strategies for rebutting science denialism in public discussions. *Nat Hum Behav* 2019;**3**:931–9.
23. Cook J, Ellerton P, Kinkead D. Deconstructing climate misinformation to identify reasoning errors. *Environ Res Lett* 2018;**13**:024018.
24. 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.
25. Bernard A. *Theory of the Hashtag*. Cambridge, United Kingdom, and Medford, MA02155, United States: Polity Press, 2019.
26. Zappavigna M, Martin JR. #Communing affiliation: Social tagging as a resource for aligning around values in social media. *Discourse Context Media* 2018;**22**:4–12.
27. Hswen Y, Xu X, Hing A *et al*. Association of "#covid19" versus "#chinesevirus" with anti-Asian sentiments on Twitter: March 9–23, 2020. *Am J Public Health* 2021;**111**:956–64.
28. Entman RM. Framing: toward clarification of a fractured paradigm. *J Commun* 1993;**43**:51–8.
29. Wang R, Xu WW. Hashtag framing and stakeholder targeting: an affordance perspective on China's digital public diplomacy campaign during COVID-19. *J Inf Technol Politics* 2022;**20**:250–68.
30. Pilgrim K, Bohnet-Joschko S. Selling health and happiness how influencers communicate on Instagram about dieting and exercise: mixed methods research. *BMC Public Health* 2019;**19**:1–9.
31. Dobrin D. The hashtag in digital activism: a cultural revolution. *J Cult Anal Soc Change* 2020;**5**:03.
32. Fuster-Casanovas A, Das R, Vidal-Alaball J *et al*. The #VaccinesWork hashtag on Twitter in the context of the COVID-19 pandemic: network analysis. *JMIR Public Health Surveill* 2022;**8**:e38153.
33. Krishna A, Thompson TL. Misinformation about health: a review of health communication and misinformation scholarship. *Am Behav Sci* 2019;**65**:316–32.
34. Harring N, Jagers SC, Löfgren Å. COVID-19: large-scale collective action, government intervention, and the importance of trust. *World Dev* 2021;**138**:105236.

35. Fiske ST, Tablante CB. *Stereotyping: Processes and Content*. Washington, DC 20002, USA: American Psychological Association eBooks, 2015, 457–507.
36. Bodenhausen GV, Kang SK, Peery D. *Social Categorization and the Perception of Social Groups*. London, United Kingdom: SAGE Publications Ltd eBooks, 2012, 311–29.
37. Earnshaw VA, Brousseau NM, Hill EC *et al*. Anticipated stigma, stereotypes, and COVID-19 testing. *Stigma and Health* 2020;**5**:390–3.
38. Pluviano S, Watt C, Della Sala S. Misinformation lingers in memory: failure of three pro-vaccination strategies. *PLoS One* 2017;**12**:e0181640.
39. Hazra A, Cherabie JN. Is mpox a sexually transmitted infection? Why narrowing the scope of this disease may be harmful. *Clin Infect Dis* 2022;**76**:1504–7.
40. Logie CH. What can we learn from HIV, COVID-19 and mpox stigma to guide stigma-informed pandemic preparedness? *J Int AIDS Soc* 2022;**25**:e26042.
41. Orsini D, Sartini M, Spagnolo AM *et al*. Mpox: “The stigma is as dangerous as the virus”. Historical, social, ethical issues and future forthcoming. *PubMed* 2023;**64**:E398–404.
42. Owen JJ, Rhoades GK, Stanley SM *et al*. “Hooking up” among college students: demographic and psychosocial correlates. *Arch Sex Behav* 2010;**39**:653–63.
43. Thakur N, Duggal YN, Liu Z. Analyzing public reactions, perceptions, and attitudes during the Mpox outbreak: findings from topic modeling of tweets. *Computers* 2023;**12**:191.
44. Zhu J, Yan L, Mu Y. Comparisons between COVID-19 stigma and other stigmas: distinct in explicit attitudes and similar in implicit process. *Front Psychol* 2022;**13**:848993.
45. Su M-H, Liu J, McLeod DM. Pathways to news sharing: issue frame perceptions and the likelihood of sharing. *Comput Hum Behav* 2019;**91**:201–10.
46. Zimmermann HML, Gültzow T, Marcos TA *et al*. Mpox stigma among men who have sex with men in the Netherlands: underlying beliefs and comparisons across other commonly stigmatized infections. *J Med Virol* 2023;**95**:e29091.
47. Vraga EK, Kim SC, Cook J. Testing logic-based and humor-based corrections for science, health, and political misinformation on social media. *J Broadcast Electron Media* 2019;**63**:393–414.
48. Yu W, Payton B, Sun M *et al*. Toward an integrated framework for misinformation and correction sharing: a systematic review across domains. *New Med Soc* 2022;**25**:2241–67.
49. Thorson EA. Belief echoes: the persistent effects of corrected misinformation. *Polit Commun* 2015;**33**:460–80.
50. Rathnayake C, Suthers DD. Twitter issue response hashtags as affordances for momentary connectedness. *Soc Media + Soc* 2018;**4**:205630511878478.
51. Susmann MW, Wegener DT. How attitudes impact the continued influence effect of misinformation: the mediating role of discomfort. *Pers Soci Psychol Bull* 2022;**49**:744–57.
52. NapoleonCat. *Instagram Users in Hong Kong*. 2023. https://napoleoncat.com/stats/instagram-users-in-hong_kong/2023/12/. Accessed: 14 December 2024.
53. Pew Research Center. *Americans and Digital Knowledge in 2019*. 2019. https://www.pewresearch.org/internet/2019/10/09/americans-and-digital-knowledge/. Accessed: 15 December 2024.
54. Muscanell NL, Guadagno RE. Make new friends or keep the old: gender and personality differences in social networking use. *Comput Hum Behav* 2012;**28**:107–12.
55. Ecker UK, O’Reilly Z, Reid JS *et al*. The effectiveness of short-format refutational fact-checks. *Br J Psychol* 2020;**111**:36–54.
56. Chao F, Zhou Q, Zhao J *et al*. Trustworthiness matters: effect of source credibility on sharing debunking information across different rumor types. *Inf Process Manage* 2024;**61**:103747.
57. Ma WW, Yuen AH. Understanding online knowledge sharing: an interpersonal relationship perspective. *Comput Educ* 2011;**56**:210–9.
58. Kim Y, Kim JH. Using photos for public health communication: a computational analysis of the Centers for Disease Control and Prevention Instagram photos and public responses. *Health Informatics J* 2020;**26**:2159–80.

Health Education Research, 2025, 40(2), cyaf009, DOI: https://doi.org/10.1093/her/cyaf009, Original Article