Journal Pre-proof

Refuting Spurious COVID-19 Treatment Claims Reduces Demand and Misinformation Sharing<!–<RunningTitle>REFUTING COVID-19 MISINFORMATION</RunningTitle>–>

Douglas MacFarlane, Li Qian Tay, Mark J. Hurlstone, Ullrich K.H. Ecker

Please cite this article as: { doi: https://doi.org/

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Refuting Spurious COVID-19 Treatment Claims Reduces Demand and Misinformation Sharing

Douglas MacFarlane[1^], Li Qian Tay[1^], Mark J. Hurlstone[2], & Ullrich K. H. Ecker[1*]

[1] School of Psychological Science, University of Western Australia, Australia

[2] Department of Psychology, Lancaster University, UK

[^] These authors contributed equally

* Corresponding author

Word count: 5,421 (Main text incl. figure captions and tables); 2,265 (Introduction and Discussion)

Corresponding author: Ullrich Ecker, UWA M304, 35 Stirling Hwy, Perth 6009, Australia; ullrich.ecker@uwa.edu.au

Abstract

The COVID-19 pandemic has seen a surge of health misinformation, which has had serious consequences including direct harm and opportunity costs. We investigated (N = 678) the impact of such misinformation on hypothetical demand (i.e., willingness-to-pay) for an unproven treatment, and propensity to promote (i.e., like or share) misinformation online.

This is a novel approach, as previous research has used mainly questionnaire-based measures of reasoning. We also tested two interventions to counteract the misinformation, contrasting a tentative refutation based on materials used by health authorities with an enhanced refutation based on best-practice recommendations. We found prior exposure to misinformation increased misinformation promotion (by 18%). Both tentative and enhanced refutations reduced demand (by 18% and 25%, respectively) as well as misinformation promotion (by 29% and 55%). The fact that enhanced refutations were more effective at curbing promotion of misinformation highlights the need for debunking interventions to follow current best-practice guidelines.

Keywords: Health misinformation; refutations; willingness to pay; online sharing; coronavirus

General Audience Summary

Health misinformation proliferates online, especially during a crisis such as the COVID-19 pandemic, when demand for effective treatments is pronounced. Such misinformation has the potential to cause harm; for example, a promoted treatment might be unproven or ineffective but have unintended side effects or prevent uptake of superior interventions (e.g., using a non-prescribed vitamin supplement to prevent viral infection while abstaining from social distancing or mask use). Combating misinformation about medical treatments is therefore a particularly exigent issue. Alas, counteracting misinformation is a non-trivial task, with psychological research demonstrating that even clearly-corrected misinformation can continue to influence reasoning and decision making. The present study targeted misinformation about high-dose vitamin E as a potent remedy for COVID-19 prevention and treatment. It shows that refutations that follow best-practice guidelines from psychological research are more effective than tentative refutations often employed by health authorities or

the media, reducing both demand for the product and the propensity to share related

misinformation online.

Refuting Spurious COVID-19 Treatment Claims Reduces Demand and Misinformation
Sharing

Misinformation—defined here as any information that is false—represents a threat to

societies that value evidence-based practice and policy making (Lazer et al., 2018;

Lewandowsky, Ecker, & Cook, 2017). Smear tactics and political fake news have affected

voter attitudes (Landon-Murray, Mujkic, & Nussbaum, 2019; Vaccari & Morini, 2014),

vaccine myths have contributed to the re-emergence of measles (Poland, Jacobson, &

Ovsyannikova, 2009), and climate change misinformation continues to pose a barrier to

mitigative action (Cook, Ellerton, & Kinkead, 2018).

COVID-19 is another case in point. At a time when demand for high-quality

information has been pronounced, online misinformation has proliferated—both as a result of

the rapidly-developing nature of the pandemic, where even the best available evidence of

today may be invalidated tomorrow, and due to the actions of ill-informed or malicious

disinformants. COVID-19 misinformation has included claims that the virus originated as a

bioweapon, or that health supplements or even onions could be effective treatments (e.g.,

Dupuy, 2020). The spread of such misinformation can fuel hostility towards groups perceived

to be responsible for the pandemic (Devakumar, Shannon, Bhopal, & Abubakar, 2020); it can

also draw people towards remedies that are unproven, ineffective, or even harmful, and

prevent uptake of superior interventions (e.g., using a vitamin supplement to prevent viral

infection while abstaining from social distancing, mask use, or vaccination), at variance with

official health advice (Roozenbeek et al., 2020; Van Bavel et al., 2020). Accordingly, the

World Health Organization (WHO) noted that the pandemic has been accompanied by an

*info*demic, characterized by a deluge of false information amplified by modern

communication technologies (Zarocostas, 2020). Given the scale of the problem,

considerable efforts are being undertaken to fact-check common claims (e.g., Therapeutic

Goods Administration, 2020; Weule, 2020). However, it is unclear how successful these

efforts will be, and to what extent individuals initially misled by misinformation will adjust

their beliefs and behaviours after receiving a refutation.

Cognitive psychology research suggests refuting misinformation is no easy task. Even

clearly corrected misinformation often continues to influence memory, reasoning, and

decision making—the so-called continued influence effect (Chan, Jones, Hall Jamieson, &

Albarracín, 2017; Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012; Walter &

Tukachinsky, 2020). Post-correction reliance on misinformation can occur despite

demonstrable memory for the correction; it can even arise with fictitious or trivial

information. This suggests that continued influence is a cognitive effect that does not

necessarily depend on individuals' motivation—it can emerge from failures of memory

updating or retrieval processes alone (Rich & Zaragoza, 2016; Swire, Ecker, &

Lewandowsky, 2017).

Much of the literature on continued influence has focused on how misinformation is

best counteracted (Ecker, O'Reilly, Reid, & Chang, 2019; Paynter et al., 2019; Walter &

Tukachinsy, 2020). Based on this research, several best-practice recommendations have been

identified. Specifically, refutations should: (a) come from a trustworthy source (Ecker &

Antonio, 2020; Guillory & Geraci, 2013) and, if applicable, discredit the disinformant by

exposing their hidden agenda (Walter & Tukachinsky, 2020); (b) make salient the

discrepancy between false and factual information, which has been shown to facilitate knowledge revision (Ecker, Hogan, & Lewandowsky, 2017; Kendeou, Butterfuss, Kim, & van Boekel, 2019); (c) explain why the misinformation is false, providing factual information to replace the false information in people's mental models (Paynter et al., 2019; Swire et al., 2017); and (d) draw attention towards any misleading strategies employed by misinformants (Cook et al., 2018; MacFarlane, Hurlstone, & Ecker, 2018, 2020a).

