



# Promoting COVID-19 vaccine confidence through public responses to misinformation: The joint influence of message source and message content

Reed M. Wood<sup>a,\*</sup>, Marie Juanchich<sup>b</sup>, Mark Ramirez<sup>c</sup>, Shenghao Zhang<sup>d</sup>

<sup>a</sup> Department of Government, University of Essex, United Kingdom

<sup>b</sup> Department of Psychology, University of Essex, United Kingdom

<sup>c</sup> School of Politics and Global Studies, Arizona State University, United States

<sup>d</sup> Institute for International Relations, Tsinghua University, China

## ARTICLE INFO

Handling Editor: M Hagger

**Keywords:**  
 Pandemic  
 COVID-19  
 Vaccines  
 Misinformation  
 Debunking  
 Politics

## ABSTRACT

**Objective:** During the pandemic healthcare professionals and political leaders routinely used traditional and new media outlets to publicly respond to COVID-19 myths and inaccuracies. We examine how variations in the sources and messaging strategies of these public statements affect respondents' beliefs about the safety of COVID-19 vaccines.

**Methods:** We analyzed the results of an experiment embedded within a multi-wave survey deployed to US and UK respondents in January–February 2022 to examine these effects. We employ a test-retest between-subjects experimental protocol with a control group. Respondents were randomly assigned to one of four experimental conditions reflecting discrete pairings of message source (political authorities vs. healthcare professionals) and messaging strategy (debunking misinformation vs. discrediting mis-informants) or a control condition. We use linear regression to compare the effects of exposure to treatment conditions on changes in respondent beliefs about the potential risks associated with COVID-19 vaccination.

**Results:** In the UK sample, we observe a statistically significant decrease in beliefs about the risks of COVID-19 vaccines among respondents exposed to debunking messages by healthcare professionals. We observe a similar relationship in the US sample, but the effect was weaker and not significant. Identical messages from political authorities had no effect on respondents' beliefs about vaccine risks in either sample. Discrediting messages critical of mis-informants likewise had no influence on respondent beliefs, regardless of the actor to which they were attributed. Political ideology moderated the influence of debunking statements by healthcare professionals on respondent vaccine attitudes in the US sample, such that the treatment was more effective among liberals and moderates than among conservatives.

**Conclusions:** Brief exposure to public statements refuting anti-vaccine misinformation can help promote vaccine confidence among some populations. The results underscore the joint importance of message source and messaging strategy in determining the effectiveness of responses to misinformation.

## 1. Introduction

Misinformation regarding the safety and effectiveness of COVID-19 vaccines circulated widely on social media even before the completion of large-scale clinical trials in late 2020. This included misleading statements that exaggerated the (very low) risk of side effects such as blood clots and myocarditis, myths that mRNA vaccines (Pfizer and Moderna) altered recipients' DNA, and false claims that vaccines caused infertility and increased miscarriage risk (Abbasi, 2022; Islam et al., 2021). The volume of misinformation surrounding COVID-19

vaccines—as well as the politicization of attitudes about COVID-19 more generally—threatened to undermine vaccination campaigns in many countries by creating public skepticism and uncertainty (Lomba et al., 2021; Neely et al., 2022).

Because vaccination was the cornerstone of many countries' COVID-19 pandemic mitigation strategy, efforts to limit the negative impact of misinformation on public confidence in vaccines became a key component of pandemic response (Cornwall, 2020). Government officials, politicians, scientists, and health professionals frequently contributed to this effort—both officially and unofficially—by directly responding to

\* Corresponding author. Department of Government, 5.424, Colchester Campus, Wivenhoe Park, Colchester, CO43SQ, United Kingdom.

E-mail address: [reed.wood@essex.ac.uk](mailto:reed.wood@essex.ac.uk) (R.M. Wood).

<https://doi.org/10.1016/j.socscimed.2023.115863>

Received 9 September 2022; Received in revised form 7 March 2023; Accepted 20 March 2023

Available online 22 March 2023

0277-9536/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

anti-vaccine misinformation (Royan et al., 2022; WHO, 2021). In addition to their efforts to develop and communicate official public health recommendations, healthcare professionals and political leaders frequently used op-eds, media interviews, blog posts, and statements on social media to publicly respond to COVID-19 myths and inaccuracies (e.g., Campbell, 2021; Clarke, 2021; Dickinson, 2022; Royan et al., 2022). Through these responses, they shared accurate information about COVID-19 vaccines and rebutted erroneous claims about vaccine risks. In some instances, they criticized and sought to undermine the credibility of the groups responsible for the misinformation.

Despite the proliferation of such commentary during the pandemic, little is known about the effects of public responses to vaccine misinformation on individuals' attitudes toward COVID-19 vaccines. Moreover, while some recent studies have evaluated the comparative effectiveness of different strategies for correcting COVID-19 vaccine misinformation (e.g., Amazeen et al., 2022; Ronzani et al., 2022), there remains significant need to identify which features of corrective messages are most effective at shaping public attitudes about vaccine safety.

We address these issues by examining how variations in the sources (political authorities vs. healthcare professionals) and messaging strategies (debunking misinformation vs. discrediting mis-informants) of public statements made in response to misinformation affect respondents' beliefs about the safety of COVID-19 vaccines. Results from an experiment embedded within a multi-wave survey deployed in the United States and the United Kingdom in January and February 2022 provide some evidence that the sources of corrective messages and the messaging strategies they adopt jointly influence the effectiveness of responses to misinformation. Moreover, they suggest that even brief exposure to public statements by credible sources (such as healthcare professionals) that refute anti-vaccine misinformation can promote public confidence in COVID-19 vaccines in some contexts.

## 2. Misinformation correction and COVID-19 vaccine hesitancy

Misinformation about COVID-19 vaccines was widely shared online by vaccine skeptics and spread through organized anti-vaccine demonstrations during the pandemic, contributing to a high level of misinformation exposure in many populations (Neely et al., 2022). Concerningly, exposure to misinformation about COVID-19 vaccines—particularly myths about vaccine safety—dissuaded many individuals from vaccinating (Loomba et al., 2021; Kreps et al., 2021). Surveys conducted in the US and UK suggest that large numbers of younger women were susceptible to misinformation about the adverse effects of COVID-19 vaccines on fertility and pregnancy; moreover, (unfounded) fears about these risks discouraged many from vaccinating (Abbasi, 2022; Holbrook, 2021). In one survey, fears about the long-term consequences of vaccines, including risks to future fertility, represented the most widely cited reasons for vaccine hesitancy among US parents of children aged 5 to 11 (Hamel et al., 2021). These examples highlight the importance of correcting misinformation as a means of promoting vaccine confidence and increasing uptake.

Previous studies demonstrate that corrections are effective at improving information accuracy on health-related topics (e.g., Walter et al., 2021). A growing body of recent evidence likewise suggests that specific strategies—including factual debunking, fact-checking, and rebuttals—effectively counter COVID-19 vaccine misinformation (Carey et al., 2022; Kreps and Douglas, 2022; Porter et al., 2022) and promote vaccine confidence (Chellenger et al., 2022; Ronzani et al., 2022). However, a recent meta-analysis concluded that corrections exerted a positive but statistically insignificant influence on respondents' beliefs in COVID-19 vaccine misinformation (Janmohamed et al., 2021). Other studies highlight the complexity of these relationships by demonstrating the moderating influence of respondents' pre-existing attitudes (Amazeen et al., 2022) or religiosity (Schmid and Betsch, 2022) on the effectiveness of pre-bunking and debunking strategies. These results demonstrate the need to clarify the conditions under which corrective

interventions succeed and to identify which strategies are most effective at countering inaccurate beliefs about COVID-19 vaccines among publics that have been widely exposed to anti-vaccine misinformation.

## 3. Public responses to COVID-19 anti-vaccine misinformation

During the pandemic, government officials and healthcare professionals often responded publicly to COVID-19 misinformation through statements posted on social media, in op-eds and blog posts, and during interviews with journalists. While these statements largely reinforced official guidance, many were made in an unofficial capacity and reflected the personal views of the individuals responsible for them. Moreover, their placement in wide variety of traditional and social media outlets allowed these messages to reach—and potentially influence—large numbers of citizens. Consequently, the public routinely encountered unofficial messaging from different expert sources and authority figures that was intended to correct anti-vaccine misinformation and promote vaccine confidence.

