



# Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA

Sahil Loomba<sup>1,5</sup>, Alexandre de Figueiredo<sup>2,5</sup>✉, Simon J. Piatek<sup>2</sup>, Kristen de Graaf<sup>2</sup> and Heidi J. Larson<sup>2,3,4</sup>✉

**Widespread acceptance of a vaccine for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) will be the next major step in fighting the coronavirus disease 2019 (COVID-19) pandemic, but achieving high uptake will be a challenge and may be impeded by online misinformation. To inform successful vaccination campaigns, we conducted a randomized controlled trial in the UK and the USA to quantify how exposure to online misinformation around COVID-19 vaccines affects intent to vaccinate to protect oneself or others. Here we show that in both countries—as of September 2020—fewer people would ‘definitely’ take a vaccine than is likely required for herd immunity, and that, relative to factual information, recent misinformation induced a decline in intent of 6.2 percentage points (95th percentile interval 3.9 to 8.5) in the UK and 6.4 percentage points (95th percentile interval 4.0 to 8.8) in the USA among those who stated that they would definitely accept a vaccine. We also find that some sociodemographic groups are differentially impacted by exposure to misinformation. Finally, we show that scientific-sounding misinformation is more strongly associated with declines in vaccination intent.**

The spread of SARS-CoV-2, the causative agent of COVID-19, has resulted in an unprecedented global public health and economic crisis<sup>1,2</sup>. The outbreak was declared a pandemic by the World Health Organization on 11 March 2020<sup>3</sup>, and development of COVID-19 vaccines has been a major undertaking in fighting the disease. As of December 2020, many candidate vaccines have been shown to be safe and effective at generating an immune response<sup>4–6</sup>, with interim analysis of phase III trials suggesting efficacies as high as 95%<sup>7–9</sup>. At least two vaccine candidates have been authorized for emergency use in the USA<sup>10,11</sup>, the UK<sup>12,13</sup>, the European Union<sup>14</sup> and elsewhere, with more candidates expected to follow soon. For these COVID-19 vaccines to be successful, they need to be not only proven safe and efficacious, but also widely accepted.

It is estimated that a novel COVID-19 vaccine will need to be accepted by at least 55% of the population to provide herd immunity, with estimates reaching as high as 85% depending on country and infection rate<sup>15,16</sup>. Reaching these required vaccination levels should not be assumed given well-documented evidence of vaccine hesitancy across the world<sup>17</sup>, which is often fuelled by online and offline misinformation surrounding the importance, safety or effectiveness of vaccines<sup>18–20</sup>. There has been widely circulating false information about the pandemic on social media platforms, such as that 5G mobile networks are linked with the virus, that vaccine trial participants have died after taking a candidate COVID-19 vaccine, and that the pandemic is a conspiracy or a bioweapon<sup>21–23</sup>. Such information can build on pre-existing fears, seeding doubt and cynicism over new vaccines, and threatens to limit public uptake of COVID-19 vaccines.

While large-scale vaccine rejection threatens herd immunity goals, large-scale acceptance with local vaccine rejection can also

have negative consequences for community (herd) immunity, as clustering of non-vaccinators can disproportionately increase the needed percentage of vaccination coverage to achieve herd immunity in adjacent geographical regions and encourage epidemic spread<sup>24</sup>. Estimates of acceptance of a COVID-19 vaccine in June 2020 suggest that 38% of the public surveyed in the UK and 34.2% of the public in the USA would accept a COVID-19 vaccine (a further 31% and 25% were, respectively, unsure that they would accept vaccination against COVID-19)<sup>25</sup>. Worryingly, more recent polling in the USA (September 2020) has shown significant falls in willingness to accept a COVID-19 among both males and females, all age groups, all ethnicities and all major political groups<sup>26</sup>, possibly due to the heavy politicization of COVID-19 vaccination in the run up to the 2020 presidential election on both sides of the political debate<sup>27,28</sup>. The public's willingness to accept a vaccine is therefore not static; it is highly responsive to current information and sentiment around a COVID-19 vaccine, as well as the state of the epidemic and perceived risk of contracting the disease. Under these current plausible COVID-19 vaccine acceptance rates, possible levels of existing protective immunity—though it is unclear whether post-infection immunity confers long-term immunity<sup>29</sup>—and the rapidly evolving nature of misinformation surrounding the pandemic<sup>23,30</sup>, it is unclear whether vaccination will reach the levels required for herd immunity.