Regrettably, real-world refutation attempts often do not take into account these recommendations. For example, it is common for scientists and official bodies to adopt a tentative, "diplomatic" approach in their communication (MacFarlane et al., 2018, 2020a; Paynter et al., 2019). To illustrate, to counter misinformation regarding the use of vitamin C for COVID-19, the Australian Therapeutic Goods Administration (TGA) wrote, "there is no robust scientific evidence to support the usage of [high dose vitamin C] in the management of COVID-19" (TGA, 2020). Such a diplomatic approach to refutation may sometimes be called for to express nuance, and this may be of particular concern during a pandemic characterized by much uncertainty (e.g., Chater, 2020). However, it is important to recognize that tentative refutations leave open the possibility that misleading information may turn out to be true, which can then be exploited by individuals or organizations with ulterior motives (MacFarlane, Hurlstone, & Ecker, 2020b; Oreskes & Conway, 2010). It is therefore clear that refutations of health misinformation need to be well-designed to both provide an accurate reflection of the best-available evidence and achieve the desired outcome, while also taking into account the cognitive biases that disinformants—such as health fraudsters trying to sell a product—seek to exploit.

Arguably the greatest weakness of existing research into the continued influence effect is its near-universal reliance on questionnaire measures of reasoning. This is an issue because some of the most harmful effects of misinformation may be on behaviours. For this

reason, a more consequential indicator of the success of a particular refutation strategy is

consumer behaviour. The need for objective behavioural indicators is particularly acute in the

context of COVID-19, where the combination of high anxiety levels and need for effective

treatment may foster misinformation-driven demand for remedies that are ineffective and

often dangerous, including bleach, methylated spirits, or essential oils (e.g., Spinney, 2020;

note, however, that for ethical reasons we selected a product for the current study that is

unproven as a COVID-19 treatment but relatively safe). Given the well-established attitude-

behaviour gap (e.g., McEachan, Conner, Taylor, & Lawton, 2011), it is critical that studies

take steps towards investigating how misinformation affects behaviour, and to what extent

corrections can undo those effects (Hamby, Ecker, & Brinberg, 2020).

Against this backdrop, and building on previous theoretical work (MacFarlane et al.,

2020b), our aim was to investigate the following two questions: (1) How does exposure to

COVID-19 misinformation influence people's (hypothetical) willingness to pay for unproven

health products, as well as their subsequent propensity to spread misinformation? (2) Are

enhanced refutations based on best-practice psychological insights better suited to counteract

the behavioural impacts of misinformation, compared to the tentative refutations often used

by authorities?

To this end, we exposed participants to misinformation[1] suggesting that high-dose

vitamin E can reliably prevent and cure COVID-19. Although vitamins are essential for

health, there is ample evidence that vitamin supplementation is contra-indicated in

individuals with no pre-existing deficiencies (Guallar 2016), and the present claim that

vitamin E will reliably cure COVID-19 is demonstrably false (Shakoor et al., 2021). The

misinformation employed a number of deceptive techniques used by real-life disinformants

---

[1] The term *disinformation*—false information crafted and disseminated with the intent to
deceive—may be more appropriate here; however, we decided to use the broader term as ill
intent cannot be assumed for all instances of advocacy for alternative health remedies.

(MacFarlane et al., 2020b). We then provided participants with either no refutation, a

tentative refutation based on reputable real-world sources, or an enhanced refutation based on

best-practice psychological insights (MacFarlane et al., 2020a, 2020b; Paynter et al., 2019;

Walter & Tukachinsky, 2020). An additional control group received only an article that

provided general information about vitamin E. Instead of using the questionnaire approach

typical of misinformation debunking studies, we examined participants' willingness to pay

for a vitamin E supplement in a hypothetical auction, and their propensity to promote a mock

social media message endorsing the use of vitamin E to treat COVID-19. We hypothesized

that misinformation exposure would increase willingness to pay and misinformation

promotion, and that the enhanced refutation would more effectively reverse this effect than

the tentative refutation.

Method

The experiment adopted a between-participants design with four conditions (control

vs. misinformation vs. tentative refutation vs. enhanced refutation), with two dependent

variables: "willingness-to-pay"—the amount bid on the vitamin E supplement—and

"misinformation-promotion"—a score derived from a participant's engagement with the

misleading social media post. Participants' general attitudes towards health supplements and

alternative medicines, as well as their concern about the COVID-19 pandemic were used as

covariates; additional predictors included frequency of past vitamin E supplement

consumption, belief in the effectiveness of routine vitamin E supplementation, and current

state of health.

Participants

A sample of 680 U.S.-based adult participants were recruited via Amazon's

Mechanical Turk (MTurk). To ensure data quality, participants had to meet eligibility criteria

of 97% approval rate and a minimum of 5,000 prior tasks completed. An a-priori power

analysis suggested a minimum sample size of 309 to detect an effect of $f = .16$ (based on the contrast between control and "full-contingency plus" conditions in MacFarlane et al., 2018) in our primary contrast (misinformation vs. enhanced refutation) with $df = 1$, $\alpha = .05$, $1 - \beta = .80$. Total sample size of 680 was calculated by extrapolating this to incorporate four conditions and an estimated 10% exclusion rate. Participants were randomly allocated to one of the four intervention conditions, subject to the constraint of approximately equal cell sizes.

Two a-priori (though not formally pre-registered) exclusion criteria were applied to remove careless responders. One participant was excluded for giving non-differentiated answers to every question in a survey block, and another for responding erratically, with overly inconsistent responses between pairs of equivalent questions (i.e., an odd/even threshold of > 2 Likert points apart). Final sample size was thus $N = 678$ (344 females, 328 males, 3 non-binary, 3 participants of undisclosed gender; age $M = 42.81$, $SD = 12.64$).

Materials

Predictors. Prior beliefs can moderate responses to health messaging (Myers, 2014). Thus, participants' general attitudes towards health supplements and alternative medicines were assessed using an 18-item questionnaire, following MacFarlane et al. (2018). Each item consisted of a declarative statement relating to a motivation for consuming alternative health products (e.g., "Vitamins are natural, and supplements are therefore safe"). Participants were asked to rate how much they agreed with each statement using a 5-point Likert scale (1 = strongly disagree; 5 = strongly agree). A composite score was calculated for each participant indicating their general attitude to health supplements and alternative medicines (hereafter, "general-attitude"). To measure response consistency, each item was paired with a reverse-phrased statement of similar meaning (i.e., 9 pairs of items). The order of items in the scale was randomized. The general-attitude scale was found to have very good internal consistency

reliability (Cronbach's α = .87). The full scale can be found in the Supplement, which is

available at https://osf.io/p89bm/.

