

Original Paper

The Challenge of Debunking Health Misinformation in Dynamic Social Media Conversations: Online Randomized Study of Public Masking During COVID-19

Mehdi Mourali¹, PhD; Carly Drake², PhD

¹Haskayne School of Business, University of Calgary, Calgary, AB, Canada

²Department of Management and Marketing, North Central College, Naperville, IL, United States

Corresponding Author:

Mehdi Mourali, PhD
Haskayne School of Business
University of Calgary
2500 University Drive NW
Calgary, AB, T2N 1N4
Canada
Phone: 1 403 220 6684
Email: mehdi.mourali@haskayne.ucalgary.ca

Abstract

Background: The spread of false and misleading health information on social media can cause individual and social harm. Research on debunking has shown that properly designed corrections can mitigate the impact of misinformation, but little is known about the impact of correction in the context of prolonged social media debates. For example, when a social media user takes to Facebook to make a false claim about a health-related practice and a health expert subsequently refutes the claim, the conversation rarely ends there. Often, the social media user proceeds by rebuking the critic and doubling down on the claim.

Objective: The aim of this study was to examine the impact of such extended back and forth between false claims and debunking attempts on observers' dispositions toward behavior that science favors. We tested competing predictions about the effect of extended exposure on people's attitudes and intentions toward masking in public during the early days of the COVID-19 pandemic and explored several psychological processes potentially underlying this effect.

Methods: A total of 500 US residents took part in an online experiment in October 2020. They reported on their attitudes and intentions toward wearing masks in public. They were then randomly assigned to one of four social media exposure conditions (misinformation only vs misinformation+correction vs misinformation+correction+rebuttal vs misinformation+correction+rebuttal+second correction), and reported their attitudes and intentions for a second time. They also indicated whether they would consider sharing the thread if they were to see it on social media and answered questions on potential mediators and covariates.

Results: Exposure to misinformation had a negative impact on attitudes and intentions toward masking ($\beta = -.35$, 95% CI $-.42$ to $-.29$; $P < .001$). Moreover, initial debunking of a false claim generally improved attitudes and intentions toward masking ($\beta = .35$, 95% CI $.16$ to $.54$; $P < .001$). However, this improvement was washed out by further exposure to false claims and debunking attempts ($\beta = -.53$, 95% CI $-.72$ to $-.34$; $P < .001$). The latter result is partially explained by a decrease in the perceived objectivity of truth. That is, extended exposure to false claims and debunking attempts appear to weaken the belief that there is an objectively correct answer to how people ought to behave in this situation, which in turn leads to less positive reactions toward masking as the prescribed behavior.

Conclusions: Health professionals and science advocates face an underappreciated challenge in attempting to debunk misinformation on social media. Although engaging in extended debates with science deniers and other purveyors of bunk appears necessary, more research is needed to address the unintended consequences of such engagement.

(*J Med Internet Res* 2022;24(3):e34831) doi: [10.2196/34831](https://doi.org/10.2196/34831)

KEYWORDS

misinformation; debunking; correction; social media; truth objectivity; COVID-19; infodemiology; health information; digital health; public health; health professional

Introduction

Context

The internet in general and social media in particular have become important sources of information for many people seeking medical and health-related information. As of 2014, 72% of internet users in the United States reported having searched for health-related information online [1]. More recently, 49% of US adults reported obtaining at least some of their news about the COVID-19 vaccine on social media, and among those who regularly obtain news from social media, 61% rated social media as an important way of keeping up with news about COVID-19 vaccines [2]. Yet, content disseminated through social media sites such as Facebook, Twitter, or Reddit remains largely unregulated, and is replete with false and misleading information [3-8].

The widespread availability and consumption of false, inaccurate, or incomplete health information (herein referred to as “health misinformation”) is a serious problem that can cause individual and social harm by promoting erroneous beliefs about health and illness, leading to detrimental behavior [9,10]. For example, the persistent circulation of unfounded claims linking vaccination to autism has convinced many parents not to immunize their children, which has resulted in a significant rise in vaccine-preventable diseases and death [11]. Believing misinformation about COVID-19 has been linked to lower adoption of health protective behaviors [12] and greater consumption of harmful products [13]. Evidence also suggests that online disinformation campaigns on a global scale have played a key role in the notable drop in vaccination coverage over time [14].

Recognizing the potentially disastrous consequences of letting misinformation proliferate on social media, researchers from diverse fields have proposed a variety of countermeasures, including technological solutions aimed at limiting exposure to misinformation, educational interventions aimed at empowering people to recognize and deal with misinformation, as well as communication tools designed to help debunk and correct misinformation [15-19]. Importantly, health experts and health care professionals have been called upon to play an active role in correcting health misinformation when they encounter it on social media [20-27].

Elsewhere, a rich literature on debunking has shown that properly designed corrections can be effective at countering misinformation [28-31]. This research, however, has examined the issue mostly from a static perspective. In a typical study, participants are first exposed to misinformation. Subsequently, some participants receive a correction, and their responses are compared to a control group that received no correction or a comparison group that received an alternate correction varying in its content, source, or some other relevant attribute (eg, [27,32-36]). Although this paradigm allows for a clean test of the relative effectiveness of specific debunking interventions,

it oversimplifies the dynamic nature of social media conversations. For example, when a social media user takes to Facebook to make a false claim about a health-related practice, and a health expert subsequently refutes the claim, the conversation rarely ends there. Often, the social media user proceeds by rebuking the critic and doubling down on the claim.

Objectives

The aim of this study was to examine the impact of such extended back and forth between false claims and debunking attempts on observers’ dispositions toward behavior favored by science. The US Centers for Disease Control and Prevention (CDC) had been recommending mask wearing in public since April 2020. However, by the time of our study in October 2020, less than half of the states had issued a mandate for mask wearing in public [37], and misinformation regarding the safety and efficacy of masking had become rampant on social media [38].

In this study, we tested competing predictions about the effect of extended exposure on people’s attitudes and intentions toward masking in public and explored several psychological processes potentially underlying this effect.

On the one hand, properly debunking a false claim may have a lasting effect, such that when the purveyor of misinformation proceeds to rebuke the critic and double down on the false claim, observers’ attitudes toward masking would remain unaffected by the new round of misinformation. A detailed refutation that includes a clear explanation of why a claim is false and what is true instead [16,39] could have a persistent impact on observers’ attitudes because, in addition to arming them with facts, it undermines the credibility of the argument underlying the false claim [40]. Moreover, research on “prebunking” has shown that it is possible to inoculate individuals against misinformation before it is even encountered [41-43]. Extending the principle of prebunking to prolonged social media debates, one could surmise that witnessing a thorough refutation of a false claim early in the debate may inoculate people against later assertions of the false claim.

On the other hand, repeated exposure to misinformation and its rebuttal could create uncertainty about the very existence of true facts [4,44]. For example, in reference to President Donald Trump’s extensive record of making false and misleading claims, Lewandowski et al [44] quoted a 2017 editorial from the Bangor Daily News suggesting that one important consequence of the repeated falsehoods is that “A third of the population will say ‘clearly the White House is lying,’ a third will say ‘if Trump says it, it must be true,’ and the remaining third will say ‘gosh, I guess this is unknowable.’” A year later, former Director of National Intelligence, James Clapper [45], warned that “many Americans are questioning if facts are even knowable, as foreign adversaries and our national leaders continue to deny objective reality while advancing their own alternative facts.”

In the health domain, exposure to conflicting information about a wide range of topics, including mammography [46,47], nutrition [48,49], and the human papillomavirus vaccine [50], has been shown to increase confusion, uncertainty, and negative attitudes toward the health topic in question. Together, these results suggest that exposure to an extended back and forth between false claims and debunking attempts may generate doubt about the very existence of a nonsubjective true answer to a health-related question. This study specifically examined the discourse surrounding the question of whether wearing masks in public should be prescribed. Based on extant theory, we hypothesized that such discourse would weaken people's attitudes and intentions toward masking in public as the prescribed behavior.

In addition to testing these competing predictions, we explored the impact of extended exposure on people's intentions to share the social media threads. This has direct implications for understanding how misinformation spreads on social media.