Recent studies have examined the effect of COVID-19 misinformation on public perceptions of the pandemic<sup>22,31,32</sup>, the tendency of certain sociopolitical groups to believe misinformation<sup>33,34</sup> and compliance with public health guidance, including willingness to accept a COVID-19 vaccine<sup>35,36</sup>. However, to our knowledge, there is no quantitative causal assessment of how exposure to

<sup>1</sup>Department of Mathematics, Imperial College London, London, UK. <sup>2</sup>The Vaccine Confidence Project, Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK. <sup>3</sup>Department of Health Metrics and Evaluation, University of Washington, Seattle, WA, USA. <sup>4</sup>Centre for the Evaluation of Vaccination, Vaccine and Infectious Disease Institute, University of Antwerp, Antwerp, Belgium. <sup>5</sup>These authors contributed equally: Sahil Loomba, Alexandre de Figueiredo. ✉e-mail: alex.defigueiredo@lshtm.ac.uk; heidi.larson@lshtm.ac.uk

misinformation affects intent to receive the vaccine and its implications for obtaining herd or community immunity if countries adopt this vaccination strategy. Moreover, it is essential to understand how misinformation differentially impacts sociodemographic groups and whether groups at high risk of developing severe complications from COVID-19 are more vulnerable to misinformation.

To fill this gap, we developed a pre–post-exposure study design and questionnaire to measure the causal impact of exposure to online pieces of misinformation relating to COVID-19 and vaccines on the intent to accept a COVID-19 vaccine, relative to factual information. In addition to assessing how misinformation might induce changes in vaccination intent, a further aim of this study is to investigate how exposure to misinformation differentially impacts individuals according to their sociodemographic characteristics (age, gender, highest education level, employment type, religious affiliation, ethnicity, income level and political affiliation), daily time spent on social media platforms<sup>37</sup>, and sources of trusted information on COVID-19. Understanding how misinformation differentially impacts sociodemographic groups and individuals according to their social media use or sources of trusted information can motivate the design of group-specific interventions to reduce the potential impact of online vaccine misinformation. Finally, we assess what makes certain information content more or less likely to influence intent to accept COVID-19 vaccination, which can be used to increase effectiveness of public health communication strategies.

For both the UK and the USA, both the treatment and control groups were nationally representative samples by gender, age and sub-national region. The causal impact of misinformation on vaccination intent was assessed on two key vaccination motives: (1) to accept a COVID-19 vaccine to protect oneself and (2) to accept a COVID-19 vaccine to protect family, friends and at-risk groups. By exploring vaccination intent to protect others, we are able to quantify how misinformation may affect altruistic vaccination behaviour—this is particularly important in the UK and the USA, where altruistic messaging prompts have been a feature of COVID-19 public health messaging campaigns<sup>38–41</sup>.

Our findings are interpreted in the light of vaccination levels required for herd immunity, and we discuss messaging strategies that may help mitigate or counter the impact of online vaccine misinformation. Throughout this study, misinformation refers to ‘false or misleading information’<sup>42</sup>, which is ‘considered incorrect based on the best available evidence from relevant experts at the time’<sup>43</sup>. Conversely, factual information refers to information that is considered correct based on the best available evidence from relevant experts at the time.

## Results

For this study, a total of 8,001 respondents were recruited via an online panel and surveyed between 7 and 14 September 2020—4,000 in the UK and 4,001 in the USA. Following randomized treatment assignment, 3,000 UK (and 3,001 US) respondents were exposed to misinformation relating to COVID-19 and vaccines (treatment group) in the UK (and the USA) and 1,000 in each country were shown factual information about COVID-19 vaccines (control group). Figure 1 presents an overview of the study design.