Participants' concern about the COVID-19 pandemic was measured as a composite

score based on participants' responses to four questions (e.g., "How severe do you think

novel coronavirus [COVID-19] will be in the U.S. general population as a whole?" on a 5-

point Likert scale [1 = very mild; 5 = very severe]), following Ecker, Butler, Cook,

Hurlstone, Kurz, & Lewandowsky (2020). All items are provided in the Supplement.[2]

Interventions. Dependent on condition, participants were presented with one or two

articles (all articles are provided in Supplement). The *control* condition used a news article

emphasizing the importance of a healthy immune system for fighting COVID-19. The article

was modelled on two pieces (Collins, 2020; Weule, 2020) from reputable and independent

media outlets, The Conversation and the Australian Broadcasting Corporation. It highlighted

the importance of maintaining a balanced diet, regular exercise, and good sleep. To provide a

background as to why participants were bidding on a vitamin E supplement, the article

included a quote from the U.S. National Institute of Health "Vitamins and Minerals" website,

"*Vitamins and minerals are…essential for health. The lack of a specific micro-nutrient may*

cause…disease. Vitamin E, for example, acts as an antioxidant and is involved in immune

function. It helps to widen blood vessels and keep blood from clotting. In addition, cells use

Vitamin E to interact with each other and to carry out many important functions." The

control article was therefore not fully neutral but contained some allusion that vitamin E

supplementation might provide some protection from viral infection. The article also noted

that whilst some people believe that vitamin supplements are essential, dietitians say that

---

[2] Three additional one-item predictors regarding previous vitamin E consumption, efficacy belief, and current state of health were included merely as a "sanity check" and for comparability with previous studies (MacFarlane et al., 2018; 2020); the predictors and some ancillary correlational analyses are reported in the Supplement for the interested reader.

taking supplements is not necessary unless you have a specific nutrient deficiency or dietary need.

The article in the *misinformation* condition additionally contained false claims suggesting vitamin E can protect against COVID-19. These were modelled on misinformation spread via social media during the early stages of the pandemic (Bogle, 2020; RMIT Fact Check, 2020; Weule, 2020). We invented a fake expert, Dr. Avery Clarke, and used several deceptive techniques employed by real-life fraudsters (MacFarlane et al., 2020b; see Table 1 for details).

In the *tentative-refutation* condition, the misinformation article was followed by a second article that highlighted the lack of evidence for claims about vitamin E and COVID-19. It was based on a tentative, "diplomatic" refutation style commonly used in real-world attempts to debunk claims about ineffective or unproven health products; specifically, it was modelled on an article featuring interviews with a disease expert and a dietitian (Weule, 2020). We invented a public health expert, Prof. Simon Corner, and outlined his response to Dr. Avery's claims, *"There is no conclusive evidence that vitamin [E] supplements can delay the onset of an infection or treat respiratory infections, such as COVID-19. Furthermore, vitamin and mineral supplements are not recommended for the general population."* The final paragraph noted that there were some exceptions (e.g., people with specific vitamin deficiencies) and recommended people talk to a doctor, pharmacist, or accredited dietitian.

In the *enhanced-refutation* condition, the misinformation article was followed by a refutational article that, instead of simply noting lack of evidence, drew attention to deceptive and misleading techniques used in the misinformation article and explained how those techniques were used to deceive readers. This enhanced refutation was based on psychological insights regarding the effective debunking of misinformation (MacFarlane et al., 2020a, 2020b; Paynter et al., 2019; Walter & Tukachinsky, 2020; see Table 2 for details).

Dependent measures. We included two novel tasks to measure participants'

hypothetical willingness to pay for a spurious treatment and their propensity to spread related

misinformation.

Willingness-to-pay. To assess willingness-to-pay, we used an experimental auction

(MacFarlane et al., 2018; also see Becker, DeGroot, & Marschak, 1964; Kagel & Levin,

2011; Thrasher, Rousu, Hammond, Navarro, & Corrigan, 2011). The auction was

hypothetical but previous experiments have shown that results obtained using this

hypothetical auction format are comparable to those under fully-incentivized conditions (i.e.,

when bidding using real money for real products; MacFarlane et al., 2020a). The decision to

use a hypothetical format was made for both practicality reasons and ethical concerns about

selling an unproven COVID-19 treatment during the pandemic. First, participants were asked

to imagine they had been given a $5 endowment. They then had an opportunity to place a bid

on a bottle of 100 vitamin E capsules. Participants were shown a plain-packaged picture of

the product and a generic text (e.g., "…The supplement is designed to be taken once a day.").

It was explained that the auction was different from other auctions in that participants could

only bid once, and it was in their best interest to bid the amount they were willing to pay for

the product. Participants entered their bid amount b in cents, with $b \in (0, 500)$. They knew

this amount would be compared against a random number $r \in (0, 500)$ drawn from a uniform

distribution, and that if $b \geq r$, they would win the auction and purchase the product for

amount b but keep $500 - b$ of their endowment; otherwise they would lose the auction but

keep the full hypothetical endowment. There was an initial practice auction to ensure

participants understood the procedure. At the end of the auction, they were informed about

the outcome, and were shown their bid amount, the random comparison amount, and their

"take-home" amount. To mitigate hypothetical bias, we employed two established bias-

reduction techniques, namely "cheap-talk"—asking participants to behave as they would if

the auction was real—and consequentiality—reminding participants the results of this study

would have implications for significant public health issues.

Misinformation-promotion. To assess participants' propensity to promote

misinformation via social media, we developed a novel measure. Participants were presented

with four social media posts, and they were asked to indicate how they would engage with

each post using one of three options, namely "share", "like", or "flag". Each option was

accompanied by a brief explanation: Participants were told they could "Share the post

publicly, so more people could read it"; "Like the post, so your contacts see that you agree

with it"; or "Flag the post if you think it is inappropriate (e.g., offensive, inaccurate,

misleading, or promoting illegal activity)". Participants were also told they could choose not

to interact with a post, thus enabling "pass" as a fourth response option.

The posts are shown in Figure 1. Three posts were on topics generally related to

health and/or COVID-19, whereas the fourth contained an endorsement of misinformation,

namely that vitamin E can treat COVID-19. Interest centred on responses to this target post,

which was always presented last. The general posts were presented first (in random order) as

decoys to help manage demand characteristics. To ensure plausibility, two of the decoy posts

were real tweets from the Wall Street Journal and the World Health Organization. To avoid a

slight confound that only the target post was from an individual doctor, the third decoy post

(relating to a tweet from New Scientist) was also from an invented doctor.

Procedure

Ethics approval was granted by the Human Research Ethics Office of the University

of Western Australia. The experiment was run using Qualtrics software (Qualtrics, Provo,

UT). Participants were initially given an ethics-approved information sheet and provided

informed consent. They then responded to questions on demographics, vitamin E

consumption, the general-attitude scale, and the COVID-19 questions. Participants then read

the article(s) associated with the condition they were assigned to (i.e., control,

misinformation, tentative-refutation, or enhanced-refutation). This was followed by the

hypothetical auction and the mock social media engagement task, in which participants bid

for a bottle of vitamin E tablets and interacted with a series of posts (response options: share,

like, flag, pass). Finally, participants were debriefed; this included a detailed refutation of the

misinformation participants had encountered (following the procedure in the enhanced-

refutation condition), and participants had to demonstrate their understanding by correctly

answering a question regarding the gist of the debriefing. The experiment took approximately

17 minutes; participants were paid US$2.50.

Results

Willingness-to-Pay

On a descriptive level, willingness-to-pay means across control (C), misinformation

(M), tentative-refutation (TR), and enhanced-refutation (ER) conditions were, in cents:

$M_C = 208.27$ ($n = 170$, $SE = 11.05$), $M_M = 221.75$ ($n = 171$, $SE = 12.19$), $M_{TR} = 181.97$

($n = 168$, $SE = 11.83$), and $M_{ER} = 167.28$ ($n = 169$, $SE = 12.33$). Thus, the refutations reduced

willingness-to-pay by 18% and 25%, relative to the misinformation condition.