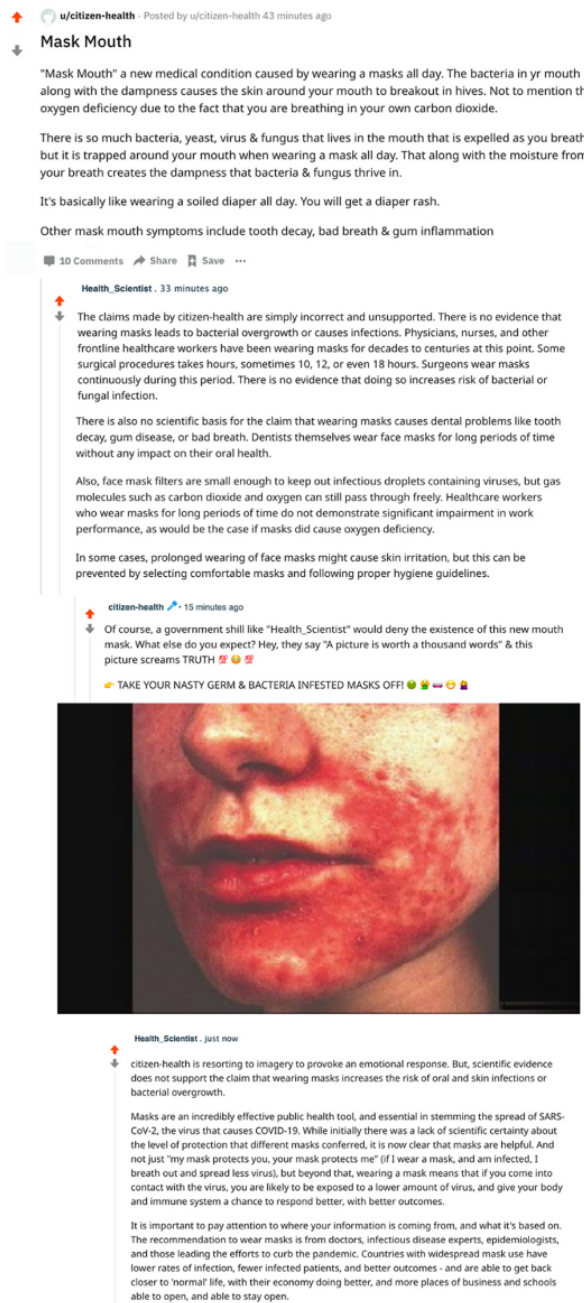
Methods

Participants and Procedure

Respondents were recruited from Prolific, a platform for recruiting online participants that explicitly caters to researchers. Prolific provides diverse and valid samples, and its data quality compares favorably with that of other online platforms such as MTurk [51,52]. US residents with an existing account on Prolific took part in the study on October 16, 2020, in exchange for monetary compensation.

Upon consenting in writing, participants answered questions about the relevance of various sources of health information, the impact of COVID-19 (including their perceived risk of infection), and reported on their attitudes and intentions toward wearing masks in public. They then completed a first attention check and were randomly assigned to one of four social media exposure conditions (misinformation only [M] vs misinformation+correction [MC] vs misinformation+correction+rebuttal [MCR] vs misinformation+correction+rebuttal+second correction [MCRC]). The attention check consisted of a statement at the end of a question asking participants to choose a specific answer to ensure they were reading the questions.

The social media content consisted of Reddit posts that were adapted from real social media posts and expert responses collected by the International Fact Checking Network [38], which are shown in full in Figure 1. Participants in the M condition reviewed a post by a user with the username *citizen-health* arguing that wearing face masks can cause a new disease called "mask mouth." Those in the MC condition saw the same post plus a correction from a user with the username *Health_Scientist*, pointing out that there is in fact no scientific evidence of any new disease caused by wearing face masks. Those in the MCR group saw a thread containing the previous two posts plus a second post from *citizen-health* rebutting *Health_Scientist* and doubling down on the original claim. Finally, those in the MCRC condition saw a thread containing the previous three posts plus a second correction from *Health_Scientist*.

Figure 1. Reddit posts for the misinformation+correction+rebuttal+second correction condition.

After reviewing the Reddit thread, participants completed a comprehension check assessing their comprehension of the position argued by *citizen-health*. They then indicated whether they would consider sharing the thread if they were to see it on social media, and reported their attitudes and intentions for a second time. They also answered questions on potential mediators and additional covariates. Mediators included perceptions of objectivity of truth, argument strength, and warmth and competence of the protagonists in the Reddit

exchange. Covariates included perceived COVID-19 risk, cognitive reflection, conspiracy mentality, and political orientation.

Ethical Considerations

The study was approved by the University of Calgary Conjoint Faculty Research Ethics Board (REB20-1178) and was conducted according to the principles expressed in the Declaration of Helsinki.

Measures

Attitudes were measured using three items (“Masking in public is necessary; good; beneficial”) rated on 7-point scales (1=*strongly disagree*, 7=*strongly agree*). Participants indicated their intention to wear a mask in public using a single item (“Over the next months, how often do you intend to wear a face mask when in public?”) also on a 7-point scale (1=*never*, 7=*all the time*). Since the attitude and intention measures were highly correlated, we combined them into a single index of disposition toward masking in public ($\alpha=.96$ preexposure and $\alpha=.97$ postexposure).

To measure perceived objectivity of truth, we asked respondents to consider the question “Should people wear masks in public?” and indicate the extent to which they think there is an objectively true answer to this question [53,54]. They reported their answers on a 7-point scale (1=*definitely no objective truth*, 7=*definitely an objective truth*).

Perceived argument strength was measured using five items adapted from Zhao et al [55]. A sample item is “The arguments of citizen-health are a convincing reason against masking in public” (1=*strongly disagree*, 7=*strongly agree*). The complete scale ($\alpha=.86$ for *citizen-health* and $\alpha=.94$ for *Health_Scientist*) is detailed in [Multimedia Appendix 1](#).

Participants rated the warmth and competence of both protagonists on 11-point bipolar scales adapted from previous research [40,56]. Perceived warmth was assessed using four items ($\alpha=.93$ for *citizen-health* and $\alpha=.96$ for *Health_Scientist*), including *unfriendly/friendly*, *cold/warm*, *irritable/good-natured*, *unsympathetic/sympathetic*. Perceived competence was assessed using six items ($\alpha=.97$ for *citizen-health* and $\alpha=.98$ for *Health_Scientist*), including *uninformed/informed*, *unqualified/qualified*, *unreliable/reliable*, *unbelievable/believable*, and *incompetent/competent*.

Respondents indicated the probability they will be infected with the coronavirus in the next 12 months on a sliding scale (0=0%, 100=100%), and rated how harmful it would be for their health if they were to become infected (1=*not at all*, 5=*extremely*). We computed a perceived COVID-19 risk score by multiplying the probability of infection by the perceived harm and dividing by 100.

Conspiracy mentality was measured using five items ($\alpha=.89$) rated on 7-point scales (1=*strongly disagree*, 7=*strongly agree*) adapted from Bruder et al [57]. A sample item is “Events which superficially seem to lack a connection are often the result of secret activities.”

We assessed cognitive reflection by combining the three items of Frederick’s [58] original Cognitive Reflection Test (CRT; eg, “A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost?”) with the four items of Thomson and Oppenheimer’s [59] nonnumeric CRT (eg, “If you’re running a race and you pass the person in second place, what place are you in?”). Answers were coded 1 for a correct answer and 0 for an incorrect answer. The final cognitive reflection score is the sum of all the correct answers.

The following two items ($r=0.89$) adapted from Schmid and Betsch [40] were used to measure political orientation: (1) If you think about your own political views, where would you classify your views on this scale? (1=*very conservative*, 7=*very liberal*), and (2) If you think about your own political identity, where would you classify your views on this scale? (1=*Republican*, 7=*Democrat*). Scores were reversed so that higher scores indicate political conservatism.

Participants rated their sharing intention on a single item: “If you were to see this post on social media, would you consider sharing it?” (1=*definitely not*, 5=*definitely yes*).

Statistical Analysis

Disposition Toward Masking in Public

Data analysis was performed using the statistical program R, version 4.0 [60], and the level of statistical significance was set at $\alpha=.05$. To answer our main research question, we examined how progressive exposure to false claims and debunking attempts affects people’s attitudes and intentions toward masking in public. Specifically, we tested whether a change in disposition toward masking in public varied from one exposure condition to the next.

Given the structure in our data (each participant provided two sets of ratings), we fit a linear mixed-effects model with disposition toward masking in public as the outcome variable; random intercepts for participants (ID); and fixed effects for exposure condition (contrast-coded using repeated contrasts), time of rating (contrast-coded using treatment contrast), and their interaction. We also added perceived COVID-19 risk, cognitive reflection, political orientation, and conspiracy mentality as mean-centered covariates in the model. The model was estimated using maximum likelihood. We compared this model’s goodness of fit to a second model that was identical but did not include the condition \times time interaction term. The likelihood ratio test indicated that model fit improved significantly when the interaction term was present ($\chi^2_3=33.71$, $P<.001$), thus suggesting a significant interaction.