The specific responses political authorities and healthcare professionals adopted—including the content, tone, and target of their messages—represented implicit strategies of misinformation correction. These messages routinely included factual statements intended to correct vaccine misinformation and debunk myths about vaccine risks. Amid public concerns about side effects associated with the AstraZeneca vaccine and high levels of misinformation about the risks of COVID-19 vaccines more generally, then-UK Prime Minister Boris Johnson repeatedly stressed in media interviews that COVID-19 vaccines were safe and the risk of side effects was extremely low (Campbell and Stewart, 2021). Similarly, former US Secretary of Health and Human Services Alex Azar (2021) published an op-ed in the *New York Times* rebutting COVID-19 vaccine misinformation and reassuring Americans about their safety and efficacy. Physicians, nurses, and public health researchers also used interviews on health websites and social media posts to directly refute misinformation about COVID-19 vaccines, such as the myth that caused infertility or increased the risk of miscarriage among pregnant women (e.g., Campbell, 2021).

Rather than debunking misinformation, political leaders and health experts also occasionally used the opportunities to condemn, sham, or otherwise discredit the agents responsible for anti-vaccine misinformation or explicitly criticize individuals who remained skeptical of COVID-19 vaccines. French President Emmanuel Macron derided vaccine skeptics and anti-vaccine protesters, stating that they had “lost their minds” (Adghirni, 2021). Then-UK Health Secretary Sajid Javid likewise referred to anti-vaccine activists as “idiots spreading vicious lies” (Hughes, 2021). Similarly, in a letter published in the *Times of London*, an NHS doctor expressed anger at “charlatans who use their platforms to sow mistruths”, whom she described as “liars and cynics with blood on their hands” (Clarke, 2021). In early 2022, scores of scientists and doctors signed an open letter labeling popular podcast host Joe Rogan a “menace to public health” for spreading vaccine misinformation (Dickinson, 2022). Such comments, which often featured ad hominem attacks rather than factual refutations of misinformation, revealed growing antipathy toward anti-vaccine activists. They also represented explicit attempts to discredit mis-informants and undermine their influence.

As this discussion highlights, experts' and authorities' public responses to COVID-19 misinformation varied along multiple dimensions. First, the figures responsible for these statements differed in terms of their perceived credibility, trustworthiness, and expertise—factors previous studies demonstrate influence the effectiveness of corrections (e.g., Ecker et al., 2022). Second, their messages varied in content, tone, and targets. Variations in these features can be viewed as roughly analogous to different strategies of misinformation correction identified in previous studies (i.e., rebutting, debunking, or discrediting) (see Ecker et al., 2022; Lewandowsky et al., 2012). Recognizing and disaggregating these dimensions provides an opportunity to examine how different attributes of public responses to COVID-19 misinformation

influence their effectiveness.

#### 4. Source and content of corrective messages

We identify two dimensions of public statements made in response to COVID-19 vaccine misinformation that potentially influence their effectiveness: message source and messaging strategy. In terms of source, we focus on political authorities and healthcare professionals because they represent distinct categories of actors that routinely responded to COVID-19 misinformation through highly visible public commentary. We also consider two distinct misinformation correction strategies that commonly featured—either intentionally or inadvertently—in the public commentary of these actors: debunking and discrediting. The first largely follows best practice in misinformation correction, utilizing factual statements to refute specific inaccuracies (see [Vraga and Bode, 2020](#)). The second, by contrast, adopts a more reactionary response that attempts to undermine the legitimacy of mis-informants through attacks on their character and motives. By distinguishing these dimensions, our discussion highlights the joint influence of message source and strategy in countering anti-vaccine misinformation.

##### 4.1. Source credibility and effective corrections during the COVID-19 pandemic

Numerous previous studies suggest that messages attributed to political authorities or health experts are more effective at combating health-related misinformation than messages from social peers or anonymous social media users ([Van der Meer and Jin, 2020](#); [Vraga and Bode, 2017](#); [Walter et al., 2021](#)). Early evidence also suggested that debunking statements attributed to medical professionals were more effective than generic messages at correcting COVID-19 misinformation and increasing vaccination intention ([Ronzani et al., 2022](#)). A common thread among these studies is the assumption that the perceived credibility, trustworthiness, and expertise of these sources influences the effectiveness of the corrective messages attributed to them ([Gollwitzer et al., 2020](#); [Pluviano et al., 2020](#)). In the context of the pandemic, this implies that actors and institutions that the public widely views as credible information sources are more effective at countering anti-vaccine misinformation and promoting vaccine confidence.

Consistent with this logic, trust in governmental, scientific, and medical authorities strongly influenced susceptibility to COVID-19 misinformation, compliance with pandemic restrictions, and attitudes toward vaccines ([Blair et al., 2022](#); [McLamore et al., 2022](#); [Pagliaro et al., 2021](#)). Problematically, public attitudes toward these authorities varied during the pandemic: public trust in authorities initially increased at the start of the pandemic but generally waned over time as citizens grew dissatisfied with its costs, constraints, and duration ([Algan et al., 2021](#); [Nielsen and Johannes Lindvall, 2021](#)). However, throughout the pandemic scientists and health professionals maintained a distinct—and often substantial—advantage over political and governmental authorities in terms of public trust ([Algan et al., 2021](#); [Robinson et al., 2021](#); [Wellcome, 2020](#)). The comparatively high levels of trust citizens placed in doctors and scientists potentially amplifies the effectiveness of corrective messages attributed to them.

Partisanship and political ideology likewise influenced public trust in experts and, by extension, their susceptibility to COVID-19 misinformation and their attitudes toward COVID-19 vaccines ([Rodriguez et al., 2022](#); [Ruiz and Bell, 2021](#)). This relationship was starkly illustrated by the significantly lower vaccination rates and higher COVID-19 mortality rates in US counties that voted for Donald Trump in 2016 ([Gollwitzer et al., 2020](#)). These observations are consistent with previous findings demonstrating the moderating influence of ideology on misinformation correction efforts (e.g., [Nyhan and Reifler, 2010](#); [Walter et al., 2020](#)). Yet, survey evidence suggests that right-left political ideology influenced trust in science in only a relatively small number of countries

(principally the US and Canada) during the pandemic ([McLamore et al., 2022](#)). Despite the highly polarized political environment, trust in health professionals remained high among both Democrats and Republicans in the US during the pandemic ([Robinson et al., 2021](#)). Moreover, a recent multi-country study found that while comments by political elites polarized respondents' attitudes about COVID-19 restrictions, the same statements made by non-political experts tended to depolarize attitudes ([Flores et al., 2022](#)). Consequently, while corrective messages from either source may combat misinformation, messages conveyed by healthcare professionals and scientists may be comparatively more effective than those delivered by political authorities, especially in highly politicized environments.

##### 4.2. Debunking misinformation vs. discrediting mis-informants

Multiple strategies exist for correcting misinformation (e.g., [Lewandowsky et al., 2012](#); [Ecker et al., 2022](#)). Factual debunking is shown to be more effective than alternative approaches at reducing the influence of misinformation about public health issues ([Van der Meer and Jin, 2020](#); [Vraga and Bode, 2020](#)), including COVID-19 misinformation ([Kreps and Douglas, 2022](#); [MacFarlane et al., 2021](#)). However, efforts to discredit or undermine the legitimacy of the actor(s) responsible for misinformation represent an alternative approach to rebutting or debunking anti-vaccine misinformation ([Ecker et al., 2022](#); [MacFarlane et al., 2021](#)). Whether undertaken intentionally or inadvertently, government officials, politicians, and health experts occasionally employed this approach during the pandemic when they publicly disparaged vaccine skeptics or criticized the motives of anti-vaccine activists.

In the context of low information environments, where individuals experience high levels of uncertainty regarding the accuracy of competing claims, source credibility can influence message effectiveness ([Petty and Cacioppo, 1986](#)). Conversely, this logic implies that observers are more likely to dismiss information from sources that lack credibility. Corrections that undermine the credibility of the actor responsible for the misinformation may therefore prove more effective than messages that emphasize the expertise of the actors correcting the misinformation ([Walter and Tukachinsk, 2020](#)). Efforts to discredit anti-vaccine activists—such as by highlighting their lack of knowledge, framing their actions as selfish, or suggesting their goals are malign rather than simply misguided—may therefore represent a successful strategy for responding to anti-vaccine misinformation. While this approach has rarely been directly compared to other strategies, some scholars contend that discrediting mis-informants can play an important role in countering COVID-19 misinformation ([MacFarlane et al., 2021](#)).