We conducted an analysis of covariance (ANCOVA) with a four-level condition

factor, willingness-to-pay as the dependent variable, and general attitude towards health

supplements (hereafter: general attitude) and COVID-19 concern as covariates. There was a

significant main effect of condition, $F(3,672) = 5.76$, $p < .001$, $\omega^2 = .02$ (see Figure 2). There

was also a significant effect of general attitude, $F(1,672) = 150.45$, $p < .001$, $\omega^2 = .18$,

indicating that greater support for supplements was associated with greater willingness-to-

pay. There was no significant effect of COVID-19 concern, $F < 1$.

Tukey post-hoc tests revealed mean willingness-to-pay was greater in the

misinformation condition than the tentative-refutation condition ($\Delta = 43.86$, 95% CI 4.73 –

82.98, $t = 2.89$, $p = .021$, $d = 0.28$) and the enhanced-refutation condition ($\Delta = 59.46$, 95% CI

20.41 − 98.51, $t = 3.92$, $p < .001$, $d = 0.37$). There were no significant differences between the

other conditions (see Table 3 for details).[3]

Misinformation-Promotion

In terms of participant proportions across response categories and conditions, 50.0%

of participants in the control condition engaged in liking or sharing of the misinformation-

endorsing social media post. This increased to 59.1% in the misinformation condition (i.e., an

18.1% increase) but dropped to 41.7% and 26.6% in tentative-refutation and enhanced-

refutation conditions, respectively (i.e., 29.5% and 54.9% reductions relative to the

misinformation condition). See Table 4 for further details.

To formally evaluate participants' propensity to promote the misinformation post, we

converted the four response options into a single measure ("misinformation-promotion"). We

re-coded each response option to reflect the progression towards an increasing propensity to

spread misinformation: flag = -1 (actively combatting the spread of misinformation); pass = 0

(not contributing to the spread of misinformation); like = 1 (signalling approval of

misinformation and increasing its salience); share = 2 (actively spreading misinformation).

An asymptotic Kruskal-Wallis test found that condition significantly influenced

misinformation-promotion, $H(3) = 54.25$, $p < .001$ (see Figure 3). [4]

To assess the locus of condition differences, post-hoc Mann-Whitney U tests with

Bonferroni adjustment were run. Misinformation-promotion was significantly greater for

---

[3] The Supplement reports additional exploratory analyses, namely an ANOVA without the covariates, as well as a full-factorial ANCOVA. Results were comparable, although the difference in willingness-to-pay between misinformation and tentative refutation conditions in the ANOVA post-hoc test became non-significant.

[4] The Supplement also reports a model comparison based on an ordinal regression approach, which demonstrates that the condition main effect also arises in an analysis including the covariates of general attitude (which was also a significant predictor of misinformation promotion) and COVID concern.

participants in the misinformation condition, *Mdn* = 1, compared to control, *Mdn* = 0.5,

Ws = 12434, *p* = .017, *r* = -.09, the tentative-refutation condition, *Mdn* = 0, *Ws* = 17712,

p < .001, *r* = -.15, and the enhanced-refutation condition, *Mdn* = -1, *Ws* = 20566, *p* < .001,

r = -.27. Participants in the enhanced-refutation condition showed significantly lower

misinformation-promotion than those in the tentative-refutation condition, *Ws* = 17386,

p < .001, *r* = -.15, and the control condition *Ws* = 18494, *p* < .001, *r* = -.19. There was no

significant difference between the control condition and the tentative-refutation condition,

Ws = 15473, *p* = .168, *r* = -.05.

Discussion

      The present study is one of the first to incorporate behavioural measures into the

continued influence paradigm. Specifically, we tested whether exposure to misinformation

that presented vitamin E as a potent remedy to prevent and cure COVID-19 affected

(hypothetical) product demand and misinformation promotion behaviour. To test the efficacy

of interventions counteracting the misinformation, we contrasted a tentative refutation based

on real-world sources and an enhanced refutation based on best-practice recommendations.

We included general attitude towards supplements and COVID-19 concern as covariates.

      We found that pre-existing attitudes predicted demand and propensity to promote

misinformation, whereas we found no effects of COVID-19 concern. Even though this is

perhaps surprising, we refrain from drawing strong conclusions based on this null finding; it

is of course possible that degree of concern has little influence on the processing of related

misinformation and subsequent decisions (similarly, Ecker, Lewandowsky, & Apai, 2010,

reported limited impact of emotionality on the continued influence effect).

      More importantly, with regards to our experimental manipulation, we found that

exposure to misinformation significantly increased participants' subsequent propensity to

promote a mock social media post endorsing the misinformation. Although willingness to pay

was not reliably affected by the misinformation relative to control, this is best explained by the control article's allusion that vitamins provide some protection from viral infection, meaning that the control condition arguably did not provide a fully-neutral comparison baseline. Compared to the misinformation condition, both refutation types substantially reduced willingness-to-pay and misinformation-promotion, which underscores the general utility of refutations beyond the realm of inferential reasoning measures. The enhanced refutation was more effective than the tentative refutation in reducing misinformation promotion,[5] reinforcing the best-practice recommendations used.

Our results demonstrate the persuasive appeal of (COVID-19) misinformation and its potential to influence behaviours (Zarocostas, 2020; see also Pennycook, McPhetres, Zhang, Lu, & Rand, 2020). They also demonstrate the value of an evidence-based approach to debunking. This is in line with previous research (Paynter et al., 2019; Swire et al., 2017) but one of the first demonstrations of this kind with behavioural measures (also see Hamby et al., 2020). This is of particular relevance given (1) the need for studies to take steps towards addressing the attitude-behaviour gap (McEachan et al., 2011), and (2) previous work questioning the utility of refutations when it comes to behaviours (Swire-Thompson, DeGutis, & Lazer, 2020; Swire-Thompson, Ecker, Lewandowsky, & Berinsky, 2020).

The practical implications are clear. Individuals exposed to refutations are less likely to waste money on spurious products. Also, individuals exposed to misinformation may be more likely to later on spread it to those around them, but this harmful sharing behaviour can be mitigated through refutations. Although sharing decisions may also be affected by factors other than veracity (e.g., inattention, worldview; Mercier; 2020; Pennycook et al., 2020), the present study suggests that well-designed refutations are an indispensable tool. The fact that

---

[5] Again, the effect of the tentative refutation on willingness-to-pay was significant only if covariates were included.

our enhanced refutation was more effective at reducing misinformation promotion illustrates the need for debunking practices to be based on insights from psychology (MacFarlane et al., 2020b). This will be especially critical during crises such as the current infodemic.