The P values for the mixed-effects model with interaction were estimated via t tests using the Satterthwaite approximations to degrees of freedom. Effect sizes for the fixed effects are indicated by the standardized regression coefficients (β values) and their 95% CIs

Perceived Objectivity of Truth

Should people wear masks in public? We speculated that during extended debates, the reiteration of false information and rebuke of experts might shake people’s confidence, not only in the veracity of any answer to the question but also in the very existence of an objectively true answer. To test this idea, we performed an analysis of covariance (ANCOVA) on perceived objectivity of truth with exposure condition as the independent variable. Our model controlled for preexposure disposition toward masking in public, perceived COVID-19 risk, cognitive reflection, political orientation, and conspiracy mentality.

Perceived Argument Strength

To test whether exposure to extended debates influences perceptions of the strength of *citizen-health*'s arguments, we performed an ANCOVA on perceived argument strength, with condition as the independent variable and controlling for the same set of covariates as indicated above.

Mediation Analysis

We examined whether multiple processes may underly the effect of exposure to misinformation and debunking attempts on the change in people's disposition toward masking in public. Specifically, we tested the idea that correcting the original false claim may improve disposition toward masking, in part, by undermining the strength of the arguments put forth by the misinformation purveyor. Yet, further exposure may generate some doubt about the very existence of an objectively true answer, which, in turn, may weaken people's attitudes and intentions toward masking.

We performed two parallel mediation analyses. The first focused on the difference between the MC and M conditions, whereas the second focused on the difference between the MCR and MC conditions. The mediation models included the change in

disposition toward masking in public as the dependent variable, exposure condition as the independent variable, perceived objectivity as the first mediator, and perceived argument strength as the second mediator. Change in disposition toward masking in public was computed by subtracting participants' initial disposition scores from their postexposure scores. The models were estimated using maximum likelihood with robust standard errors.

Results

Sample Characteristics

A total of 500 participants responded to the survey. Four participants failed the initial attention check and 17 failed the comprehension check. Of those, 6 were in the MC condition, 3 were in the MCR condition, and 8 were in the MCRC condition. After removing responses from participants who failed at least one attention check, we were left with a final sample of 479 participants. The sample's demographic characteristics are shown in Table 1. It is worth noting that including data from participants who failed the attention or comprehension checks did not materially change the size, direction, or statistical significance of the reported effects.

Table 1. Sample characteristics (N=479).

Characteristic	Value
Age (years), mean (SD)	32.1 (12.3)
Gender, n (%)	
Female	257 (53.7)
Male	212 (44.6)
Other	8 (1.7)
Prefer not to answer	2 (0.4)
Education, n (%)	
Less than high school	8 (1.7)
High school graduate	55 (11.5)
Some college but no degree	124 (25.9)
Associate degree	42 (8.8)
Bachelor's degree	158 (33.0)
Master's degree	73 (15.2)
Doctoral degree	3 (0.6)
Professional degree	16 (3.3)
Employment, n (%)	
Employed full time	183 (38.2)
Employed part time	101 (21.1)
Unemployed looking for work	61 (12.7)
Unemployed not looking for work	28 (5.8)
Retired	17 (3.5)
Student	80 (16.7)
Disabled	9 (1.9)

Disposition Toward Masking in Public

Figure 2 shows the estimated marginal means and their 95% CIs. The means and SDs for all variables are presented in Table S1 of Multimedia Appendix 1. Prior to reviewing the Reddit threads, participants in all conditions reported similarly high dispositions toward masking in public. The estimated marginal means did not differ significantly between successive conditions, as shown in Table 2 (ie, P values for MC vs M, MCR vs MC, and MCRC vs MCR are all greater than .05). This confirmed that random assignment produced groups with equivalent baselines. The model statistics are summarized in Table 3. Furthermore, exposure to misinformation without any correction resulted in lower disposition toward masking in public (Time 2 in Table 2), while correcting false information improved disposition toward masking. Indeed, participants in the MC condition expressed more positive attitudes and intentions at time 2 than those in the M condition (Time 2: MC vs M in Table 2). Interestingly, the positive effect of correction seemed to vanish when the source of misinformation rebukes the correction

and doubles down on the false claim. Participants in the MCR condition reported significantly lower disposition scores than those in the MC condition (Time 2: MCR vs MC in Table 2). Perhaps even more concerning, a second round of corrections did not appear to effectively counter the impact of the rebuke and doubling down. Indeed, disposition toward masking in public did not differ significantly between participants in the MCRC and MCR conditions.

Although not our primary focus, it is worth noting that all four covariates had significant effects in the expected directions. Disposition toward masking in public was positively related to perceived COVID-19 risk and cognitive reflection (CRT). The latter is consistent with recent research showing CRT to be negatively correlated with the perceived accuracy of fake news, and positively correlated with the ability to discern fake news from real news [61]. Conversely, disposition toward masking in public was negatively related to conspiratorial thinking and political conservatism.

Figure 2. Estimated marginal means and 95% CIs for disposition toward masking in public across conditions and measurement times. M: misinformation-only experimental condition; MC: misinformation+correction experimental condition; MCR: misinformation+correction+rebutal experimental condition; MCRC: misinformation+correction+rebutal+second correction experimental condition.

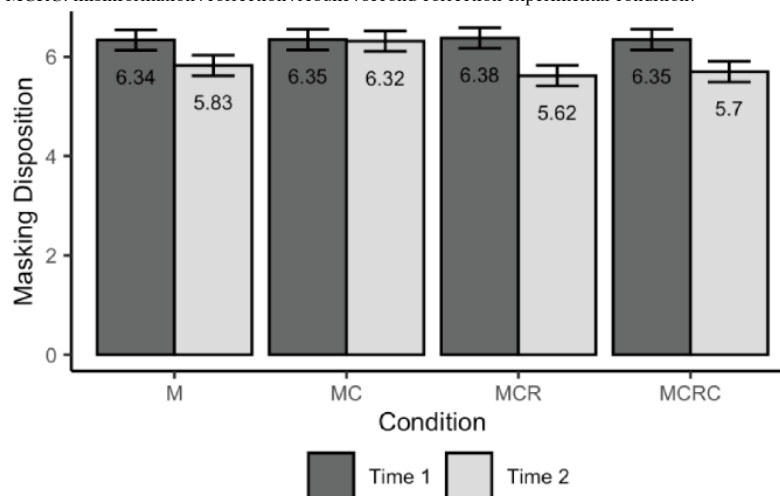


Table 2. Fixed effects for disposition toward masking in public from the mixed-effects regression model.

Predictors	Estimate, <i>b</i> (SE)	<i>t</i> (<i>df</i> =479)	<i>P</i> value	β (95% CI)
Intercept	6.33 (0.05)	121.13	<.001	.18 (.10 to .25)
COVID-19 risk	0.15 (0.05)	3.00	.003	.11 (.04 to .18)
CRT ^a	0.06 (0.02)	2.74	.006	.10 (.03 to .17)
Political orientation	−0.31 (0.03)	−10.48	<.001	−.39 (−.46 to −.32)
Conspiracy mentality	−0.12 (0.04)	−2.99	.003	−.11 (−.18 to −.04)
Time 2	−0.49 (0.05)	−10.38	<.001	−.35 (−.42 to −.29)
MC ^b versus M ^c	0.01 (0.15)	0.07	.94	.01 (−.20 to −.22)
MCR ^d versus MC	0.03 (0.15)	0.22	.83	.02 (−.19 to −.24)
MCRC ^e versus MCR	−0.03 (0.15)	−0.21	.83	−.02 (−.24 to .19)
Time 2: MC versus M	0.48 (0.13)	3.63	<.001	.35 (.16 to .54)
Time 2: MCR versus MC	−0.73 (0.13)	−5.48	<.001	−.53 (−.72 to −.34)
Time 2: MCRC versus MCR	0.11 (0.13)	0.84	.40	.08 (−.11 to .27)

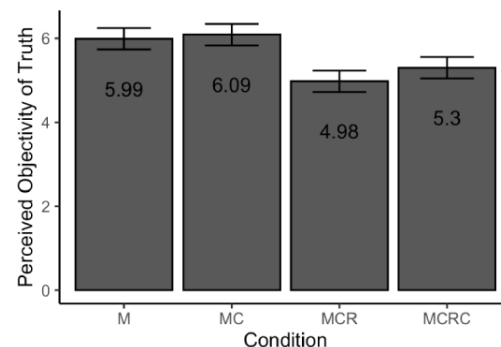
^aCRT: Cognition Reflection Test.^bMC: misinformation+correction.^cM: misinformation only.^dMCR: misinformation+correction+rebutal.^eMCRC: misinformation+correction+rebutal+second correction.**Table 3.** Random effects of the mixed-effects regression model for disposition toward masking in public.