All respondents in both groups were asked to provide their intent to receive a COVID-19 vaccine before and after being exposed to vaccine information (misinformation or factual): ‘If a new corona-

virus (COVID-19) vaccine became available, would you accept the vaccine for yourself?’ (SELF) and ‘If a new coronavirus (COVID-19) vaccine became available, would you accept the vaccine if it meant protecting friends, family, or at-risk groups?’ (OTHERS). Responses were on a four-point scale: ‘yes, definitely’, ‘unsure, but leaning towards yes’, ‘unsure, but leaning towards no’ and ‘no, definitely not’. This scale was chosen to remove subjective ambiguity involved with Likert scales and to allow respondents to explicitly detail their intent, thereby allowing a more meaningful interpretation of results.

All information (misinformation and factual) was identified using Meltwater via a Boolean search string eliciting information and misinformation around a COVID-19 vaccine (Methods, ‘Selection of images’). A systematic selection approach was used to identify the COVID-19 vaccine information on social media with high circulation and engagement between 1 June and 30 August 2020. Information was classified as misinformation or factual after consulting reputable online sources of knowledge, such as peer-reviewed scientific research, webpages of public health organizations and fact-checking websites (or media outlets employing fact checkers) to verify the content and the context in which it was presented (Methods ‘Selection of images’). A final set of five pieces of misinformation comprising non-overlapping messaging and themes was selected to represent the diverse messaging found in COVID-19 vaccine misinformation (such as information questioning the importance or safety of a vaccine; Supplementary Table 1). As misinformation can be highly country- and context-dependent, it was decided to expose UK and US respondents to different sets of misinformation to reflect the different audiences targeted by the sources of misinformation, while factual information was the same for both groups. Each piece of (mis)information was shown on a separate page to facilitate image comprehension. For each exposure image, respondents were asked to rate the extent that: they agreed with the information displayed; they were inclined to be vaccinated; they believed the information to be trustworthy; they would fact check the information; and they would share the image. After exposure, the respondents were also asked if they had seen similar content on social media in the past month. The full questionnaire is shown in the Supplementary Materials.