Nonetheless, limitations must be acknowledged. First, our enhanced refutation included many elements and it is unclear which ones contributed to the enhanced refutation's success. There may also be nuisance factors such as perceived confidence in evidence, memorability, and article length that differentiated tentative and enhanced refutations. Further research should thus include component analyses to aid the construction of specific best-practice guidelines (see also Lewandowsky et al., 2020; Paynter et al., 2019). This would also help add theoretical nuance beyond the scope of the current study, furthering understanding of why certain interventions work better than others.

Second, we relied only on a hypothetical auction and a mock social media scenario to assess behaviours. Although an important first step, willingness-to-pay methods may display substantial heterogeneity despite bias-reduction techniques (e.g., Kanya, Sanghera, Lewin, & Fox-Rushby, 2019; Voelckner, 2006), and the same may apply to our social media scenario. It is therefore critical that future studies use measures that correspond with behaviours of interest even more closely (e.g., incentivized auctions, social-media simulations) and validate results using real-world data (e.g., Mosleh, Pennycook, & Rand, 2020).

Third, we did not find a difference in willingness-to-pay between the enhanced and tentative refutation conditions, suggesting that the auction measure may not have been sufficiently sensitive. Potentially related to this, we also did not find a reliable willingness-to-pay difference between misinformation and control conditions. However, consistent with effects of implied and subtle misinformation (e.g., Ecker, Lewandowsky, Chang, & Pillai, 2014; Powell, Keil, Brenner, Lim, & Markman, 2018; Rich & Zaragoza, 2016), this may have also been driven by the insinuation in the control article that vitamins can provide some

infection protection. This suggests that the framing of "neutral" educational communications by impartial media sources needs to be carefully considered. The fact that the control condition was non-neutral means there was no neutral no-misinformation baseline; this makes it impossible to conclude with certainty whether there was continued influence (i.e., above-baseline reliance on refuted misinformation) in either of the refutation conditions.

Furthermore, there was no time lag between misinformation and correction, nor between interventions and behavioural tasks. This may constrain applicability of current findings for real-world scenarios; future research should explore the impact of longer time delays (see Paynter et al., 2019).

Finally, we recommend that communicators dealing with misinformation consider additional strategies beyond refutations. The speed of misinformation generation and spread presents a challenge to fact-checkers and their capacity to refute misleading content (Lazer et al., 2018). Future work should thus examine inoculation strategies, which pre-emptively explain misleading techniques common to misinformation campaigns and thus increase resilience to misdirection (e.g., Basol, Roozenbeek, & van der Linden, 2020; Comptom, 2013; Cook, Lewandowsky, & Ecker, 2017; Maertens, Roozenbeek, Basol, & van der Linden; 2020; van der Linden, Leiserowitz, Rosenthal, & Maibach, 2017). In particular, more research is needed to assess the impact of inoculations on behaviours.

Supplement

Both the data and a Supplement containing the materials and supplementary analyses are available at https://osf.io/p89bm/.

supplementary analyses. LQT and DM drafted the manuscript, and MH and UKHE

provided critical revisions. All authors approved the final version of the manuscript

for submission.

Data Statement: Both the data and a Supplement containing the materials are available at

https://osf.io/p89bm/.

References

Basol, M., Roozenbeek, R., & van der Linden., S. (2020). Good News about Bad News:
Gamified Inoculation Boosts Confidence and Cognitive Immunity Against Fake
News. *Journal of Cognition, 3*, 2. https://doi.org/10.5334/joc.91

Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring utility by a single-
response sequential method. *Behavioral Science, 9*, 226-232.
https://doi.org/10.1002/bs.3830090304

Bogle, A. (2020, March 21). *As coronavirus fears grow, family group chats spread support
but also misinformation.* ABC News. https://www.abc.net.au/news/science/2020-03-
21/coronavirus-health-misinformation-spreading-whatsapp-text-groups/12066386

Chan, M. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). Debunking: A meta-
analysis of the psychological efficacy of messages countering misinformation.
Psychological Science, 28, 1531-1546. https://doi.org/10.1177/0956797617714579

Chater, N. (2020). Facing up to the uncertainties of COVID-19. *Nature Human Behaviour, 4*,
439-439. https://doi.org/10.1038/s41562-020-0865-2

Collins, C. (2020, March 17). *5 ways nutrition could help your immune system fight off the
coronavirus*. The Conversation. https://theconversation.com/5-ways-nutrition-could-
help-your-immune-system-fight-off-the-coronavirus-133356

Compton, J. (2013). Inoculation theory. *The Sage Handbook of Persuasion: Developments in
Theory and Practice*, *2*, 220-237.

Cook, J., Ellerton, P., & Kinkead, D. (2018). Deconstructing climate misinformation to
identify reasoning errors. *Environmental Research Letters, 13,* 024018.
https://doi.org/10.1088/1748-9326/aaa49f

Cook, J., Lewandowsky, S., & Ecker, U. K. H. (2017). Neutralizing misinformation through

inoculation: Exposing misleading argumentation techniques reduces their influence.

PLOS ONE, 12, e0175799. doi:10.1371/journal.pone.0175799

Devakumar, D., Shannon, G., Bhopal, S., & Abubakar, I. (2020). Racism and discrimination

in COVID-19 responses. *The Lancet, 395*, 1194-1194. https://doi.org/10.1016/S0140-

6736(20)30792-3

Dupuy, B (2020, May 13). *No evidence onions can cure ailments or kill viruses.* AP News.

https://apnews.com/article/8905130246

Ecker, U. K. H., & Antonio, L. M. (2020). Can you believe it? An investigation into the

impact of retraction source credibility on the continued influence effect. *Memory &

Cognition*.

Ecker, U. K. H., Butler, L. H., Cook, J., Hurlstone, M. J., Kurz, T., & Lewandowsky, S.

(2020). Using the COVID-19 economic crisis to frame climate change as a secondary

issue reduces mitigation support. *Journal of Environmental Psychology, 70,* 101464.

https://doi.org/10.1016/j.jenvp.2020.101464

Ecker, U. K. H., Hogan, J. L., & Lewandowsky, S. (2017). Reminders and repetition of

misinformation: Helping or hindering its retraction? *Journal of Applied Research in

Memory and Cognition, 6*, 185-192. https://doi.org/10.1016/j.jarmac.2017.01.014

Ecker, U. K. H., Lewandowsky, S., & Apai, J. (2018). Terrorists brought down the plane!—

No, actually it was a technical fault: Processing corrections of emotive information.