Random effect	Value
σ^2	0.53
τ_{00id}	0.78
Intraclass correlation coefficient	0.60
N_{id}	479
Observations	958
Marginal R^2 /Conditional R^2	0.308/0.721

Perceived Objectivity of Truth

The patterns in our data (see the estimated marginal means in Figure 3) lend support to the idea that reiteration of false

information and rebutal of experts might weaken people's confidence in the very existence of an objectively true answer.

Figure 3. Estimated marginal means and 95% CIs for perceived objectivity of truth across conditions. M: misinformation-only experimental condition; MC: misinformation+correction experimental condition; MCR: misinformation+correction+rebutal experimental condition; MCRC: misinformation+correction+rebutal+second correction experimental condition.

We found a significant effect of condition ($F_{3,470}=16.96$, $P<.001$, $\eta_p^2=0.098$). Planned contrasts with Bonferroni correction for multiple tests revealed that people's perception of the objectivity of truth did not change significantly between the MC and M conditions (mean 6.18, SD 1.14 vs mean 5.91, SD 1.34, respectively; $t_{470}=0.53$, $P>.99$, $d=0.22$, 95% CI -0.03 to 0.47). However, exposure to a second round of misinformation (MCR condition) resulted in appreciably lower perceptions of truth objectivity compared to those for the MC condition (mean 4.96, SD 1.88; $t_{470}=-6.01$, $P<.001$, $d=-0.79$, 95% CI -1.05 to -0.52). Moreover, witnessing a second correction (MCRC condition) failed to improve perceptions of truth objectivity compared to the MCR condition (mean 5.31, SD 1.81; $t_{470}=1.74$, $P=.25$, $d=0.19$, 95% CI -0.07 to 0.44). These results suggest that once undermined, perceived objectivity of truth may be difficult to restore.

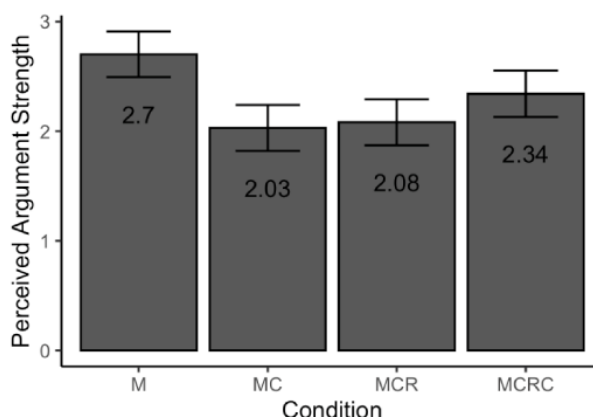
With respect to the covariates, we found that initial disposition toward masking was positively related to perceived objectivity ($b=0.31$, SE 0.06; $F_{1,470}=25.88$, $P<.001$, $\eta_p^2=0.05$), whereas political conservatism ($b=-0.17$, SE 0.04; $F_{1,470}=14.48$, $P<.001$, $\eta_p^2=0.03$) and conspiracy mentality ($b=-0.12$, SE 0.05; $F_{1,470}=5.06$, $P=.02$, $\eta_p^2=0.01$) were negatively related to perceived objectivity. The effects of cognitive reflection ($b=0.03$, SE 0.03; $F_{1,470}=0.92$, $P=.34$, $\eta_p^2=0.002$) and perceived COVID-19 risk ($b=0.10$, SE 0.07; $F_{1,470}=1.96$, $P=.16$, $\eta_p^2=0.004$) were not significant.

Perceived Argument Strength

The previous analysis suggested that reduction in the perceived objectivity of truth only occurs following exposure to the second round of misinformation. Thus, perceived objectivity of truth cannot account for the observed changes in attitude and intention across the entire range of exposure conditions. In particular, it cannot account for the improvement in disposition toward masking following correction of the initial false claim. Therefore, we next examined whether perceptions of the strength of *citizen-health's* arguments may provide an alternative account.

Exposure condition had a significant impact on perceived argument strength ($F_{3,470}=8.21$, $P<.001$, $\eta_p^2=0.05$). Planned contrasts with Bonferroni correction for multiple tests revealed that *citizen-health's* arguments were perceived to be weaker in the MC condition than in the M condition (mean 1.95, SD 1.26 vs mean 2.75, SD 1.59, respectively; $t_{470}=-4.43$, $P<.001$, $d=-0.55$, 95% CI -0.81 to -0.30). However, perceived argument strength remained low in the MCR condition and showed no difference from that in the MC condition (mean 2.12, SD 1.31; $t_{470}=0.34$, $P>.99$, $d=0.13$, 95% CI -0.12 to 0.39). It also did not differ significantly between the MCRC condition and the MCR condition (mean 2.34, SD 1.30; $t_{470}=1.72$, $P=.26$, $d=0.17$, 95% CI -0.09 to 0.42). The estimated marginal means are shown in Figure 4.

Figure 4. Estimated marginal means and 95% CIs for perceived argument strength across exposure conditions. M: misinformation-only experimental condition; MC: misinformation+correction experimental condition; MCR: misinformation+correction+rebuttal experimental condition; MCRC: misinformation+correction+rebuttal+second correction experimental condition.



As for the covariates, political conservatism ($b=0.14$, SE 0.04; $F_{1,470}=14.03$, $P<.001$, $\eta_p^2=0.029$) and conspiracy mentality ($b=0.16$, SE 0.04; $F_{1,470}=12.19$, $P<.001$, $\eta_p^2=0.025$) were positively related to perceived argument strength, whereas initial disposition to wearing masks ($b=-0.39$, SE 0.05; $F_{1,470}=62.96$, $P<.001$, $\eta_p^2=0.118$) and cognitive reflection ($b=-0.05$, SE 0.03; $F_{1,470}=4.13$, $P=.04$, $\eta_p^2=0.009$) were negatively related to perceived argument strength. Perceived COVID-19 risk ($b=0.08$,

SE 0.06; $F_{1,470}=1.92$, $P=.17$, $\eta_p^2=0.004$) was unrelated to perceived argument strength.

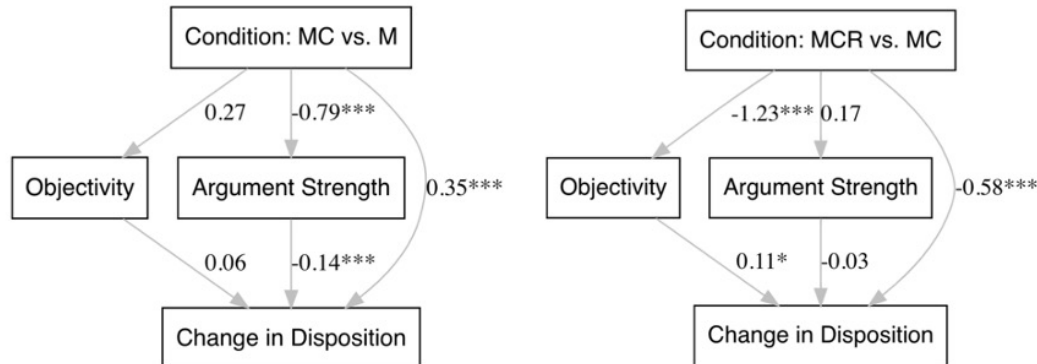
Mediation Analysis

When comparing the MC and M conditions, the effect of increased exposure on change in disposition toward masking in public was partially mediated by perceived argument strength (Figure 5). Indeed, both the indirect effect Condition through Argument Strength through Change in Disposition ($b=0.12$, SE 0.04, $z=2.96$; $P=.01$, 95% CI 0.04 – 0.19) and the direct effect

Condition through Change in Disposition ($b=0.35$, SE 0.09, $z=3.90$; $P<.001$, 95% CI 0.17-0.52) were significant. However, the second indirect effect Condition through Objectivity through Change in Disposition was not significant ($b=0.02$, SE 0.02, $z=0.93$; $P=.35$, 95% CI -0.02 to 0.05). Moreover, a formal test

of the difference between the indirect effects confirmed that the indirect effect through Argument Strength was significantly larger than the indirect effect through Perceived Objectivity ($b=0.10$, SE 0.04, $z=2.84$; $P=.004$, 95% CI 0.03-0.17).

Figure 5. Objectivity and argument strength mediate the effect of exposure on change in disposition. Values are unstandardized coefficients (b values). M: misinformation-only experimental condition; MC: misinformation+correction experimental condition; MCR: misinformation+correction+rebuttal experimental condition. * $P<.05$, *** $P<.001$.