Nonetheless, discrediting strategies are controversial and have raised concerns among some public health experts. Criticism can provoke feelings of moral reproach among unvaccinated persons, further entrenching their beliefs and reducing their intention to vaccination ([Rosenfield and Tomiyama, 2022](#)). Discrediting strategies may therefore “backfire” if unvaccinated individuals feel they are being judged or ridiculed for their decision. Even if the observer is not the explicit target of the criticism, individuals that sympathize with the group responsible for the misinformation may resent and reject the criticism. In one recent study, unvaccinated participants study felt that public discourse about COVID-19 vaccines was “unfair, moralistic, and patronizing”, and they perceived discrimination based on their vaccination status ([Henkel et al., 2022](#)).

Discrediting approaches that explicitly criticize mis-informants may also prove ineffective among the population more broadly. Studies of voter attitudes suggest that uncivil commentary and “mudslinging” by politicians reduces trust in both politicians and political institutions ([Mutz and Reeves, 2005](#)) and decreases citizens' interest in politics ([Kahn and Kenney, 1999](#)). Ridiculing and condemning vaccine skeptics and anti-vax activists may likewise undermine the authority of politicians and health professionals (from whom the public expects a more measured tone), leading the public to disregard the content of their

messages. Given the highly politicized pandemic landscape, this discussion suggests that critical comments from political authorities may not only prove ineffective but might spark offense, leading to backfire.

#### 4.3. Summary and empirical expectations

We contend that message source and message strategy jointly influence the effectiveness of messages intended to correct misinformation. Put otherwise, the effectiveness of statements intended to counter COVID-19 vaccine misinformation may depend not only on the source of the message but also on its content. Based on the above discussion, we therefore propose.

**H1.** Compared to a control condition, exposure to statements by healthcare professionals that debunk anti-vaccine, misinformation reduces respondent beliefs about harms associated with COVID-19 vaccines.

**H2.** Compared to a control condition, exposure to statements by political authorities (without clear partisan affiliation) that debunk anti-vaccine misinformation reduces respondent beliefs about harms, associated with COVID-19 vaccines.

Furthermore, we expect that statements that attempt to discredit mis-informants will ultimately prove either ineffective or counterproductive. However, we expect the adverse effects of discrediting statements to be greater when they originate from political sources compared to health professionals. Consequently, we propose.

**H3.** Compared to a control condition, exposure to statements by healthcare professionals that criticize the actors, responsible for anti-vaccine misinformation neither reduces nor increases respondent beliefs about harms, associated with COVID-19 vaccines.

**H4.** Compared to a control condition, exposure to statements by political authorities (without clear partisan affiliation) that discredit the actors responsible for anti-vaccine misinformation increases beliefs about, harms associated with COVID-19 vaccines.

Finally, based on the collective logic of our arguments, we also consider the relative effectiveness of the specific source-message pairings captured by our four treatment conditions. As we argue, the specific combination of message source and messaging strategy influence the effectiveness of corrections. Given the results of research highlighting the importance of apolitical experts at conveying (correct) COVID-19 information (e.g., Flores et al., 2022) and evidence that citizens trusted scientists and doctors significantly more than government authorities during the pandemic (e.g., Algan et al., 2021), these arguments suggests that debunking messages attributed to healthcare professionals may be more effective at reducing COVID-19 risk perception than those attributed to political authorities. Moreover, because we hypothesize that discrediting statements by either political authorities or healthcare professionals are ineffective or backfire, the discussion implies that messaging by healthcare professionals that debunks misinformation rather than discredits mis-informants will prove more effective at reducing COVID-19 vaccine risk perception than each of the other message-source combinations.

## 5. Methodology

We test these hypotheses via an experiment embedded within multi-wave online survey carried out in both the United States and United Kingdom by the survey firm YouGov. We select the US and UK as the study sites for several reasons. First, vaccines were widely available in both countries at the time of the study. Vaccination therefore represented a personal choice, and uptake was rarely affected by supply constraints or accessibility issues. While vaccination rates were high in both countries—reaching 77% of the UK population and 72% in the US—by early 2022, millions nonetheless remained unvaccinated (see

Mathieu et al., 2021). Moreover, misinformation about COVID-19 vaccines circulated widely on social media, podcasts, and talk shows, and anti-vaccine activism, including protests and demonstrations intended to disseminate misinformation, were commonplace in both countries (Dickinson, 2022; Loomba et al., 2021).

Despite these similarities, important differences between the countries warrant exploration of the potential heterogeneous effect of debunking strategies. For example, while COVID-19 vaccines became highly politicized in the US (Gollwitzer et al., 2020; Ruiz and Bell, 2021), political ideology appeared to have only a minimal influence on British attitudes toward vaccines (Klymak and Vlandas, 2022). Similarly, while high-profile conservative political elites in the US (e.g., Donald Trump) often derided official health recommendations and expressed sympathy with vaccine skeptics, British political elites across the political spectrum were largely united in their support for the national vaccination program (Klymak and Vlandas, 2022). Lastly, the structure of the health systems in each country—largely privatized in the US vs. free and publicly funded in the UK—produce important differences in the ways citizens access and pay for healthcare. These differences allow us to assess the effectiveness of our treatments across heterogeneous populations and in diverse socio-political contexts.

YouGov administered the initial survey instrument to 5900 (US  $n = 2947$ ; UK  $n = 2953$ ) adult (18+) respondents. This sample was drawn from YouGov's respondent pool and matched on demographic and political variables to each country's populations using data from census data and various commercial surveys. The samples were representative of the national population based on several standard demographic factors (e.g., education, age, gender, etc.). However, due to the previously reported vaccination inequalities between white and minority respondents in both countries, we oversampled respondents from minority groups in both countries. All participants from Wave 1 were invited to participate in a follow up survey, which included the experimental component. Wave 1 of the survey was deployed between January 17, 2022 and 21 January 2022, while Wave 2 was deployed between February 7, 2022 and February 16, 2022. A total of 4519 respondents participated in the second wave (US  $n = 2216$ ; UK  $n = 2303$ ), yielding a recontact rate of 77%. Successful recontacts were randomly assigned to multiple different groups corresponding to different study protocols: 55% of successful recontacts were randomly assigned to one of five conditions used in this study (described below) while 45% were assigned to conditions associated with different studies. Consequently, 2505 respondents associated with this study (US  $n = 1229$ ; UK  $n = 1226$ ) successfully completed study waves 1 and 2.

The sample of US respondents included in the analyses discussed below comprise 54% female and 49% non-white respondents with an average age of 48; the relevant UK sample comprises 54% female and 17% non-white respondents with an average age of 51. We report additional descriptive statistics by country and experimental condition in our online appendix. We excluded speeders (respondents that completed Wave 2 in fewer than 4 min) and participants that failed an attention check that asked respondents to identify the issue area of the protest (from a list provided) from our analyses. These criteria led to the exclusion of 14% of US respondents and 6% of UK respondents. The samples we ultimately analyzed therefore included 1054 US respondents and 1203 UK respondents. The results are extremely regardless of the exclusion criteria (Table A1.22).

### 5.1. Experiment design

Our design reflects a “test-retest” between-subjects experimental protocol with a control group. In the first wave of the survey, all respondents answered a battery of standard socio-demographic, political, and psychological questions as well a series of questions probing their attitudes and beliefs about COVID-19 vaccines. In the second wave, respondents read a news story about an anti-vaccine protest, after which they completed attention checks and were queried again about their

attitudes toward COVID-19 vaccines using the same questions deployed in Wave 1. Participants were randomly assigned to either a control condition or one of four experimental conditions defined by the intersection of the two dimensions identified above: 1) message sources (political authority vs. healthcare professional); 2) messaging strategy (debunking vs. discrediting). This represents a mixed factorial where the intervention was manipulated between-subjects while timing (before and after) was within-subject: 2 Sources x 2 Corrective messages + 1 Control x 2 Time points. Random allocation produced similar sample sizes across conditions (Tables A1.3-A1.12).

Our experimental conditions emulated a newspaper article, thus providing a natural way to present a misinformation-only control condition as well as multiple experimental conditions that include response to misinformation that vary by message source and messaging strategy. Each vignette described a hypothetical recent anti-vaccine protest in either London (UK respondents) or Washington DC (US respondents) and was accompanied by the same photo of a protest (without specific visual details about its location or purpose). Both the control and treatment conditions included misleading statements about the risks/harms of COVID-19 vaccines, which were delivered as direct quotes from protest participants. These included: 1) vaccines cause female infertility, 2) vaccines cause miscarriages, and 3) vaccines commonly cause heart damage in children. The control condition only described the protest event and included misinformation about COVID-19 vaccines. The treatment conditions included these components as well as one of four response statements, which varied along the two dimensions discussed above.