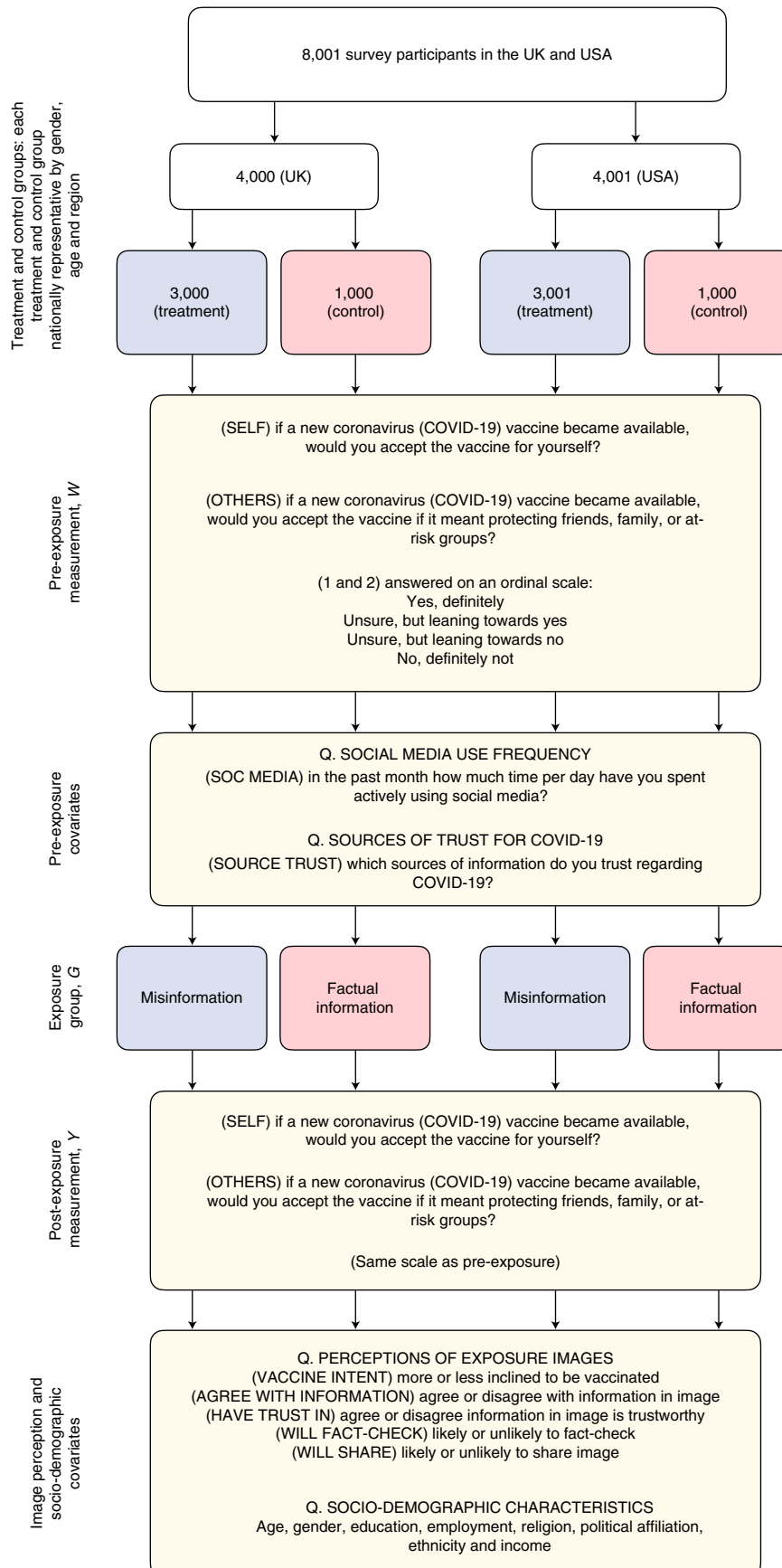
**Misinformation lowers intent to accept a COVID-19 vaccine.** Before treatment, 54.1% (95% percentile interval (PI) 52.5 to 55.7) of respondents in the UK and 42.5% (95% PI 41.0 to 44.1) in the US reported that they would ‘definitely’ accept a COVID-19 vaccine to protect themselves; whereas 6.0% (95% PI 5.3 to 6.8) and 15.0% (95% PI 14.0 to 16.1) said they would ‘definitely not’ accept a COVID-19 vaccine (Table 1). The remaining respondents were ‘unsure’ about whether they would accept a COVID-19 vaccine (Table 1). Higher intent to accept a COVID-19 vaccine in the UK than the USA has been reported recently<sup>25</sup>.

The treatment of misinformation exposure induces a decrease in the number of respondents who would ‘definitely’ take the vaccine relative to the control group in both countries by 6.2 percentage points (95% PI 3.9 to 8.5) in the UK and 6.4 percentage points (95% PI 4.0 to 8.8) in the USA (Table 1). There are corresponding increases in some lower-intent response categories. In the UK, we observe an increase of 2.7 percentage points (95% PI 1.0 to 4.5) in those ‘unsure, but leaning towards no’ and of 3.3 percentage points (95% PI 2.0 to 4.6) in those saying they ‘definitely will not’

**Fig. 1 | Overview of pre- and post-exposure study design.** A total of 8,001 participants across the USA and the UK were divided into treatment and control groups and had their intent to accept a COVID-19 vaccine measured. Respondents were then exposed to either misinformation or factual information before their vaccination intent was re-recorded. Additional survey items asked respondents to detail the frequency with which they use social media, their sources of trust for information around COVID-19 and their sociodemographic characteristics. The full questionnaire is reproduced in the Supplementary Information.

accept a vaccine, while in the USA there is a rise of 2.3 percentage points (95% PI 0.7 to 4.0) in those ‘unsure, but leaning towards no’ (Table 1).