Conversely, when considering the MCR and MC conditions, perceived objectivity partially mediated the effect of exposure on change in disposition. The indirect effect Condition through Objectivity through Change in Disposition was statistically significant ($b=-0.14$, SE 0.07, $z=-2.09$; $P=.04$, 95% CI -0.27 to -0.01), as was the direct effect Condition through Change in Disposition ($b=-0.58$, SE 0.13, $z=-4.59$; $P<.001$, 95% CI -0.83 to -0.33). Moreover, the indirect effect Condition through Argument Strength through Change in Disposition was not significant ($b=-0.004$, SE 0.01, $z=-0.47$; $P=.64$, 95% CI -0.02 to 0.01), and the indirect effect through Perceived Objectivity was larger than the indirect effect through Argument Strength ($b=0.14$, SE 0.07, $z=2.00$; $P=.05$, 95% CI 0.003-0.27).

These results suggest that no single process can fully account for the observed patterns in the data. Although a decrease in perceived argument strength partially explains why seeing a correction following initial exposure to misinformation improves attitudes and intentions toward masking, this path accounted for only 24.0% of the total effect of seeing a correction on change in attitudes and intentions. Moreover, a decrease in the perceived objectivity of truth partially explains why exposure to a second round of misinformation that includes rebuttal of the correction and doubling down on the original false claim negatively impacts attitudes and intentions toward masking. However, this path accounted for only 19.4% of the total effect of further exposure to misinformation on change in disposition toward masking.

Given the significant direct effect of exposure on change in disposition and the modest proportion mediated in both cases,

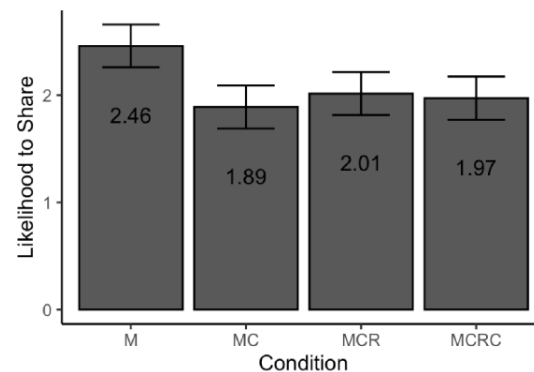
it is likely that there are additional mediators that could contribute to understanding the dynamic effects of extended exposure to misinformation on people's attitudes and intentions toward behaviors favored by science.

Additional Analyses

We tested for other potential mediators, but found that increased exposure did not influence participants' perceptions of *citizen-health's* warmth ($F_{3,470}=1.73$, $P=.16$, $\eta_p^2=0.01$), *citizen-health's* competence ($F_{3,470}=0.92$, $P=.43$, $\eta_p^2=0.006$), *Health_Scientist's* warmth ($F_{2,349}=0.63$, $P=.53$, $\eta_p^2=0.004$), *Health_Scientist's* competence ($F_{2,349}=0.37$, $P=.69$, $\eta_p^2=0.002$), or strength of *Health_Scientist's* arguments ($F_{2,349}=1.39$, $P=.25$, $\eta_p^2=0.008$).

Finally, we analyzed participants' intention to share misinformation on social media (see Figure 6). Exposure had a significant effect on sharing intention ($F_{3,470}=6.24$, $P<.001$, $\eta_p^2=0.038$). Planned contrasts with Bonferroni correction for multiple comparisons showed a decrease in the intention to share the original posting after seeing a correction (MC mean 1.88, SD 1.15 vs M mean 2.48, SD 1.40; $t_{470}=-3.93$, $P<.001$, $d=-0.47$, 95% CI -0.72 to -0.21). Furthermore, there was no significant difference in sharing intention between the MCR and MC conditions (mean 2.03, SD 0.99; $t_{470}=0.86$, $P>.99$, $d=0.14$, 95% CI -0.11 to 0.39), or between the MCRC and MCR conditions (mean 1.96, SD 1.09; $t_{470}=-0.29$, $P>.99$, $d=-0.07$, 95% CI -0.32 to 0.19).

Figure 6. Marginal means and 95% CIs of intention to share across exposure conditions. M: misinformation-only experimental condition; MC: misinformation+correction experimental condition; MCR: misinformation+correction+rebuttal experimental condition; MCRC: misinformation+correction+rebuttal+second correction experimental condition.



Examining the covariates, we found that initial disposition to wearing masks ($b=-0.23$, SE 0.05; $F_{1,470}=23.76$, $P<.001$, $\eta_p^2=0.046$) and cognitive reflection ($b=-0.06$, SE 0.02; $F_{1,470}=6.32$, $P=.01$, $\eta_p^2=0.013$) were negatively associated with intention to share. Neither perceived COVID-19 risk ($b=0.06$, SE 0.05; $F_{1,470}=1.33$, $P=.25$, $\eta_p^2=0.003$) nor political orientation ($b=0.04$, SE 0.03; $F_{1,470}=1.55$, $P=.21$, $\eta_p^2=0.003$) or conspiracy mentality ($b=0.01$, SE 0.04; $F_{1,470}=0.03$, $P=.86$, $\eta_p^2=0.000$) was significantly related to sharing intentions.

These results indicate that people are more likely to share misinformation when its content is consistent with their existing beliefs about the issue. Importantly, they also suggest that correcting a false claim can reduce the extent of its spread on social media, and this effect seems resistant to further exposure to the same misinformation.

Discussion

Principal Findings

Past research from across a variety of domains has shown that debunking misinformation works. That is, well-crafted corrections delivered by trusted sources can positively impact beliefs, attitudes, and intentions toward behavior favored by science [26-36]. With the explosion of false and misleading health claims on social media, especially since the start of the COVID-19 pandemic [62,63], scientists, experts, and health care professionals have been called upon to increase their presence on social media and help combat this “infodemic” [20-27]. A recent study found that US physicians and nurses are generally willing to take on the task even if it comes with important challenges [64]. However, little is known about the impact of extended social media debates on observers’ attitudes and intentions toward the debated issue. In this study, we tested such an impact in the context of a debate about the safety and effectiveness of wearing face masks in public during the early days of the pandemic.

We found that exposure to misinformation has a negative impact on attitudes and intentions toward masking. This result is consistent with prior research finding that exposure to

misinformation negatively impacts attitudes and intentions toward behaviors favored by science [40].

Also in line with prior work [28-31], we found that initial debunking of a false claim generally improves attitudes and intentions toward masking. This effect is partially explained by a decrease in the perceived strength of the argument underlying the false claim. However, this improvement is washed out by further exposure to false claims and debunking attempts. The latter result is partially explained by a decrease in the perceived objectivity of truth. That is, extended exposure to false claims and debunking attempts appears to weaken the belief that there is a nonsubjective, correct answer to how people ought to behave in this situation, which in turn leads to less positive reactions toward masking as the prescribed behavior. Interestingly, exposure to contradictory information affects perceived truth objectivity in a nonlinear fashion. For instance, exposure to a false claim and its initial debunking does not weaken the belief that there is an objectively true answer. It appears that the level of exposure to contradictory information needs to reach a certain threshold before it affects perceived truth objectivity.

Finally, we found that people are more likely to share misinformation when its content is consistent with their existing beliefs. However, correcting misinformation reduces its likelihood of being shared on social media, and this effect persists even after multiple exposures. These results, while highlighting the value of debunking in combating the spread of misinformation on social media, suggest that, unlike attitudes, sharing intentions may be insensitive to extended exposure to a back and forth between misinformation and correction. This pattern may reflect a floor effect, in that people had expressed very low intentions to share corrected misinformation (mean of 1.89 on a 7-point scale). Exposure to further debate resulted in more negative attitudes toward masking, but did not impact sharing intentions because sharing intentions were already extremely low.

Comparison With Prior Work

Our findings have important implications for research on debunking misinformation. Extant literature has noted that even though corrections generally reduce people’s beliefs in false information, the misinformation often continues to influence

their thinking, a phenomenon known as the continued influence effect [65]. Once people process information that appears vaguely credible to them, it becomes difficult to retract it. A popular explanation of the continued influence effect assumes that people build mental models of the world and want their models to be complete. They are willing to accept false information if it allows them to build complete models. When that false information is later debunked, it creates a gap in their understanding of the world. Since people dislike gaps in their understanding and prefer their mental models to be complete, they continue to rely on information they know is false. [65,66].