Table 1 describes the specific content of each of these statements and illustrates the 2 × 2 nature of the conditions, where each pairing (source-strategy) represents a discrete treatment condition. Respondents in each country received distinct versions of the survey instrument and vignettes that differed only in terms of country-specific terminology (e.g., jab/vaccination or GP/doctor). In our empirical analyses, we compare each of these four treatment conditions to the misinformation-only control condition. We include these vignettes in the online appendix.

Debunking statements adopted a reassuring tone that asserted the general safety of the vaccines and specifically refute misinformation about the risks of the vaccine to young children and pregnant women—groups that were often singled out as especially vulnerable to harms from vaccines in the misinformation disseminated by anti-vaccine activists. By contrast, discrediting statements criticized and sought to delegitimize the source of the misinformation by referring to anti-vax

**Table 1**  
Text of statements and sources used in the experimental conditions.

Corrective Message Source	
<i>Healthcare Professionals</i>	<i>Political Authorities</i>
1) Pediatrician/GP Theresa Robinson (US/UK)	1) Congresswoman/MP Theresa Robinson (US/UK)
2) Damien Gordon, head of Intensive Care at St. Mary's Hospital (US & UK)	2) Damien Gordon, a spokesperson for the Office of the President/Cabinet Office (US/UK)
Corrective Messaging Strategy	
<i>Debunk Misinformation</i>	<i>Discredit Mis-informants</i>
1) "Millions of children—including both of my own kids—have been vaccinated against COVID-19. The vaccines are very safe and provide children and their families high levels of protection against a very serious disease."	1) "I have lost all patience with anti-vaxxers and their nonsense. They are discouraging people from protecting themselves and putting lives at risk. They should be ashamed."
2) "There is zero evidence that vaccination poses a risk to pregnant women. Getting vaccinated is the best way for expectant mothers to protect themselves and their babies against COVID-19."	2) "They are selfish idiots, and their actions are dangerous. So many of us have worked tirelessly throughout the pandemic to save people. Groups like that are an insult to our sacrifices."

groups as "idiots" and highlighting that their actions threaten the health of others. Because the discrediting messages did not include the same information as the factual debunking messages, they do not represent truly parallel conditions. We therefore did not compare the effectiveness of messages that combine comments discrediting mis-informants and factual debunking to the effectiveness of messages that *only* include factual debunking. Rather, we conceived of critical commentary intended to disparage and discredit mis-informants as a distinct messaging strategy, which we compare directly to the strategy of debunking misinformation using factual statements.

5.2. *Dependent variables*

Our dependent variable reflects respondents' perceptions of serious health risks associated with COVID-19 vaccines. Specifically, respondents indicated their beliefs about the likelihood that receiving a COVID-19 vaccine would cause serious harm to each of the following groups: adult females, pregnant women, and children under age 12. They were also asked their beliefs about the likelihood that receiving a COVID-19 would cause the following specific severe side effects: infertility in women, miscarriage in pregnant women, and heart damage in children. All responses were given on a 6-point Likert scale, ranging from "extremely unlikely" to "extremely likely". We selected the groups identified above because anti-vaccine misinformation often focuses on their particular vulnerability to harms associated with COVID-19 vaccines. Similarly, we focused on the chosen side effects because they represent common (erroneous) beliefs about COVID-19 vaccines, and they reflect common features of anti-vaccine misinformation. Because respondents were not permitted to skip questions and were not provided with "unsure" (or similar options) for the questions that constitute the dependent variables, we have no missing observations for the respondents that completed Wave 2.

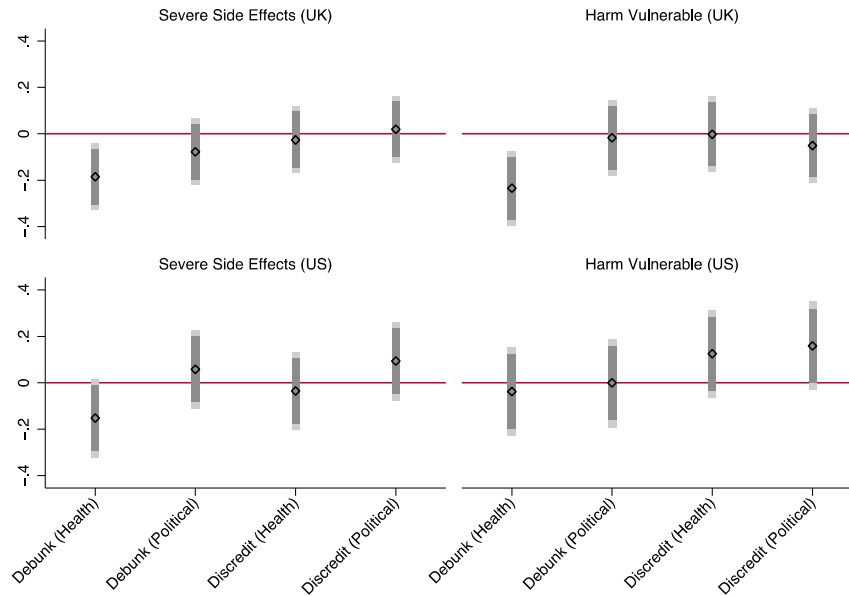
To examine the latent structure of these beliefs, we performed an exploratory factor analysis (EFA) on both sets of items (using values recorded in Wave 2). The EFA yielded a one-factor solution for the items regarding harm for specific groups [UK: Eigenvalue = 2.43; US: Eigenvalue = 2.26]. EFA also revealed a one-factor solution for the items related to specific side effects [UK: Eigenvalue = 2.64; US: Eigenvalue = 2.41]. Each of the factor loadings were above 0.8 for all of the items. We therefore averaged the items in each of the categories to create two composite measures: *Harm Vulnerable* [UK:  $\alpha = 0.94$ , US:  $\alpha = 0.92$ ] and *Severe Side Effects* [UK:  $\alpha = 0.96$ , US:  $\alpha = 0.94$ ]. These measures serve as the dependent variables in our empirical analyses.

5.3. *Estimation method*

We test our hypotheses with a set of linear regression models. In each model, we include the post-treatment values of either *Harm Vulnerable* or *Severe Side Effects* as the dependent variable and regress these indicators on a series of dummy variables indicating membership in one of the treatment conditions. We also included the pre-treatment value of the corresponding dependent variable in these models, which allowed us to adjust for each respondents' attitude prior to exposure to the treatment condition. Owing to heterogeneity in the sample populations, particularly with respect to baseline differences in vaccine risk perception (higher among US respondents [Figure A1.2]), as well as the key differences in healthcare systems and citizens' healthcare expectations the two countries, we conducted separate analyses of the two samples.

6. **Results**

We present coefficient plots for the results of the linear models in Fig. 1. Each plots shows the effect (as a coefficient estimate) of each treatment condition on the stated outcome relative to the control condition while controlling for respondents' pre-treatment risk perceptions. Results for the UK sample are reported in the top row and US results in



**Fig. 1.** Effects of each treatment condition on respondent perceptions of COVID-19 vaccine risk relative to control condition.

Coefficients estimates (black diamonds) with 90% (dark grey bars) 95% (light grey bars) CIs from OLS models predicting the effect of treatment conditions (x-axis) on respondent beliefs about vaccine risks (y-axis), controlling for respondent beliefs in pre-treatment study wave.

the bottom row. We include these results in tabular form along with additional models adjusting for covariates in our online appendix (Tables A1.13 and A1.23).

These results offer partial support for H1, which posited that factual debunking statements by health professionals were effective at reducing respondent perceptions of vaccine risks. In the UK sample, beliefs that vaccines cause severe adverse effects or harm specific groups of potentially vulnerable individuals were significantly lower among participants exposed to factual debunking statements from healthcare professionals relative to those in the control group ( $p < 0.05$ ). In substantive terms, exposure to this condition [*Debunk (Health)*] reduced respondents' belief that vaccines cause severe side effects by 0.19 points (95% CI [-0.33, -0.04]) and reduced belief that vaccines harm vulnerable groups by 0.23 points (CI [-0.40, -0.07]). While this represents only a modest reduction in respondents' perceptions of vaccine risk, the results nonetheless suggest that this treatment effectively shifted UK respondents' attitudes about COVID-19 vaccines during a period in which citizens' attitudes toward vaccines had already crystallized. Among US respondents, exposure to debunking statements by healthcare professionals likewise exerted a negative but only marginally significant effect on respondents' beliefs about serious side effects from vaccines relative to the control group ( $p < 0.10$ ). The coefficient in the model predicting respondent beliefs about the harms COVID-19 vaccines pose to vulnerable groups was also negative but insignificant ( $p = 0.699$ ). Thus, while the effect is consistently in the same direction, the effectiveness of debunking by healthcare professionals is weaker in the US compared to the UK.