While these values give the net effect of exposure to misinformation compared with the control, they conceal the full picture of the four post-exposure responses (Y) conditional on the four



**Table 1 | Exposure to COVID-19 vaccine misinformation reduces intent to accept a COVID-19 vaccine relative to exposure to factually correct information**

Distribution of vaccination intent	Vaccination intent	UK		USA	
		Self	Others	Self	Others
Pre-exposure $P(W)$	Yes, definitely	54.1 (52.5, 55.7)	63.7 (62.2, 65.1)	42.5 (41.0, 44.1)	53.3 (51.8, 54.9)
	Unsure, lean yes	31.9 (30.5, 33.3)	24.7 (23.4, 26.1)	30.0 (28.6, 31.4)	24.7 (23.3, 26.0)
	Unsure, lean no	8.0 (7.1, 8.9)	7.0 (6.2, 7.8)	12.4 (11.4, 13.5)	10.0 (9.2, 11.0)
	No, definitely not	6.0 (5.3, 6.8)	4.5 (3.9, 5.2)	15.0 (14.0, 16.1)	12.0 (11.1, 13.1)
Post-exposure $P(Y T)$ (treatment group)	Yes, definitely	48.6 (46.9, 50.2)	55.7 (54.1, 57.4)	39.8 (38.2, 41.5)	46.4 (44.8, 48.1)
	Unsure, lean yes	31.1 (29.5, 32.6)	27.9 (26.4, 29.4)	28.7 (27.1, 30.3)	26.7 (25.1, 28.2)
	Unsure, lean no	11.3 (10.2, 12.4)	9.6 (8.6, 10.6)	14.0 (12.9, 15.2)	11.7 (10.7, 12.8)
	No, definitely not	9.1 (8.2, 10.0)	6.8 (6.0, 7.6)	17.5 (16.3, 18.7)	15.2 (14.1, 16.4)
Post-exposure $P(Y C)$ (control group)	Yes, definitely	54.8 (52.6, 57.1)	61.5 (59.3, 63.7)	46.3 (44.0, 48.5)	52.9 (50.6, 55.0)
	Unsure, lean yes	30.9 (28.6, 33.2)	25.6 (23.4, 27.9)	25.4 (23.0, 27.8)	22.5 (20.2, 24.9)
	Unsure, lean no	8.5 (7.0, 10.2)	7.2 (5.8, 8.7)	12.5 (10.7, 14.4)	11.3 (9.5, 13.1)
	No, definitely not	5.8 (4.6, 7.0)	5.7 (4.6, 6.9)	15.9 (14.2, 17.5)	13.4 (11.8, 15.0)
Treatment effect $\Delta(Y) = P(Y T) - P(Y C)$	Yes, definitely	<b>-6.2 (-8.5, -3.9)</b>	<b>-5.7 (-7.9, -3.5)</b>	<b>-6.4 (-8.8, -4.0)</b>	<b>-6.5 (-8.8, -4.1)</b>
	Unsure, lean yes	0.2 (-2.5, 2.8)	2.3 (-0.3, 4.8)	<b>3.3 (0.6, 5.9)</b>	<b>4.2 (1.5, 6.9)</b>
	Unsure, lean no	<b>2.7 (1.0, 4.5)</b>	<b>2.3 (0.7, 4.0)</b>	1.5 (-0.6, 3.5)	0.4 (-1.7, 2.4)
	No, definitely not	<b>3.3 (2.0, 4.6)</b>	1.1 (-0.1, 2.3)	1.6 (-0.1, 3.3)	<b>1.9 (0.3, 3.5)</b>

The pre- and post-exposure intent to accept a COVID-19 vaccine are shown for the UK and the USA in the first three rows. The treatment effects of exposure to misinformation relative to factual information on the post-exposure responses  $Y$  that is  $\Delta(Y)$  are shown in the final row as percentage point changes in the number of people with intent  $Y$  after exposure to misinformation relative to factual information. Model probabilities have been quoted as percentages to aid interpretation. Values in parentheses indicate 95% percentile intervals (PI) with values in bold indicating PIs that do not include 0.