Another explanation of the continued influence effect argues that attempts to correct misinformation often end up reinforcing it through repetition [67,68]. From this perspective, repeating misinformation when attempting to correct it makes it feel more familiar and fluent. By inadvertently increasing the ease with which misinformation is processed, correction attempts also increase the likelihood of it being accepted as true. Our study suggests yet another possible explanation of the stickiness of misinformation. In some situations, witnessing a heated debate with arguments for and against a controversial issue could undermine people's confidence in the existence of an objectively true answer, which may weaken their commitment to either side of the debate.

Limitations and Future Research

We set out to study the impact of exposure to extended debates on social media. However, our study was limited to a single platform (Reddit), and the debate was restricted to four exchanges between only two protagonists. This limits the generalizability of our findings. Interaction norms likely differ across social media platforms, which may impact how users interpret the conversation. Future research could attempt to replicate our findings using different social media platforms (eg, Facebook and Twitter), and examine the consequences of extending the debate to include more than two protagonists and more than four exchanges. Relatedly, although extended debates such as those described in this research are familiar to regular users of social media, we do not know how often they happen. Future research would benefit from quantifying the frequency of their occurrence and how it may vary across platforms.

A possible limitation of our experimental design is that the rebuke message included an image, whereas all other messages were strictly text-based. Thus, it is impossible to disentangle the impact of exposure to the rebuke message from the presence of an image. However, pinpointing which message element (image vs text) accounts for the effect of exposure was not a goal of this study. Instead, we sought to pit strong refutation against persuasive misinformation. We chose to include a graphic image in the rebuke message precisely because of its persuasive power.

Although the debunking messages used in this study were developed using recommended best practices, it is possible that different debunking techniques would have resulted in different outcomes. Future research would greatly benefit from manipulating features of the debunking argument as well as the source of debunking. For example, previous research has shown that debunking messages from anonymous social media users are more effective when they include a link to a trusted source such as the CDC [26]. Future research could test whether the positive impact of providing links to credible sources persists in the context of extended social media debates. Other features of debunking messages previously found to sometimes reduce misperceptions include the use of humor [69] and infographics [70]. However, whether such debunking techniques can be effective in the context of extended debates remains an open question.

Preemptively refuting misinformation or even just warning people that they might be misinformed has been shown to decrease later reliance on misinformation. Future research could test the effectiveness of such prebunking in the context of extended debates, where—to extend the biomedical analogy—the viral load is higher.

Conclusions

In sum, dynamic conversations present a heretofore underappreciated challenge faced by health professionals and science advocates attempting to debunk misinformation on social media. Engaging in extended debates with science deniers and other purveyors of bunk appears necessary, but more research is needed to address the unintended consequences of such engagement.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Complete scales for measures and descriptive statistics.
[DOCX File , 27 KB-Multimedia Appendix 1]

References

1. Fox S. The social life of health information. Pew Research Center. 2014 Jan 15. URL: <https://www.pewresearch.org/fact-tank/2014/01/15/the-social-life-of-health-information/> [accessed 2021-09-21]
2. Mitchell A, Liedke J. About four-in-ten Americans say social media is an important way of following COVID-19 vaccine news. Pew Research Center. 2021 Aug 24. URL: <https://www.pewresearch.org/fact-tank/2021/08/24/about-four-in-ten-americans-say-social-media-is-an-important-way-of-following-covid-19-vaccine-news/> [accessed 2021-09-30]

3. Sylvia Chou W, Gaysynsky A, Cappella JN. Where we go from here: health misinformation on social media. *Am J Public Health* 2020 Oct;110(S3):S273-S275. [doi: [10.2105/ajph.2020.305905](https://doi.org/10.2105/ajph.2020.305905)]
4. Southwell BG, Thorson EA, Sheble L. The persistence and peril of misinformation. *Am Sci* 2017;105(6):368-371. [doi: [10.1511/2017.105.6.372](https://doi.org/10.1511/2017.105.6.372)]
5. Vosoughi S, Roy D, Aral S. The spread of true and false news online. *Science* 2018 Mar 09;359(6380):1146-1151. [doi: [10.1126/science.aap9559](https://doi.org/10.1126/science.aap9559)] [Medline: [29590045](https://pubmed.ncbi.nlm.nih.gov/29590045/)]
6. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. *J Med Internet Res* 2021 Jan 20;23(1):e17187 [FREE Full text] [doi: [10.2196/17187](https://doi.org/10.2196/17187)] [Medline: [33470931](https://pubmed.ncbi.nlm.nih.gov/33470931/)]
7. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. *Soc Sci Med* 2019 Nov;240:112552 [FREE Full text] [doi: [10.1016/j.socscimed.2019.112552](https://doi.org/10.1016/j.socscimed.2019.112552)] [Medline: [31561111](https://pubmed.ncbi.nlm.nih.gov/31561111/)]
8. Cummings M. Misinformation on Reddit has become unmanageable, Alberta moderators say. *CBC News*. 2021 Sep 21. URL: <https://www.cbc.ca/news/canada/edmonton/misinformation-alberta-reddit-unmanageable-moderators-1.6179120> [accessed 2021-09-21]
9. Broniatowski DA, Jamison AM, Qi S, AlKulaib L, Chen T, Benton A, et al. Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. *Am J Public Health* 2018 Oct;108(10):1378-1384. [doi: [10.2105/AJPH.2018.304567](https://doi.org/10.2105/AJPH.2018.304567)] [Medline: [30138075](https://pubmed.ncbi.nlm.nih.gov/30138075/)]
10. Tan ASL, Lee C, Chae J. Exposure to health (mis)information: lagged effects on young adults' health behaviors and potential pathways. *J Commun* 2015 Jul 06;65(4):674-698. [doi: [10.1111/jcom.12163](https://doi.org/10.1111/jcom.12163)]
11. Ratzan SC. Setting the record straight: vaccines, autism, and the Lancet. *J Health Commun* 2010 Apr 27;15(3):237-239. [doi: [10.1080/10810731003780714](https://doi.org/10.1080/10810731003780714)] [Medline: [20432106](https://pubmed.ncbi.nlm.nih.gov/20432106/)]
12. Lee JJ, Kang K, Wang MP, Zhao SZ, Wong JYH, O'Connor S, et al. Associations between COVID-19 misinformation exposure and belief with COVID-19 knowledge and preventive behaviors: cross-sectional online study. *J Med Internet Res* 2020 Nov 13;22(11):e22205 [FREE Full text] [doi: [10.2196/22205](https://doi.org/10.2196/22205)] [Medline: [33048825](https://pubmed.ncbi.nlm.nih.gov/33048825/)]
13. Luk TT, Zhao S, Weng X, Wong JY, Wu YS, Ho SY, et al. Exposure to health misinformation about COVID-19 and increased tobacco and alcohol use: a population-based survey in Hong Kong. *Tob Control* 2021 Nov;30(6):696-699. [doi: [10.1136/tobaccocontrol-2020-055960](https://doi.org/10.1136/tobaccocontrol-2020-055960)] [Medline: [32855353](https://pubmed.ncbi.nlm.nih.gov/32855353/)]
14. Wilson SL, Wiysonge C. Social media and vaccine hesitancy. *BMJ Glob Health* 2020 Oct;5(10):e004206 [FREE Full text] [doi: [10.1136/bmjgh-2020-004206](https://doi.org/10.1136/bmjgh-2020-004206)] [Medline: [33097547](https://pubmed.ncbi.nlm.nih.gov/33097547/)]
15. Pennycook G, Rand DG. Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proc Natl Acad Sci U S A* 2019 Feb 12;116(7):2521-2526 [FREE Full text] [doi: [10.1073/pnas.1806781116](https://doi.org/10.1073/pnas.1806781116)] [Medline: [30692252](https://pubmed.ncbi.nlm.nih.gov/30692252/)]
16. Lewandowsky S, Cook J, Ecker U, Albarracín D, Amazeen M, Kendeou P, et al. The Debunking Handbook 2020. *Skeptical Science*. 2020 Oct 14. URL: <https://sks.to/db2020> [accessed 2021-09-30]
17. Kozyreva A, Lewandowsky S, Hertwig R. Citizens versus the internet: confronting digital challenges with cognitive tools. *Psychol Sci Public Interest* 2020 Dec;21(3):103-156 [FREE Full text] [doi: [10.1177/1529100620946707](https://doi.org/10.1177/1529100620946707)] [Medline: [33325331](https://pubmed.ncbi.nlm.nih.gov/33325331/)]
18. Lazer DMJ, Baum MA, Benkler Y, Berinsky AJ, Greenhill KM, Menczer F, et al. The science of fake news. *Science* 2018 Mar 09;359(6380):1094-1096. [doi: [10.1126/science.aao2998](https://doi.org/10.1126/science.aao2998)] [Medline: [29590025](https://pubmed.ncbi.nlm.nih.gov/29590025/)]
19. Pennycook G, Epstein Z, Mosleh M, Arechar AA, Eckles D, Rand DG. Shifting attention to accuracy can reduce misinformation online. *Nature* 2021 Apr 17;592(7855):590-595. [doi: [10.1038/s41586-021-03344-2](https://doi.org/10.1038/s41586-021-03344-2)] [Medline: [33731933](https://pubmed.ncbi.nlm.nih.gov/33731933/)]
20. Gunter J. Medical misinformation and the internet: a call to arms. *Lancet* 2019 Jun 08;393(10188):2294-2295. [doi: [10.1016/S0140-6736\(19\)31206-1](https://doi.org/10.1016/S0140-6736(19)31206-1)] [Medline: [31180024](https://pubmed.ncbi.nlm.nih.gov/31180024/)]
21. Armstrong PW, Naylor CD. Counteracting health misinformation: a role for medical journals? *JAMA* 2019 May 21;321(19):1863-1864. [doi: [10.1001/jama.2019.5168](https://doi.org/10.1001/jama.2019.5168)] [Medline: [31009036](https://pubmed.ncbi.nlm.nih.gov/31009036/)]
22. Pagoto S, Waring ME, Xu R. A call for a public health agenda for social media research. *J Med Internet Res* 2019 Dec 19;21(12):e16661 [FREE Full text] [doi: [10.2196/16661](https://doi.org/10.2196/16661)] [Medline: [31855185](https://pubmed.ncbi.nlm.nih.gov/31855185/)]
23. O'Connor C, Murphy M. Going viral: doctors must tackle fake news in the covid-19 pandemic. *BMJ* 2020 Apr 24;369:m1587. [doi: [10.1136/bmj.m1587](https://doi.org/10.1136/bmj.m1587)] [Medline: [32332066](https://pubmed.ncbi.nlm.nih.gov/32332066/)]
24. Rubin R. Getting social: physicians can counteract misinformation with an online presence. *JAMA* 2019 Aug 20;322(7):598-600. [doi: [10.1001/jama.2019.10779](https://doi.org/10.1001/jama.2019.10779)] [Medline: [31365030](https://pubmed.ncbi.nlm.nih.gov/31365030/)]
25. Swire-Thompson B, Lazer D. Public health and online misinformation: challenges and recommendations. *Annu Rev Public Health* 2020 Apr 02;41(1):433-451. [doi: [10.1146/annurev-publhealth-040119-094127](https://doi.org/10.1146/annurev-publhealth-040119-094127)] [Medline: [31874069](https://pubmed.ncbi.nlm.nih.gov/31874069/)]
26. Vraga EK, Bode L. I do not believe you: how providing a source corrects health misperceptions across social media platforms. *Inf Commun Soc* 2017 Apr 19;21(10):1337-1353. [doi: [10.1080/1369118x.2017.1313883](https://doi.org/10.1080/1369118x.2017.1313883)]
27. Vraga EK, Bode L. Using expert sources to correct health misinformation in social media. *Sci Commun* 2017 Sep 14;39(5):621-645. [doi: [10.1177/1075547017731776](https://doi.org/10.1177/1075547017731776)]