The results offer no support for H2. The direction of the effect of debunking messages attributed to political authorities [*Debunk (Political)*] is inconsistent across the samples, and in no case does the coefficient achieve statistical significance. Thus, there is no evidence that debunking statements by political authorities affect respondent beliefs about vaccine risks. Consistent with this finding, respondent expressed substantially greater trust in healthcare professionals (US:  $\bar{x} = 3.59$ ,  $SD = 1.07$ ; UK:  $\bar{x} = 3.70$ ,  $SD = 0.96$ ) than in the national government (US:  $\bar{x} = 2.60$ ,  $SD = 1.25$ ; UK:  $\bar{x} = 2.69$ ,  $SD = 1.12$  (5-point scale). (Tables A1.1-A1.2). This observation accords with previous findings about

comparative trust during the pandemic as well as research demonstrating the influence of trust on citizens' attitudes toward COVID-19 vaccines and mitigation measures (Algan et al., 2021; Pagliaro et al., 2021).

Regardless of the source, response statements attempting to discredit mis-informants (anti-vaccine activists) did not have a significant effect on respondent beliefs relative to the control. According to the results in Fig. 1, exposure to discrediting statements by health professionals [*Discredit (Health)*] exerted a statistically insignificant influence on respondents' beliefs about vaccine risks relative of the control condition. To further investigate this apparent null effect, we conduct a series of Bayesian ANCOVA (BANCOVA) analyses using baseline risk perception as a covariate and the intervention vs. control as independent variables for both samples. Each BANCOVA compares four models: full null effects, effect of base respondent beliefs only, effect of the intervention only, or both. For both outcome variables and across both samples, the analyses show that the model including only the baseline perceptions (without the intervention) was the most likely given the data (*Severe Side Effects*,  $BF_M = 20.84$  [US] and 28.65 [UK]; *Harm Vulnerable*,  $BF_M = 16.77$  [US] and 30.46 [UK]). Analyses of these effects further suggest that the null effect of the intervention is 6–10 times more likely given the data (*Severe Side Effects*:  $BF_{excl} = 6.95$  [US] and 9.55 [UK] and *Harm Vulnerable*:  $BF_{excl} = 5.59$  [US] and 10.11 [UK]). These results provide further evidence of a null effect, thus supporting H3. We present full results in the appendix (Table A1.26).

The results also fail to support H4, which posited that discrediting statements by political authorities [*Discredit (Political)*] were not only ineffective but often counterproductive. Coefficients for this treatment were positive in three of four models but insignificant in all cases. Thus, even in a highly politicized context in which backfire is most likely, we find no evidence that corrective statements increase respondent belief in misinformation. This finding accords with previous studies that demonstrate the empirical rarity of backfire effects (e.g., Swire-Thompson et al., 2020). However, because we intentionally avoided explicit references to the partisan affiliation of the political authorities in our vignettes, we cannot rule out the possibility that provocative statements by politicians could produce backlash among individuals

with different partisan affiliation.

As an exploratory extension of our principal hypotheses, we also considered the effectiveness of the treatment conditions relative to one another. In addition to the results in Fig. 1, results from models directly comparing debunking by healthcare professionals to each of the other treatment conditions (excluding the control group) further demonstrate the relative effectiveness of this message-source combination (Figures A1.19-A1.21). Across both samples, *Debunk (Health)* was negatively associated with *Severe Side Effects* (US:  $-0.223$  CI  $[-0.40, -0.05]$ ; UK:  $-0.208$  CI  $[-0.35, -0.07]$ ) and *Harmed Vulnerable* (US:  $-0.199$  CI  $[-0.40, -0.00]$ ; UK:  $-0.185$  CI  $[-0.35, -0.02]$ ) when compared to *Discredit (Political)*. It was also negatively associated with *Harm Vulnerable* (US:  $-0.203$  CI  $[-0.40, -0.00]$ ; UK:  $-0.228$  CI  $[-0.39, -0.06]$ ) in both samples and was negatively associated with *Severe Side Effects* (UK:  $-0.162$  CI  $[-0.31, -0.02]$ ) in the UK sample when compared to the condition labeled *Discredit (Health)*. Lastly, *Debunk (Health)* was negatively associated with *Severe Side Effects* (UK:  $-0.220$  CI  $[-0.40, -0.04]$  relative to debunking by political authorities) in the US sample while in the UK sample it was negatively associated with *Harm Vulnerable* (UK:  $-0.206$  CI  $[-0.37, -0.04]$ ). Collectively, these results provide tentative support for our exploratory hypothesis and generally suggest that debunking messages attributed to health professionals are comparatively more effective than each of the other source-message combinations we discuss. Caution must be taken, however, as these results are preliminary.

Finally, recognizing that using listwise deletion for missing at random data might result in biased estimates (Sidi and Harel, 2018), we repeated our analyses using imputed data that attempted to simulate the attitudes of respondents that failed to return for the second wave. Similarly, due to missing observations for political ideology and vaccine status, which we use to examine moderating influences on the

relationships reported in Fig. 1 (see below), we impute missing values and re-estimate our models. We discuss our imputation method and report results for these models in our appendix (Figures A1.16-A1.18).

### 6.1. Moderating influences

To supplement our primary analyses, we assessed a series of models that explicitly consider the heterogeneity within our samples. First, because the effectiveness of interventions designed to counteract COVID-19 misinformation may partly depend on respondents' initial beliefs (Amazeen et al., 2022), we evaluated a set of models that interact each of the treatment conditions with the Wave 1 measure of the corresponding dependent variable. In no model did the interaction term reach statistical significance (Figures A1.12-A1.15), implying that the effectiveness of this intervention does not significantly depend on respondents' prior levels of beliefs about COVID-19 vaccine risks. As another means of assessing the potential moderating influence of respondents' pre-existing vaccine attitudes, we also examine models that evaluated the moderating influence of respondent vaccine status. These interactions were likewise insignificant (Figures A1.4-A1.7).

Second, we consider whether ideology moderates the relationships between the treatments and the dependent variables. We measure respondent political ideology via a Likert scale Liberal-Conservative index, which we interact with the variables representing the treatment conditions. Political ideology is a standard indicator provided by YouGov for US-based panellists but was added to the UK survey by the study authors. US respondents identified their political ideology on a 5-point scale ("Very Liberal" to "Very Conservative") while UK respondents identified their political ideology on a 7-point scale ("Extremely Liberal" to "Extremely Conservative").