Quarterly Journal of Experimental Psychology, 64, 283–310.

https://doi.org/10.1080/17470218.2010.497927

Ecker, U. K. H., Lewandowsky, S., Chang, E. P., & Pillai, R. (2014). The effects of subtle

misinformation in news headlines. *Journal of Experimental Psychology: Applied, 20*,

323-335. https://doi.org/10.1037/xap0000028

Ecker, U. K. H., O'Reilly, Z., Reid, J., & Chang, E. P. (2019). The effectiveness of short-format refutational fact- checks. *British Journal of Psychology, 111*, 36-54. https://doi.org/10.1111/bjop.12383

Guallar, E., Stranges, S., Mulrow, C., Appel, L. J., & Miller III, E. R. (2013). Enough is enough: Stop wasting money on vitamin and mineral supplements. *Annals of Internal Medicine*, *159*, 850–851. https://doi.org/10.7326/0003-4819-159-12-201312170-00011

Guillory, J., & Geraci, L. (2013). Correcting erroneous inferences in memory: The role of source credibility. *Journal of Applied Research in Memory and Cognition, 2*, 201-209. https://doi.org/10.1016/j.jarmac.2013.10.001

Hamby, A., Ecker, U. K. H., & Brinberg, D. (2020). How stories in memory perpetuate the continued influence of false information. *Journal of Consumer Psychology, 30*, 240-259. https://doi.org/10.1002/jcpy.1135

Kagel, J. H., & Levin, D. (2011). Auctions: A survey of experimental research, 1995-2010. *Handbook of Experimental Economics*, *2*, 563-637.

Kanya, L., Sanghera, S., Lewin, A., & Fox-Rushby, J. (2019). The criterion validity of willingness to pay methods: A systematic review and meta-analysis of the evidence. *Social Science & Medicine, 232*, 238-261.

Kendeou, P., Butterfuss, R., Kim, J., & Van Boekel, M. (2019). Knowledge revision through the lenses of the three-pronged approach. *Memory & Cognition, 47*, 33-46. https://doi.org/10.3758/s13421-018-0848-y

Landon-Murray, M., Mujkic, E., & Nussbaum, B. (2019). Disinformation in contemporary U.S. foreign policy: Impacts and ethics in an era of fake news, social media, and artificial intelligence. *Public Integrity, 21*, 512-522. https://doi.org/10.1080/10999922.2019.1613832

Lazer, D., Baum, M., Benkler, Y., Berinsky, A., Greenhill, K., Menczer, F., … & Zittrain, J.

    (2018). The science of fake news. *Science, 359*, 1094-1096.

    https://doi.org/10.1126/science.aao2998

Lewandowsky, S., Cook, J., Ecker, U. K. H., Albarracín, D., Amazeen, M. A., Kendeou, P.,

    … & Zaragoza, M. S. (2020). *The Debunking Handbook 2020*. https://sks.to/db2020;

    doi:10.17910/b7.1182

Lewandowsky, S., Ecker, U. K. H., & Cook, J. (2017). Beyond misinformation:

    Understanding and coping with the post-truth era. *Journal of Applied Research in*

    Memory and Cognition, 6, 353-369. https://doi.org/10.1016/j.jarmac.2017.07.008

Lewandowsky, Ecker, U. K. H., Seifert, C., Schwarz, N., & Cook, J. (2012). Misinformation

    and its correction: Continued influence and successful debiasing. *Psychological*

    Science in the Public Interest, 13, 106-131.

    https://doi.org/10.1177/1529100612451018

MacFarlane, D., Hurlstone, M. J., & Ecker, U. K. H. (2018). Reducing demand for

    ineffective health remedies: Overcoming the illusion of causality. *Psychology &*

    Health, 33, 1472-1489. https://doi.org/10.1080/08870446.2018.1508685

MacFarlane, D., Hurlstone, M. J., & Ecker, U. K. H. (2020a). Countering demand for

    ineffective health remedies: Do consumers respond to risks, lack of benefits, or both?

    Psychology and Health. https://doi.org/10.1080/08870446.2020.1774056

MacFarlane, D., Hurlstone, M. J., & Ecker, U. K. H. (2020b). Protecting consumers from

    fraudulent health claims: A taxonomy of psychological drivers, interventions,

    barriers, and treatments. *Social Science and Medicine,* 259, 112790.

    https://doi.org/10.1016/j.socscimed.2020.112790

Maertens, R., Roozenbeek, J., Basol, M., & van der Linden, S. (2020). Long-term
effectiveness of inoculation against misinformation: Three longitudinal experiments.
Journal of Experimental Psychology: Applied.

McEachan, R., Conner, M., Taylor, N., & Lawton, R. (2011). Prospective prediction of
health-related behaviours with the theory of planned behaviour: A meta-analysis.
Health Psychology Review, 5, 97-144. https://doi.org/10.1080/17437199.2010.521684

Mercier, H., (2020). *Not Born Yesterday: The Science of Who We Trust and What We
Believe.* Princeton University Press. https://doi.org/10.2307/j.ctvn1tbqq

Mosleh, M., Pennycook, G., & Rand, D. (2020). Self-reported willingness to share political
news articles in online surveys correlates with actual sharing on Twitter. *PLOS ONE,
15*), e0228882. https://doi.org/10.1371/journal.pone.0228882

Myers, L. B. (2014). Changing smokers' risk perceptions—for better or worse? *Journal of
Health Psychology, 19*, 325-332. https://doi.org/10.1177/1359105312470154

Oreskes, N., & Conway, E. (2010). Defeating the merchants of doubt. *Nature*, 465, 686–687.
https://doi.org/10.1038/465686a

Paynter, J., Luskin-Saxby, S., Keen, D., Fordyce, K., Frost, G., Imms, C., … & Ecker, U. K.
H. (2019). Evaluation of a template for countering misinformation—Real-world
Autism treatment myth debunking. *PLOS ONE, 14*, e0210746.
https://doi.org/10.1371/journal.pone.0210746

Pennycook, G., McPhetres, J., Zhang, Y., Lu, J., & Rand, D. (2020). Fighting COVID-19
Misinformation on social media: Experimental evidence for a scalable accuracy-
nudge intervention. *Psychological Science, 31*, 770-780.
https://doi.org/10.1177/0956797620939054

Poland, G., Jacobson, R., & Ovsyannikova, I. (2009). Trends affecting the future of vaccine

development and delivery: The role of demographics, regulatory science, the anti-

vaccine movement, and vaccinomics. *Vaccine, 27*, 3240-3244.

https://doi.org/10.1016/j.vaccine.2009.01.069

Powell, D., Keil, M., Brenner, D., Lim, L., & Markman, E. (2018). Misleading health

consumers through violations of communicative norms: A case study of online

diabetes education. *Psychological Science, 29*, 1104-1112.

https://doi.org/10.1177/0956797617753393

Rich, P. R., & Zaragoza, M. S. (2016). The continued influence of implied and explicitly

stated misinformation in news reports. *Journal of Experimental Psychology:

Learning, Memory, and Cognition, 42*, 62-74. doi:10.1037/xlm0000155

Roozenbeek, J., Schneider, C. R., Dryhurst, S., Kerr, J., Freeman, A. L., Recchia, G., ... &

van der Linden, S. (2020). Susceptibility to misinformation about COVID-19 around

the world. *Royal Society Open Science, 7*, 201199.