28. Chan MS, Jones CR, Hall Jamieson K, Albarracín D. Debunking: a meta-analysis of the psychological efficacy of messages countering misinformation. *Psychol Sci* 2017 Nov;28(11):1531-1546 [[FREE Full text](#)] [doi: [10.1177/0956797617714579](https://doi.org/10.1177/0956797617714579)] [Medline: [28895452](#)]
29. Lewandowsky S, Ecker UKH, Seifert CM, Schwarz N, Cook J. Misinformation and its correction: continued influence and successful debiasing. *Psychol Sci Public Interest* 2012 Dec;13(3):106-131. [doi: [10.1177/1529100612451018](https://doi.org/10.1177/1529100612451018)] [Medline: [26173286](#)]
30. Walter N, Tukachinsky R. A meta-analytic examination of the continued influence of misinformation in the face of correction: how powerful is it, why does it happen, and how to stop it? *Commun Res* 2019 Jun 22;47(2):155-177. [doi: [10.1177/0093650219854600](https://doi.org/10.1177/0093650219854600)]
31. Vraga EK, Bode L. Correction as a solution for health misinformation on social media. *Am J Public Health* 2020 Oct;110(S3):S278-S280. [doi: [10.2105/AJPH.2020.305916](https://doi.org/10.2105/AJPH.2020.305916)] [Medline: [33001724](#)]
32. Bode L, Vraga EK. See something, say something: correction of global health misinformation on social media. *Health Commun* 2018 Sep;33(9):1131-1140. [doi: [10.1080/10410236.2017.1331312](https://doi.org/10.1080/10410236.2017.1331312)] [Medline: [28622038](#)]
33. Bode L, Vraga EK. In related news, that was wrong: the correction of misinformation through related stories functionality in social media. *J Commun* 2015 Jun 23;65(4):619-638. [doi: [10.1111/jcom.12166](https://doi.org/10.1111/jcom.12166)]
34. van der Meer TGLA, Jin Y. Seeking formula for misinformation treatment in public health crises: the effects of corrective information type and source. *Health Commun* 2020 May 14;35(5):560-575. [doi: [10.1080/10410236.2019.1573295](https://doi.org/10.1080/10410236.2019.1573295)] [Medline: [30761917](#)]
35. Gesser-Edelsburg A, Diamant A, Hijazi R, Mesch GS. Correcting misinformation by health organizations during measles outbreaks: a controlled experiment. *PLoS One* 2018 Dec 19;13(12):e0209505 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0209505](https://doi.org/10.1371/journal.pone.0209505)] [Medline: [30566485](#)]
36. Vraga EK, Bode L, Tully M. The effects of a news literacy video and real-time corrections to video misinformation related to sunscreen and skin cancer. *Health Commun* 2021 Apr 12:1-9. [doi: [10.1080/10410236.2021.1910165](https://doi.org/10.1080/10410236.2021.1910165)] [Medline: [33840310](#)]
37. Fischer CB, Adrien N, Silguero JJ, Hopper JJ, Chowdhury AI, Werler MM. Mask adherence and rate of COVID-19 across the United States. *PLoS One* 2021 Apr 14;16(4):e0249891 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0249891](https://doi.org/10.1371/journal.pone.0249891)] [Medline: [33852626](#)]
38. The CoronaVirusFacts/DatosCoronaVirus Alliance Database. Poynter. URL: <https://www.poynter.org/ifcn-covid-19-misinformation/> [accessed 2021-09-30]
39. Paynter J, Luskin-Saxby S, Keen D, Fordyce K, Frost G, Imms C, et al. Evaluation of a template for countering misinformation-Real-world Autism treatment myth debunking. *PLoS One* 2019;14(1):e0210746 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0210746](https://doi.org/10.1371/journal.pone.0210746)] [Medline: [30699155](#)]
40. Schmid P, Betsch C. Effective strategies for rebutting science denialism in public discussions. *Nat Hum Behav* 2019 Sep 24;3(9):931-939. [doi: [10.1038/s41562-019-0632-4](https://doi.org/10.1038/s41562-019-0632-4)] [Medline: [31235861](#)]
41. Lewandowsky S, van der Linden S. Countering misinformation and fake news through inoculation and prebunking. *Eur Rev Soc Psychol* 2021 Feb 22;32(2):348-384. [doi: [10.1080/10463283.2021.1876983](https://doi.org/10.1080/10463283.2021.1876983)]
42. Roozenbeek J, Schneider CR, Dryhurst S, Kerr J, Freeman ALJ, Recchia G, et al. Susceptibility to misinformation about COVID-19 around the world. *R Soc Open Sci* 2020 Oct 14;7(10):201199 [[FREE Full text](#)] [doi: [10.1098/rsos.201199](https://doi.org/10.1098/rsos.201199)] [Medline: [33204475](#)]
43. Cook J, Lewandowsky S, Ecker UKH. Neutralizing misinformation through inoculation: exposing misleading argumentation techniques reduces their influence. *PLoS One* 2017;12(5):e0175799 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0175799](https://doi.org/10.1371/journal.pone.0175799)] [Medline: [28475576](#)]
44. Lewandowsky S, Ecker UK, Cook J. Beyond misinformation: understanding and coping with the “post-truth” era. *J Appl Res Mem Cogn* 2017 Dec;6(4):353-369. [doi: [10.1016/j.jarmac.2017.07.008](https://doi.org/10.1016/j.jarmac.2017.07.008)]
45. Clapper JR. Facts and fears: hard truths from a life in intelligence. New York: Viking Penguin; 2018.
46. Nagler RH, Yzer MC, Rothman AJ. Effects of media exposure to conflicting information about mammography: results from a population-based survey experiment. *Ann Behav Med* 2019 Aug 29;53(10):896-908 [[FREE Full text](#)] [doi: [10.1093/abm/kay098](https://doi.org/10.1093/abm/kay098)] [Medline: [30596830](#)]
47. Nagler R, Vogel R, Gollust S, Yzer M, Rothman A. Effects of prior exposure to conflicting health information on responses to subsequent unrelated health messages: results from a population-based longitudinal experiment. *Ann Behav Med* 2021 Aug 16:kaab069. [doi: [10.1093/abm/kaab069](https://doi.org/10.1093/abm/kaab069)] [Medline: [34398961](#)]
48. Lee C, Nagler RH, Wang N. Source-specific exposure to contradictory nutrition information: documenting prevalence and effects on adverse cognitive and behavioral outcomes. *Health Commun* 2018 Apr 02;33(4):453-461 [[FREE Full text](#)] [doi: [10.1080/10410236.2016.1278495](https://doi.org/10.1080/10410236.2016.1278495)] [Medline: [28151010](#)]
49. Chang C. Motivated processing. *Sci Commun* 2015 Jul 30;37(5):602-634. [doi: [10.1177/1075547015597914](https://doi.org/10.1177/1075547015597914)]
50. Nan X, Daily K. Biased assimilation and need for closure: examining the effects of mixed blogs on vaccine-related beliefs. *J Health Commun* 2015 Apr 09;20(4):462-471. [doi: [10.1080/10810730.2014.989343](https://doi.org/10.1080/10810730.2014.989343)] [Medline: [25751250](#)]
51. Peer E, Brandimarte L, Samat S, Acquisti A. Beyond the Turk: alternative platforms for crowdsourcing behavioral research. *J Exp Soc Psychol* 2017 May;70:153-163. [doi: [10.1016/j.jesp.2017.01.006](https://doi.org/10.1016/j.jesp.2017.01.006)]