The moderating influence of ideology varies across treatment

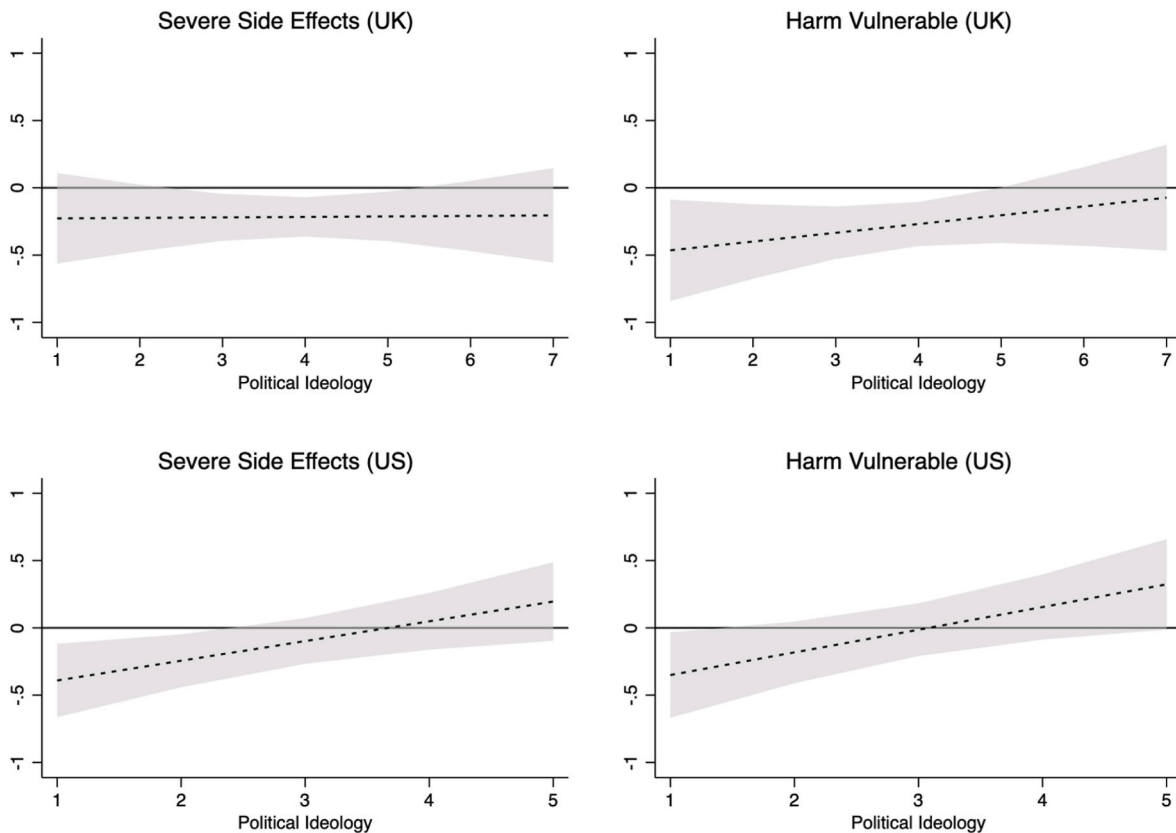


Fig. 2. Moderating effect of political ideology on the relationship between debunking by health professionals and respondent vaccine beliefs. Plots of the marginal effect (y-axis) of debunking treatment attributed to healthcare professionals over the scale of respondent political ideology (Liberal-Conservative) (x-axis). Dashed line depicts effect of treatment group relative to control group. Grey shading represents 95% confidence intervals.

conditions and between the samples. Among US respondents, ideology moderates the relationship between the healthcare professional debunking condition and each of the dependent variables. We present the marginal effects plots of these relationships in Fig. 2. Exposure to debunking message by healthcare professionals reduced misperception of vaccine risks/harms among Liberal respondents; however, it had the opposite influence on Conservatives, suggesting potential backfire. This result conforms to prior findings that ideology can moderate the effectiveness of corrections (Nyhan and Reifler, 2010) and suggests potential conditions under which efforts to correct misinformation might prove counterproductive. By contrast, we observe no such relationship in the UK sample. These results are consistent with previous findings, highlighting the intense politicization of vaccine attitudes in the US (Ruiz and Bell, 2021) and comparatively limited influence of political ideology on vaccine attitudes in the UK (Klymak and Vlandas, 2022). There is no evidence that ideology significantly influences the effectiveness of the other conditions (Figures A1.8-A1.11).

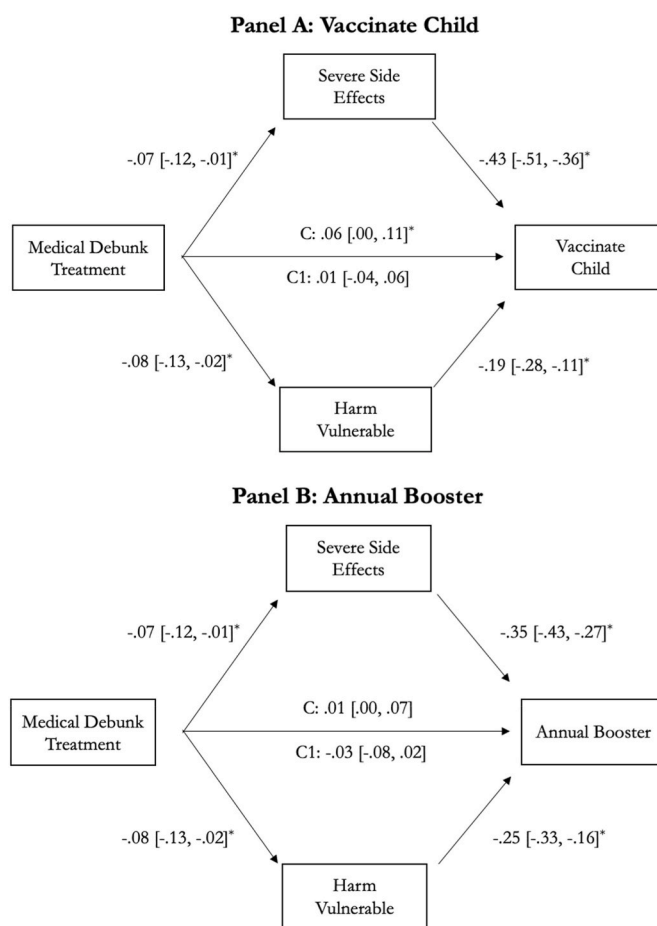
## 6.2. Indirect influences on vaccination intention

The broader goal of interventions designed to counter vaccine misinformation is to influence public behavior and increase vaccine uptake. We therefore test a mediation model where the independent variable—specifically the healthcare professional debunking treatment—can potentially directly or indirectly influence respondents' behavioral intentions through multiple mediator variables (Preacher and Hayes, 2008)—in this case, respondent beliefs about COVID-19 vaccine risks. The results of these analyses are presented in Fig. 3. We examine two distinct outcomes: Panel A considers the direct and indirect effects of the treatment on the likelihood that a respondent would vaccinate a 12-year-old child if they were the child's parent of guardian; Panel B considers these effects on respondents' intention to receive an annual COVID-19 booster vaccine if one "were available free of charge and recommended by healthcare professionals." Responses to these questions were given as 6-point ("extremely likely" to "extremely unlikely") and 5-point ("definitely not get it" to "deficiently get it") Likert scale responses respectively.

In both cases, the positive effect of debunking statements attributed to health professionals on vaccination intention (vs. control) was mediated by a lowered perception of severe side effects and reduced beliefs that the vaccine would harm vulnerable groups—two inaccurate claims that were explicitly debunked in the treatment. A Sobel test reveals that the indirect mediated effect of *Debunking (Health)* on willingness to vaccinate a child is positive and statistically significant for the mediators *Severe Side Effects* (0.07 CI [0.01, 0.13]) and *Harm Vulnerable* (0.06 CI [0.01, 0.10]). The indirect mediated effect of the treatment on annual booster intention is likewise significant for *Severe Side Effects* (0.10 CI [0.01, 0.20]) and *Harm Vulnerable* (0.06 CI [0.01, 0.10]). These results highlight the indirect influence of our treatment on behavioral intentions regarding COVID-19 vaccination among UK respondents. Specifically, exposure to statements by healthcare professionals that debunks vaccine misinformation reduces UK respondents' concerns about the risks associated with the vaccines, and through reducing these fears indirectly increase respondents' willingness to vaccinate.

## 7. Discussion and conclusions

Given the persistence of COVID-19 and its ongoing health risks, correcting vaccine misinformation remains an important task for public health authorities. Our study offers encouraging and actionable results for these efforts. The results highlight the joint importance of the corrective message source and messaging strategy in correcting inaccurate beliefs and promoting confidence in COVID-19 vaccines. Our results support prior studies that find that source credibility influences the effectiveness of corrective statements. However, we refine this relationship by demonstrating the importance of choosing the most



**Fig. 3.** Mediation Analysis of the Effect of Debunking by a Health Professional (vs. control) on Behavioral Intentions to Vaccinate Via Beliefs about Vaccine Risks.

Results of mediational analysis illustrating pathways through which the treatment influences the respondent's willingness to vaccinate a child (age 12) (Panel A) and intention to receive an annual booster vaccine (Panel B). Standardized coefficients with 95% confidence intervals shown. \* =  $p \leq 0.05$ . N = 1203. U.K. respondents only. C is the estimated effect of the treatment on the outcome without controlling for mediators. CI is the estimated effect of the treatment on the outcome after controlling for the mediators.

credible information sources among multiple potential authorities. In the case of a highly politicized pandemic, statements by healthcare professionals carry more weight than comments by political figures.

Our results also suggest that the content of the message influences its effectiveness. In general, disparaging comments intended to discredit mis-informants were ineffective, regardless of the credibility of the source of the comments. This accords with arguments that the public views ad hominem and accusatory personal political attacks as irrelevant and inappropriate (Kahn and Kenney, 1999; Mutz and Reeves, 2005). They might therefore disregard corrective messages they perceive as uncivil or intentionally disparaging, even when the messages are directed at the source of the misinformation. Notably, our results suggest that healthcare professionals might inadvertently undermine the effectiveness of their messages by resort to disparaging commentary. They should therefore ensure that responses to misinformation are made in good faith and avoid overtly critical comments in order to maximize message effectiveness. It is reassuring, however, that even in such a highly politicized environment, highly critical discrediting statements did not backfire.