RMIT Fact Check. (2020, August 9). *Facebook post spruiking vitamin C as a coronavirus

cure - Fact Check.* ABC News. https://www.abc.net.au/news/2020-04-09/corona-5-

b/12130952?nw=0

Spinney, L. (2020, June 8). Bleach baths and drinking hand sanitiser: poison centre cases rise

under Covid-19. *The Guardian.*

https://www.theguardian.com/world/2020/jun/08/bleach-baths-and-drinking-hand-

sanitiser-poison-centre-cases-rise-under-covid-19

Shakoor, H., Feehan, J., Al Dhaheri, A. S., Ali, H. I., Platat, C., Ismail, L. C., … Stojanovska,

L. (2021). Immune-boosting role of vitamins D, C, E, zinc, selenium and omega-3

fatty acids: Could they help against COVID-19? *Maturitas*, *143*, 1–9.

Swire, B., Ecker, U. K. H., & Lewandowsky, S. (2017). The role of familiarity in correcting inaccurate information. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 43*, 1948-1961. https://doi.org/10.1037/xlm0000422

Swire-Thompson, B., DeGutis, J., & Lazer, D. (2020). Searching for the backfire effect: Measurement and design considerations. *Journal of Applied Research in Memory and Cognition*. https://doi.org/10.1016/j.jarmac.2020.06.006

Swire- Thompson, B., Ecker, U. K. H., Lewandowsky, S., & Berinsky, A. J. (2020). They might be a liar but they're my liar: Source evaluation and the prevalence of misinformation. *Political Psychology, 41*, 21-34. https://doi.org/10.1111/pops.12586

Therapeutic Goods Administration (2020). *No evidence to support intravenous high-dose vitamin C in the management of COVID-19*. https://www.tga.gov.au/alert/no-evidence-support-intravenous-high-dose-vitamin-c-management-covid-19

Thrasher, J. F., Rousu, M. C., Hammond, D., Navarro, A., & Corrigan, J. R. (2011). Estimating the impact of pictorial health warnings and "plain" cigarette packaging: evidence from experimental auctions among adult smokers in the United States. *Health Policy, 102*, 41-48. https://doi.org/10.1016/j.healthpol.2011.06.003

Vaccari, C., & Morini, M. (2014). The power of smears in two American presidential campaigns. *Journal of Political Marketing, 13*, 19-45. https://doi.org/10.1080/15377857.2014.866021

Van Bavel, J., Baicker, K., Boggio, P., Capraro, V., Cichocka, A., Cikara, M.,… & Willer, R. (2020). Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour, 4*, 460-471. https://doi.org/10.1038/s41562-020-0884-z

van der Linden, S., Leiserowitz, A., Rosenthal, S, & Maibach, E. (2017). Inoculating the

    public against misinformation about climate change. *Global Challenges, 1*, 1600008.

    https://doi.org/10.1002/gch2.201600008

Voelckner, F. (2006). An empirical comparison of methods for measuring consumers'

    willingness to pay. *Marketing Letters, 17*, 137–149. https://doi.org/10.1007/s11002-

    006-5147-x

Walter, N., & Tukachinsky, R. (2020). 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? *Communication Research, 47*, 155-177.

    https://doi.org/10.1177/0093650219854600

Weule, G. (2020, March 25). Can you boost your immune system to help fight coronavirus?

    Australian Broadcasting Corporation. https://www.abc.net.au/news/health/2020-03-

    25/can-you-boost-your-immune-system-to-help-fight-coronavirus/12085036

Zarocostas, J. (2020). How to fight an infodemic. *The Lancet*, *395*, 676–676.

    https://doi.org/10.1016/s0140-6736(20)30461-x

Figure 1. The social media posts shown to participants. Posts (a) to (c) are decoys; posts (b) and (c) are real tweets, whereas post (a) is a real tweet associated with a fictional handle. Post (d) is the target post containing endorsement of the misinformation.
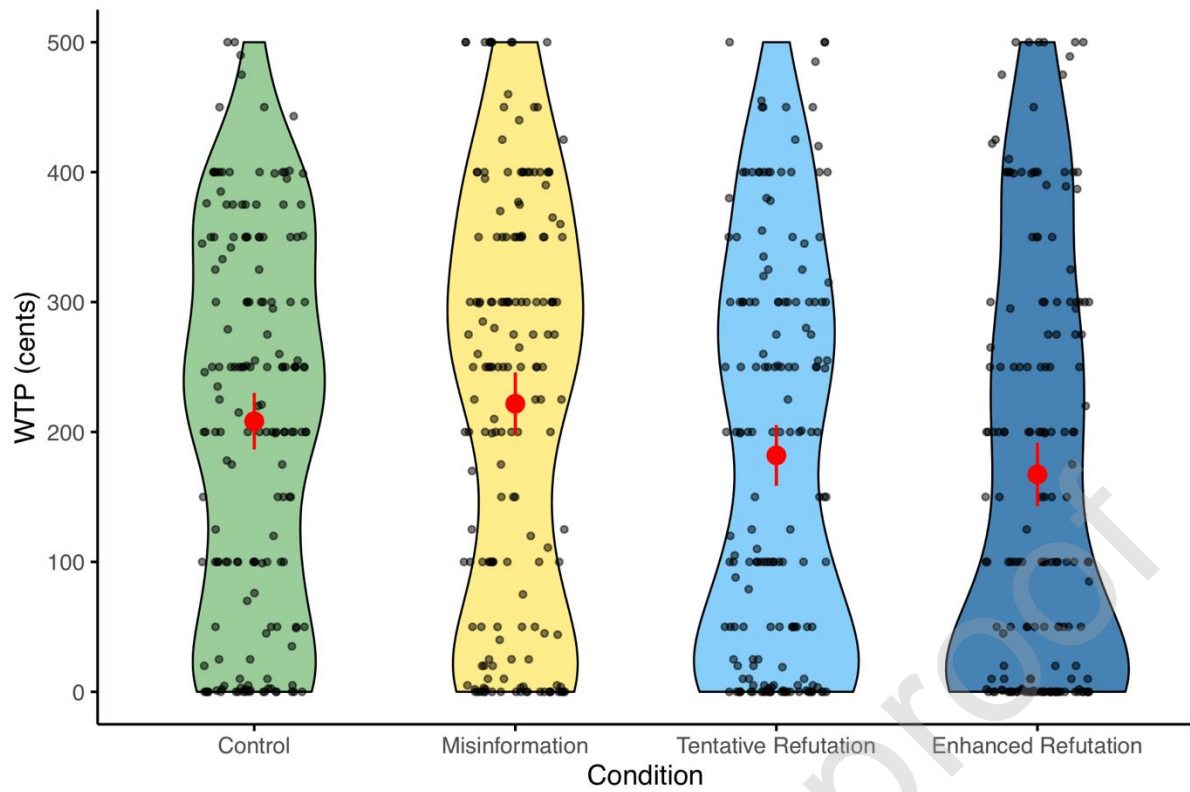
Figure 2. Violin plots showing willingness-to-pay (WTP) for the vitamin E supplement across conditions. Red markers and error bars indicate means and 95% confidence intervals.
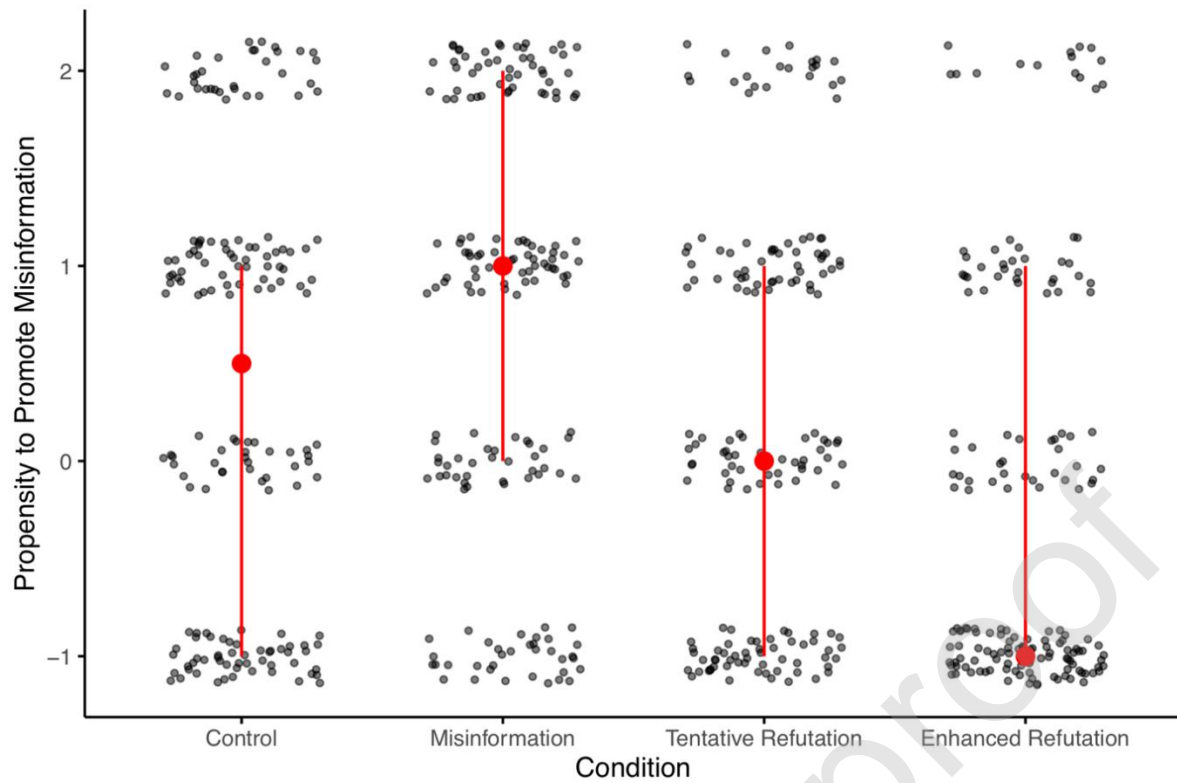
Figure 3. Bee-swarm plots showing engagement with the misinformation-endorsing social-media post across conditions. Higher scores indicate greater propensity to promote the misinformation (flag = -1; pass = 0; like = 1; share = 2; see text for details). Red markers and whiskers indicate medians and quartiles; individual data points are jittered along both axes.