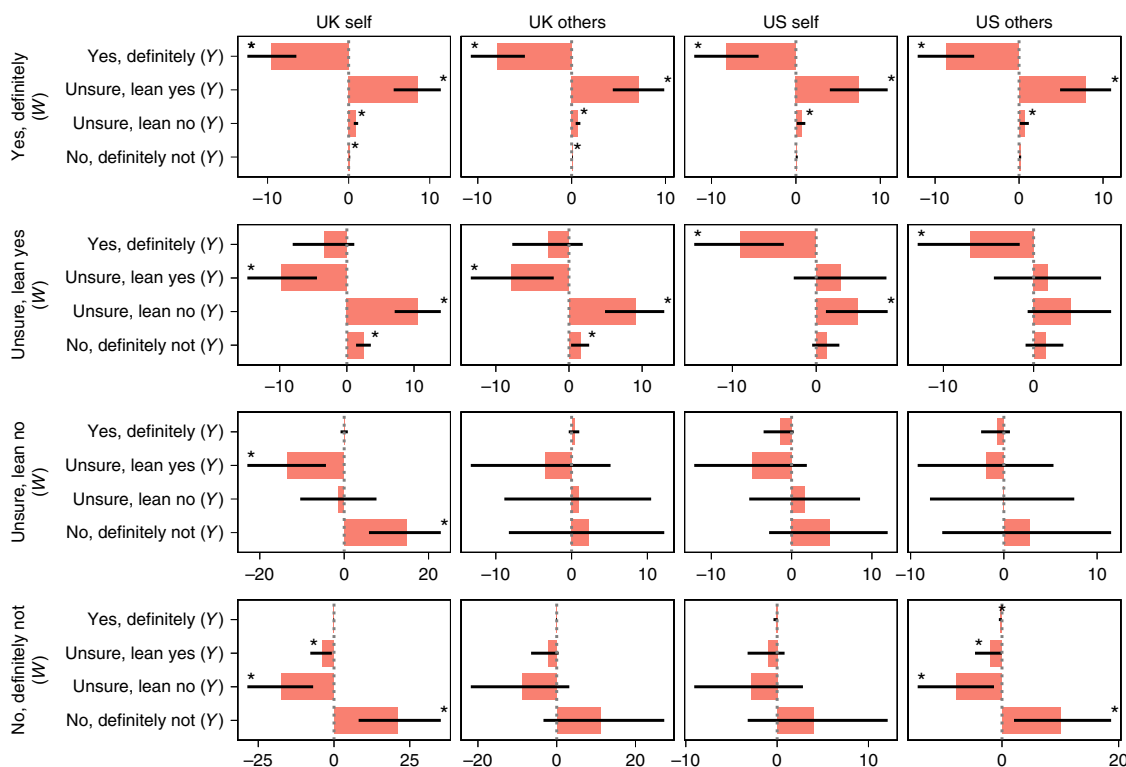
pre-exposure response ( $W$ ) for the treatment group compared to the control group, since exposure to information on COVID-19 vaccines may affect those with different prior vaccination intents differently. The changes in respondents' post-exposure response stratified by pre-treatment response are shown in Fig. 2 and Supplementary Table 2, where values indicate the percentage point change in the number of people with prior intent  $W$  who change intent to  $Y$  after exposure to misinformation, relative to factual information (Methods, 'Estimating treatment effects').

For any pre-treatment response, there is a net movement towards the response category immediately below (except for the pre-treatment 'no, definitely not' where there is a net increase in this response after exposure for the treatment group compared with the control). For example, in the UK there is a net increase of 8.5 percentage point (95% PI 5.5 to 11.4) in the post-exposure response 'unsure, but leaning towards yes' for respondents with pre-treatment response 'yes, definitely'. Similarly, there is a 10.6 percentage point (95% PI 7.1 to 14.0) increase in the post-exposure response 'unsure, but leaning towards no' for respondents with pre-treatment response 'unsure, but leaning towards yes' (Fig. 2). The same substantive results hold for the USA (Fig. 2).

Interestingly, more respondents in both countries would accept a vaccine if it meant protecting family, friends or at-risk groups (than if the vaccine was for themselves): 63.7% (95% PI 62.2 to 65.1) of respondents in the UK and 54.1% (95% PI 52.5 to 55.7) in the USA say that they would 'definitely' get vaccinated to protect others (Table 1). The exposure to misinformation again induces a decrease in intent to accept the vaccine to protect others, by 5.7 percentage points (95% PI 3.5 to 7.9) in the UK and 6.5 percentage points (95% PI 4.1 to 8.8) in the USA (Table 1) for the treatment group relative to the control. The treatment effects when conditioned on pre-treatment vaccination intent show a similar picture. For instance, in the USA there is a net decrease in those who previously responded 'definitely' by 8.7 percentage points (95% PI 5.3 to 12.1) and a net increase in those who previously responded 'no,

definitely no' by 10.0 percentage points (95% PI 2.1 to 18.7). The same substantive results hold for the UK.