In addition, we find evidence that passive corrections of misinformation—in the form of statements that individuals readily



encounter in the real world—can reduce vaccine hesitancy. We intentionally designed the experiment to mimic conditions under which individuals are likely to encounter COVID-19 misinformation as well as different responses to this misinformation. While previous studies have focused on online social media environments, we focus instead on traditional news media. Media outlets routinely report on groups that possess extreme or heterodox perspectives—such as COVID deniers, lockdown protests, and anti-vaccine activism. Media attention allows anti-vaccine activists to spread misinformation and sow uncertainty about vaccines, which could potentially promote vaccine hesitancy among the broader population. One advantage of traditional media over social media is that its coverage often includes rebuttals or dissenting opinions, which in the case of anti-vaccine activism can include factual or critical responses from political authorities or health experts. A key limitation, however, is that younger people and individuals with more extreme political views—who are generally less likely to be vaccinated—are comparatively less likely to engage with traditional news media. Nonetheless, evidence that responses messages delivered in this format can—at least in some contexts—shift public beliefs about vaccines suggests that governments, public health organizations, and individual health and science experts should view such interventions as one of several potentially useful strategies for debunking misinformation and thus promoting vaccine confidence.

We also acknowledge some important limitations of our study. First, the debunking and discrediting treatments do not constitute truly parallel conditions. Principally, the former included factual statements rebutting the misinformation, but the latter were largely information free. While we designed our discrediting treatments to evaluate how tone and target influence the effectiveness of public statements in response to vaccine misinformation, this omission prevents us from drawing clear conclusions about the effectiveness of discrediting as a strategy of misinformation correction. Rather, we can only conclude that statements ridiculing or condemning mis-informants has little influence on respondent beliefs about vaccine side-effects/harms. It is possible that the insignificant relationship results from our description of authority figures acting contrary to social norms by criticizing citizens/patients rather than from the attempt to discredit the mis-informant. Our characterization of mis-informants may have also diluted the effect of the intervention. Anti-vaxxers were criticized but implicitly framed as spreading misinformation rather than disinformation, and they were explicitly described as “dangerous idiots” but implicitly presented as having “good intentions” rather than explicitly attempting to deceive the public. Attacking their motives (rather than attacking their competence) might have been a more effective strategy for discrediting them (see Campos-Castillo et al., 2021).

Similarly, the inconsistent effect of the debunking treatments across the cases could be driven by their brevity and the omission of certain features identified as best practice in designing misinformation corrections (see Vraga and Bode, 2020). Our treatments include repeated factual refutations of incorrect information by credible sources. However, we might have strengthened them by providing clear explanations for why the misinformation is false (e.g., low rates of side effects in clinical trials). In addition, best practice suggests that corrective interventions should occur before misperceptions become entrenched. Our study occurred more than a year after the first COVID-19 vaccines were deployed in the US and UK, by which time myths and misinformation were widespread and public attitudes about vaccines were largely established (e.g., Loomba et al., 2021; Neely et al., 2022). Given that individuals’ beliefs about vaccines had largely crystallized by this point, we believe it is noteworthy that any interventions were successfully able to produce observable differences in respondents’ beliefs.

#### Credit author statement

**Reed M. Wood:** Conceptualization, Methodology, Formal Analysis, Writing (original and draft), Supervision, Investigation, Funding

Acquisition. **Marie Juanchich:** Conceptualization, Investigation, Methodology, Writing (review and editing), Funding Acquisition. **Mark Ramirez:** Investigation, Methodology, Formal Analysis; Writing (review and editing). **Shenghao Zhang:** Formal Analysis, Visualization, Data Curation.

#### Data availability

The authors have made all replication data and code available at <https://osf.io/nwkp3/>

#### Acknowledgements

The authors wish to thank the British Academy for funding this research (CRUSA210017). The Social Sciences Ethics Sub-Committee at the University of Essex reviewed and approved the application for ethical approval of this study (ETH-2122-0259). Data, codebooks, and commands to replicate all analyses reported in the manuscript and the online appendix are located at: <https://osf.io/nwkp3/>.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.socscimed.2023.115863>.

#### References

- Abbasi, Jennifer, 2022. Widespread misinformation about infertility continues to create COVID-19 vaccine hesitancy. *J. Am. Med. Assoc.* 327 (11), 1013–1015.
- Adghirni, Samy, 2021. French president Macron says anti-vax protests have ‘lost their minds’. *Bloomberg UK* 4 August 21. Available: <https://www.bloomberg.com/news/articles/2021-08-04/france-s-macron-says-anti-vax-protesters-have-lost-their-minds#xj4y7vzkg>.
- Algan, Yann, Cohen, Danel, Davoine, Eva, Foucault, Martial, Stantcheva, Stefanie, 2021. Trust in scientists in time of pandemic: Panel evidence from 12 countries. *Proc. Natl. Acad. Sci. USA* 118 (40).
- Amazeen, Michelle, Krishna, Arunima, Eschmann, Rob, 2022. Cutting the bunk: comparing the solo and aggregate effects of prebunking and debunking covid-19 vaccine misinformation. *Sci. Commun.* 44 (4), 387–414.
- Azar, Alex, 2021. I was the architect of operation warp speed. I Have a Message for All Americans. *New York Time* 2 August 2021. Available: <https://www.nytimes.com/2021/08/03/opinion/covid-vaccine-safety.html>.
- Blair, Robert, Curtice, Travis, Dow, David, Grossman, Guy, 2022. Public trust, policing, and the COVID-19 pandemic: evidence from an electoral authoritarian regime. *Soc. Sci. Med.* 305, 115045.
- Campbell, Leah, 2021. Debunking COVID-19 Vaccine Myths Spreading on Parent Facebook Groups Healthline 21 May 2021. Available: <https://www.healthline.com/health-news/debunking-covid-19-vaccine-myths-spreading-on-parent-facebook-groups>.
- Campbell, Denis, Stewart, Heather, 2021. Boris Johnson stresses COVID vaccine safety as tensions with NHS spill over. *Guardian* 18 March 2021. Available: <https://www.theguardian.com/world/2021/mar/18/boris-johnson-stresses-covid-vaccine-safety-as-tensions-with-nhs-spill-over>.
- Campos-Castillo, Shuster, Celeste, Stef, 2021. So what if they’re lying to us? Comparing rhetorical strategies for discrediting sources of disinformation and misinformation using an affect-based credibility rating. *Am. Behav. Sci.* <https://doi.org/10.1177/00027642211066058>.
- Carey, John, Guess, Andrew, Peter, Loewen, Merkle, Eric, Nyhan, Brendan, Phillips, Joseph, Reifler, Jason, 2022. The ephemeral effects of fact-checks on COVID-19 misperceptions in the United States, great britain and Canada. *Nat. Human Behav.* 6, 236–243.
- Chellenger, Aimée, Sumner, Petroc, Bott, Lewis, 2022. COVID-19 myth-busting: an experimental study. *BMC Publ. Health* 22 (131).
- Clarke, Rachel, 2021. I Suppress of Howl of Disbelief: She Would rather Die than Get a COVID Vaccine”. *The Times of London*, 27 November 2021. Available: <https://www.thetimes.co.uk/article/i-suppress-a-howl-of-disbelief-she-would-rather-die-than-get-a-covid-vaccine-n2l8r2wcj>.
- Cornwall, Warren, 2020. Officials gird for war on vaccine misinformation. *Science* 369, 6499.
- Dickinson, E.J., 2022. A menace to public health: doctors demand spotify puts an end to covid lies on ‘Joe rogan experience’. *Roll. Stone* 12 January 2022. Available: <https://www.rollingstone.com/culture/culture-news/covid-misinformation-joe-rogan-spotify-petition-1282240/>.
- Ecker, Ullrich, Lewandowsky, Stephan, Cook, John, Schmid, Philipp, Fazio, Lisa, Brashier, Nadia, Kendeou, Panayiota, Vraga, Emily, Amazeen, Michelle, 2022. The psychological driver of misinformation belief and its resistance to correction. *Nature Reviews: Psychology* 1, 13–29.