Table 1

Deceptive Techniques Used in the Misinformation Text

Misinformation Technique	Excerpt
Appeal to authority	"Dr. Avery Clarke…said…experts in Taiwan have shown that COVID-19 can be slowed, or stopped completely, with immediate widespread use of high doses of Vitamin E"
Illusion of causality through false testimony	"I have seen patients who showed early symptoms of COVID coming on; I told them to try large doses of Vitamin E, and symptoms went away in just a few days"
Appeal to nature	"what's more, this remedy is completely natural"
Conspiratorial thinking	"I am urgently trying to get this message out now because it doesn't get much airtime—after all, it's cheap and readily available, so there's not much money in it"
Appeal to morality	"Imagine if you, or one of your loved ones, gets sick or dies, from something that is completely preventable."

Table 2

Key Components of the Enhanced Refutation Text

Debunking Technique	Excerpt
Highlight:	
(a) the trustworthiness of the refutation source, and	"Professor Simon Corner, a public health expert at the University of Chicago"
(b) the untrustworthiness of the misinformation source	"Dr. Clarke intentionally hides the fact that he is not a medical doctor and that his many false claims are rejected by the medical community"
Make salient the discrepancy between fact and fiction	"Dr. Clarke is spreading false and misleading information to deliberately deceive people. […] Dr. Clarke's message starts with some true facts about the benefits of vitamins […but] there is no clinical evidence that Vitamin E supplements have any impact on COVID-19"
Provide a factual alternative account	"The only proven way to keep safe from COVID-19 is to maintain physical (social) distancing and ensure proper hygiene practices"
Debunk the appeal to nature	"Using terms such as 'natural' is designed to get people to associate the proposed treatment with being harmless"
Highlight key overlooked risks	"high doses of Vitamin E supplements have been linked to adverse side-effects…and some more serious diseases"
Debunk the appeal to morality	"sharing such bogus remedies with your family would put them at risk of serious side-effects, while providing no benefit"
Counteract the illusion of causality	"the majority of people will recover from COVID-19 thanks to their own immune systems, irrespective of any dietary supplements or unproven remedies they may or may not have taken"

Table 3

Post Hoc Comparisons Testing the Influence of Condition on Willingness-To-Pay

		Mean Difference	95% CI for Mean Difference		SE	t	Cohen's d	p_{Tukey}
			Lower	Upper				
Control	Misinformation	-23.65	-62.68	15.39	15.16	-1.56	-0.16	0.40
	Tentative Refutation	20.21	-19.01	59.43	15.23	1.33	0.16	0.55
	Enhanced Refutation	35.81	-3.31	74.93	15.19	2.36	0.24	0.09
Misinformation	Tentative Refutation	43.86	4.73	82.98	15.19	2.89	0.28	0.02*
	Enhanced Refutation	59.46	20.41	98.51	15.16	3.92	0.37	< .001***
Tentative Refutation	Enhanced Refutation	15.60	-23.61	54.82	15.23	1.03	0.10	0.74

Note. Confidence interval adjustment: Tukey method for comparing a family of 4 estimates. Cohen's *d* does not correct for multiple comparisons. * $p < .05$, ** $p < .01$, *** $p < .001$

Table 4. *Engagement with Misinformation-endorsing Social-media Post across Conditions*

Condition	Share %	Like %	Pass %	Flag %	Share+Like %	Pass+Flag %	Share % (vs Misinform.)	Flag % (vs Misinform.)	Share+Like % (vs Control)	Share+Like % (vs Misinfo.)
Control	19.41	30.59	18.82	31.18	50.00	50.00	-32.26	+52.32	-	-18.13
Misinformation	28.65	30.41	20.47	20.47	59.06	40.94	-	-	+18.13	-
Tentative-Refutation	13.10	28.57	25.00	33.33	41.67	58.33	-54.30	+62.86	-16.66	-29.46
Enhanced-Refutation	8.88	17.75	18.34	55.03	26.63	73.37	-69.03	+168.86	-79.21	-54.92