52. Palan S, Schitter C. Prolific.ac—A subject pool for online experiments. *J Behav Exp Finance* 2018 Mar;17:22-27. [doi: [10.1016/j.jbef.2017.12.004](https://doi.org/10.1016/j.jbef.2017.12.004)]
53. Fisher M, Knobe J, Strickland B, Keil FC. The influence of social interaction on intuitions of objectivity and subjectivity. *Cogn Sci* 2017 May 01;41(4):1119-1134. [doi: [10.1111/cogs.12380](https://doi.org/10.1111/cogs.12380)] [Medline: [27246257](https://pubmed.ncbi.nlm.nih.gov/27246257/)]
54. Sarkissian H, Park J, Tien D, Wright J, Knobe J. Folk moral relativism. *Mind Lang* 2011;26:482-505. [doi: [10.1111/j.1468-0017.2011.01428.x](https://doi.org/10.1111/j.1468-0017.2011.01428.x)]
55. Zhao X, Strasser A, Cappella JN, Lerman C, Fishbein M. A measure of perceived argument strength: reliability and validity. *Commun Methods Meas* 2011 Mar 04;5(1):48-75 [FREE Full text] [doi: [10.1080/19312458.2010.547822](https://doi.org/10.1080/19312458.2010.547822)] [Medline: [25568663](https://pubmed.ncbi.nlm.nih.gov/25568663/)]
56. McCroskey JC, Jenson TA. Image of mass media news sources. *J Broadcasting* 1975 Mar;19(2):169-180. [doi: [10.1080/08838157509363777](https://doi.org/10.1080/08838157509363777)]
57. Bruder M, Haffke P, Neave N, Nouripannah N, Imhoff R. Measuring individual differences in generic beliefs in conspiracy theories across cultures: conspiracy mentality questionnaire. *Front Psychol* 2013;4:225. [doi: [10.3389/fpsyg.2013.00225](https://doi.org/10.3389/fpsyg.2013.00225)] [Medline: [23641227](https://pubmed.ncbi.nlm.nih.gov/23641227/)]
58. Frederick S. Cognitive reflection and decision making. *J Econ Persp* 2005 Nov 01;19(4):25-42. [doi: [10.1257/089533005775196732](https://doi.org/10.1257/089533005775196732)]
59. Thomson K, Oppenheimer D. Investigating an alternate form of the cognitive reflection test. *Judgm Decis Mak* 2016;11(1):99-113 [FREE Full text]
60. The R Project for Statistical Computing Version 4. URL: <https://www.R-project.org/> [accessed 2022-02-19]
61. Pennycook G, Rand DG. Lazy, not biased: susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition* 2019 Jul;188:39-50. [doi: [10.1016/j.cognition.2018.06.011](https://doi.org/10.1016/j.cognition.2018.06.011)] [Medline: [29935897](https://pubmed.ncbi.nlm.nih.gov/29935897/)]
62. Al-Zaman MS. Prevalence and source analysis of COVID-19 misinformation in 138 countries. *IFLA J* 2021 Aug 27;034003522110411. [doi: [10.1177/03400352211041135](https://doi.org/10.1177/03400352211041135)]
63. Cinelli M, Quattrocioni W, Galeazzi A, Valensise CM, Brugnoti E, Schmidt AL, et al. The COVID-19 social media infodemic. *Sci Rep* 2020 Oct 06;10(1):16598. [doi: [10.1038/s41598-020-73510-5](https://doi.org/10.1038/s41598-020-73510-5)] [Medline: [33024152](https://pubmed.ncbi.nlm.nih.gov/33024152/)]
64. Bautista JR, Zhang Y, Gwizdka J. US physicians' and nurses' motivations, barriers, and recommendations for correcting health misinformation on social media: qualitative interview study. *JMIR Public Health Surveill* 2021 Sep 01;7(9):e27715 [FREE Full text] [doi: [10.2196/27715](https://doi.org/10.2196/27715)] [Medline: [34468331](https://pubmed.ncbi.nlm.nih.gov/34468331/)]
65. Johnson HM, Seifert CM. Sources of the continued influence effect: when misinformation in memory affects later inferences. *J Exp Psychol Learn Mem Cogn* 1994 Nov;20(6):1420-1436. [doi: [10.1037/0278-7393.20.6.1420](https://doi.org/10.1037/0278-7393.20.6.1420)]
66. Ecker UKH, Lewandowsky S, Tang DTW. Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Mem Cogn* 2010 Dec;38(8):1087-1100. [doi: [10.3758/mc.38.8.1087](https://doi.org/10.3758/mc.38.8.1087)]
67. Skurnik I, Yoon C, Park D, Schwarz N. How warnings about false claims become recommendations. *J Consum Res* 2005 Mar;31(4):713-724. [doi: [10.1086/426605](https://doi.org/10.1086/426605)]
68. Schwarz N, Newman E, Leach W. *Behav Sci Policy* 2016;2(1):85-95 [FREE Full text] [doi: [10.1353/bsp.2016.0009](https://doi.org/10.1353/bsp.2016.0009)]
69. 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 Sep 20;63(3):393-414. [doi: [10.1080/08838151.2019.1653102](https://doi.org/10.1080/08838151.2019.1653102)]
70. Vraga EK, Bode L. Addressing COVID-19 misinformation on social media preemptively and responsively. *Emerg Infect Dis* 2021 Feb;27(2):396-403. [doi: [10.3201/eid2702.203139](https://doi.org/10.3201/eid2702.203139)] [Medline: [33395379](https://pubmed.ncbi.nlm.nih.gov/33395379/)]

Abbreviations

ANCOVA: analysis of covariance

CDC: Centers for Disease Control and Prevention

CRT: Cognitive Reflection Test

M: misinformation-only experimental condition

MC: misinformation+correction experimental condition

MCR: misinformation+correction+rebuttal experimental condition

MCRC: misinformation+correction+rebuttal+second correction experimental condition

Edited by M Gisondi, J Faust; submitted 09.11.21; peer-reviewed by L Bode, M Waring, S Pesälä; comments to author 27.12.21; revised version received 08.01.22; accepted 13.02.22; published 02.03.22

Please cite as:

Mourali M, Drake C

The Challenge of Debunking Health Misinformation in Dynamic Social Media Conversations: Online Randomized Study of Public Masking During COVID-19

J Med Internet Res 2022;24(3):e34831

URL: <https://www.jmir.org/2022/3/e34831>

doi: [10.2196/34831](https://doi.org/10.2196/34831)

PMID: [35156933](https://pubmed.ncbi.nlm.nih.gov/35156933/)

©Mehdi Mourali, Carly Drake. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 02.03.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.