- Flores, Alexandra, et al., 2022. Politicians polarize and experts depolarize public support for COVID-19 management policies across countries. *Proc. Natl. Acad. Sci. USA* 119 (3).
- Gollwitzer, Anton Cameron, et al., 2020. Partisan differences in physical distancing are linked to health outcomes during the COVID-19 pandemic. *Nat. Human Behav.* 4, 1186–1197.
- Hamel, Liz, Lopes, Lunna, Grace Sparks, Kirzinger, Ashley, Kearney, Audrey, Stokes, Melisha, Brodie, Mollyann, 2021. FKK COVID-19 Vaccine Monitor October 2021. October 2021. Available: <https://www.kff.org/coronavirus-covid-19/poll-finding/kff-covid-19-vaccine-monitor-october-2021/>.
- Henkel, Luca, Sprengholz, Philipp, Korn, Lars, Betsch, Cornelia, Böhm, Robert, 2022. The association between vaccination status identification and societal polarization. *Nat. Human Behav.* <https://doi.org/10.1038/s41562-022-01469-6>.
- Holbrook, Chris, 2021. Predicting Better than Expected Vaccine Uptake, 16 February 2021. Available: <https://findoutnow.co.uk/blog/we-found-out-over-65s-vaccine-up-take-would-be-better/>.
- Hughes, David, 2021. Anit-vaxxers are 'idiots spreading lies, says health secretary (Lond. Engl.) 25 October 2021 Available: <https://www.standard.co.uk/news/uk/sajid-javid-keir-starmer-downing-street-sky-news-prite-patel-b962279.html>.
- Islam, Md Saiful, et al., 2021. COVID-19 vaccine rumors and conspiracy theories: the need for cognitive inoculation against misinformation to improve vaccine adherence. *PLoS One* 16 (5). <https://doi.org/10.1371/journal.pone.0251605>.
- Janmohamed, Kamila, et al., 2021. Interventions to mitigate COVID-19 misinformation: a systematic review and meta-analysis. *J. Health Commun.* 26, 846–857.
- Kahn, Kim Fridkin, Kenney, Patrick, 1999. Do negative campaigns mobilize or suppress turnout? Clarifying the relationship between negativity and participation. *Am. Polit. Sci. Rev.* 93 (4), 877–889.
- Klymak, Margaryta, Vlandas, Tim, 2022. Partisanship and the COVID-19 vaccination in the UK. *Nature: Sci. Rep.* 12, 19785.
- Kreps, Sarah, Douglas, Kriner, 2022. The COVID-19 infodemic and the efficacy of interventions intended to reduce misinformation. *Publ. Opin. Q.* 81 (1), 162–175.
- Kreps, Sarah, Goldfarb, Jillian, Brownstein, John, Douglas, Kriner, 2021. The relationship between US adults' misconceptions about COVID-19 vaccines and vaccination preferences. *Vaccines* 9, 901.
- Lewandowsky, Stephan, Ecker, Ullrich, Seifert, Colleen, Schwarz, Norbert, Cook, John, 2012. Misinformation and its correction: continued influence and success debiasing. *Psychol. Sci. Publ. Interest* 13 (3), 106–131.
- Loomba, Sahil, de Figueiredo, Alexandre, Simon, Piatek, Kristen de Graaf, Larson, Heidi, 2021. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nat. Human Behav.* 5, 337–348.
- MacFarlane, Douglas, Tay, Li Quan, Hurlstone, Mark, Ecker, Ulrich, 2021. Refuting spurious COVID-19 treatment claims reduces demand and misinformation sharing. *Journal of Applied Research in Memory and Cognition* 10, 248–258.
- Mathieu, Edourd, et al., 2021. A global database of COVID-19 vaccinations. *Nat. Human Behav.* 5, 947–953.
- McLamore, Quinnehtukut, et al., 2022. Trust in scientific information mediates associations between conservatism and coronavirus responses in the US, but few other nations. *Nature: Sci. Rep.* 12, 3724.
- Mutz, Diana, Reeves, Byron, 2005. The new videomalaise: effects of televised incivility on political trust. *Am. Polit. Sci. Rev.* 99 (1), 1–15.
- Neely, Stephen, Eldredge, Christina, Ersing, Robin, Remington, Christa, 2022. Vaccine hesitancy and exposure to misinformation: a survey analysis. *J. Gen. Intern. Med.* 37, 179–187.
- Nielsen, Julie, Johannes Lindvall, 2021. Trust in government in Sweden and Denmark during the COVID-19 pandemic. *W. Eur. Polit.* 44, 1180–1204.
- Nyhan, Brendan, Reifler, Jason, 2010. When corrections fail: the persistence of political misperception. *Polit. Behav.* 32 (2), 3030–3330.
- Pagliaro, S., Sacchi, S., Pacilli, M.G., Brambilla, M., Lionetti, F., Bettache, K., et al., 2021. Trust predicts COVID-19 prescribed and discretionary behavioral intentions in 23 countries. *PLoS One* 16 (3), e0248334. <https://doi.org/10.1371/journal.pone.0248334>.
- Petty, Richard, Cacioppo, John, 1986. *Communication and Persuasion*. Springer, New York, NY.
- Pluviano, Sara, Della Sala, Sergio, Watt, Caroline, 2020. The effects of source expertise and trustworthiness on recollection: the case of vaccine misinformation. *Cognit. Process.* 21, 321–330.
- Porter, Ethan, Velez, Yamil, Wood, Thomas, 2022. Factual corrections eliminate false beliefs about COVID-19 vaccines. *Publ. Opin. Q.* 86 (3), 763–773.
- Preacher, Kristopher J., Hayes, Andrew F., 2008. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav. Res. Methods* 40 (3), 879–891.
- Robinson, Scott E., Gupta, Kuhika, Ripberger, Joseph, Ross, Jennifer A., Fox, Andrew, Jenkins-Smith, Hank, Silva, Carol, 2021. *Trust in Government Agencies in the Time of COVID-19*. Cambridge University Press, Cambridge, UK.
- Rodriguez, Cristian, Kushner Gadarian, Shana, Wallace Goodman, Sara, Pepinsky, Thomas, 2022. Morbid polarization: exposure to COVID-19 and partisan disagreement about pandemic response. *Polit. Psychol.* 43 (6), 1169–1189.
- Ronzani, Piero, Panizza, Folco, Martini, Carlo, Savadori, Lucia, Motterlini, Matteo, 2022. Countering vaccine hesitancy through medical expert endorsement. *Vaccine* 40 (32), 4635–4643.
- Rosenfield, Daniel, Tomiyama, Janet A., 2022. Jab my arm, not my morality: perceived moral reproach as a barrier to COVID-19 vaccine uptake. *Soc. Sci. Med.* 294, 114699.
- Royan, et al., 2022. Use of twitter amplifiers by medical professionals to combat misinformation during the COVID-19 pandemic. *J. Med. Internet Res.* 24 (7), 10.2196%2F38324.
- Ruiz, Jeanette, Bell, Robert, 2021. Predictors of intention to vaccinate against COVID-19: results from a nationwide survey. *Vaccine* 39 (2021), 1080–1086.
- Schmid, Philipp, Betsch, Cornelia, 2022. Benefits and pitfalls of debunking interventions to counter mRNA vaccination misinformation during the COVID-19 pandemic. *Sci. Commun.* 44 (5), 531–558.
- Sidi, Yulia, Harel, Ofer, 2018. The treatment of incomplete data: reporting, analysis, reproducibility, and replicability. *Soc. Sci. Med.* 209, 169–173.
- Swire-Thompson, Briony, Joseph, DeGutis, Lazer, David, 2020. Searching for the backfire effect: measurement and design considerations. *Journal of Applied Research in Memory and Cognition* 9, 286–299.
- Van der Meer, Toni, Jin, Yan, 2020. Seeking formula for misinformation treatment in public health crises: the effects of information type and source. *Health Commun.* 35 (5), 560–575.
- Vraga, Emily, Bode, Leticia, 2017. Using expert sources to correct health misinformation in social media. *Sci. Commun.* 39 (5), 621–645.
- Vraga, Emily, Bode, Leticia, 2020. Correction as a solution for health misinformation on social media. *Am. J. Publ. Health* 110, S3.
- Walter, Nathan, Tukachinsk, Riva, 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? *Commun. Res.* 47 (2), 155–177.
- Walter, Nathan, Cohen, Jonathan, Holbert, R. Lance, Morag, Yasmin, 2020. Fact-checking: A Meta-analysis of What Works and for Whom. *Polit. Commun.* 37 (3), 350–375.
- Walter, Nathan, Brooks, John, Saucier, Camille, Suresh, Sapna, 2021. Evaluating the impact of attempts to correct health misinformation on social media: a meta-analysis. *Health Commun.* 36 (13), 1776–1784.
- Wellcome, 2020. Wellcome global monitor. How COVID-19 Affected People's Lives and their Views about Science. Available: <https://wellcome.org/reports/wellcome-global-monitor-covid-19/2020>.
- WHO, 2021. Fighting Misinformation in the Time of COVID-19, One Click at a Time World Health Organization, 27 April, 2021. Available: <https://www.who.int/news-room/feature-stories/detail/fighting-misinformation-in-the-time-of-covid-19-one-click-at-a-time>.