**The impact of misinformation by sociodemographic characteristics.** A Bayesian ordered logistic regression model is used to establish whether the treatment of exposure to misinformation relative to factual information differentially impacted subjects' intent to accept a vaccine for themselves according to their sociodemographic background. We computed the heterogeneous treatment effects (HTEs), denoted by the statistic  $\Delta$  (equation (6), Methods), which represent the impact of exposure to misinformation relative to factual information, for a group of interest relative to its reference group. If  $\Delta$  is greater than 0, then the treatment of exposure to misinformation induces a lowering of vaccination intent, relative to the control for a specific group relative to the reference group (male, 18–24 years of age, highest education, employed, Christian, white, Conservative (UK) or Republican (USA) and highest income). In Fig. 3, we show this statistic for impact on vaccination intent to protect oneself—denoted by  $\Delta_s$ —and to protect others—denoted by  $\Delta_o$ —for each sociodemographic characteristic. (Raw parameter values can be found in Supplementary Tables 3 and 4). Below, we describe only those effects where the 95% PIs exclude zero, which we deem statistically credible. Since the HTEs are computed as a difference of log cumulative odds ratios between the treatment and control groups, we include these statistics separately for the treatment and control groups in Supplementary Figs. 1 and 2 and Supplementary Tables 3 and 4. Although they do not measure causal effects, these log cumulative odds ratios show how sociodemographic groups respond to misinformation or factually correct information relative to the reference group undergoing the same treatment. This reveals additional knowledge about those sociodemographic groups which—while not displaying a HTE—may be more inclined than the reference group to change their vaccination intent in the same direction upon exposure to either kind of information (full model details in Methods, 'Estimating treatment effects').



**Fig. 2 | Exposure to COVID-19 vaccine misinformation induces a net decrease in intent to accept a COVID-19 vaccine for all levels of pre-exposure intent.** Points indicate the relative change in probabilities (denoted as percentage point changes to aid interpretation) in the number of people with prior intent *W* who change it to *Y* after exposure to misinformation, relative to factual information (Methods, ‘Estimating treatment effects’). Bars indicate 95% PI; asterisks indicate PIs that do not include 0. Values are presented in Supplementary Table 2.

In both countries, we find evidence that some sociodemographic groups are differentially impacted by exposure to misinformation, relative to factual information. In the USA, females are less robust to misinformation than males when considering vaccination intent to protect others:  $\Delta_o = 0.42$  (95% PI 0.02 to 0.81). There is also evidence that lower-income groups (levels 0 to 2) are less likely to lower their vaccination intent to protect themselves or others upon exposure to misinformation than the highest income group (level 4): for level 0,  $\Delta_s = -0.83$  (95% PI  $-1.57$  to  $-0.12$ ); level 1,  $\Delta_o = -0.65$  (95% PI  $-1.33$  to  $-0.02$ ); level 2,  $\Delta_s = -0.86$  (95% PI  $-1.53$  to  $-0.20$ ) and  $\Delta_o = -0.80$  (95% PI  $-1.48$  to  $-0.13$ ). Interestingly, some groups respond similarly to misinformation to the reference group but show comparatively different inclinations to vaccinate upon exposure to factual information. Consequently, such groups are differentially more robust than their reference counterparts to exposure to misinformation relative to factual information, such as those from ‘other’ ethnic minorities in the USA when compared to whites:  $\Delta_s = -0.99$  (95% PI  $-1.65$  to  $-0.31$ ). Similar results are found in the UK, where unemployed respondents are more robust to misinformation than employed respondents, with  $\Delta_s = -0.99$  (95% PI  $-1.78$  to  $-0.19$ ); ‘other’ religious affiliations are more robust to misinformation than Christians, with  $\Delta_s = -0.76$  (95% PI  $-1.29$  to  $-0.23$ ); and those who are Jewish are more robust to misinformation than Christians, with  $\Delta_o = -1.58$  (95% PI  $-3.14$  to  $-0.02$ ).

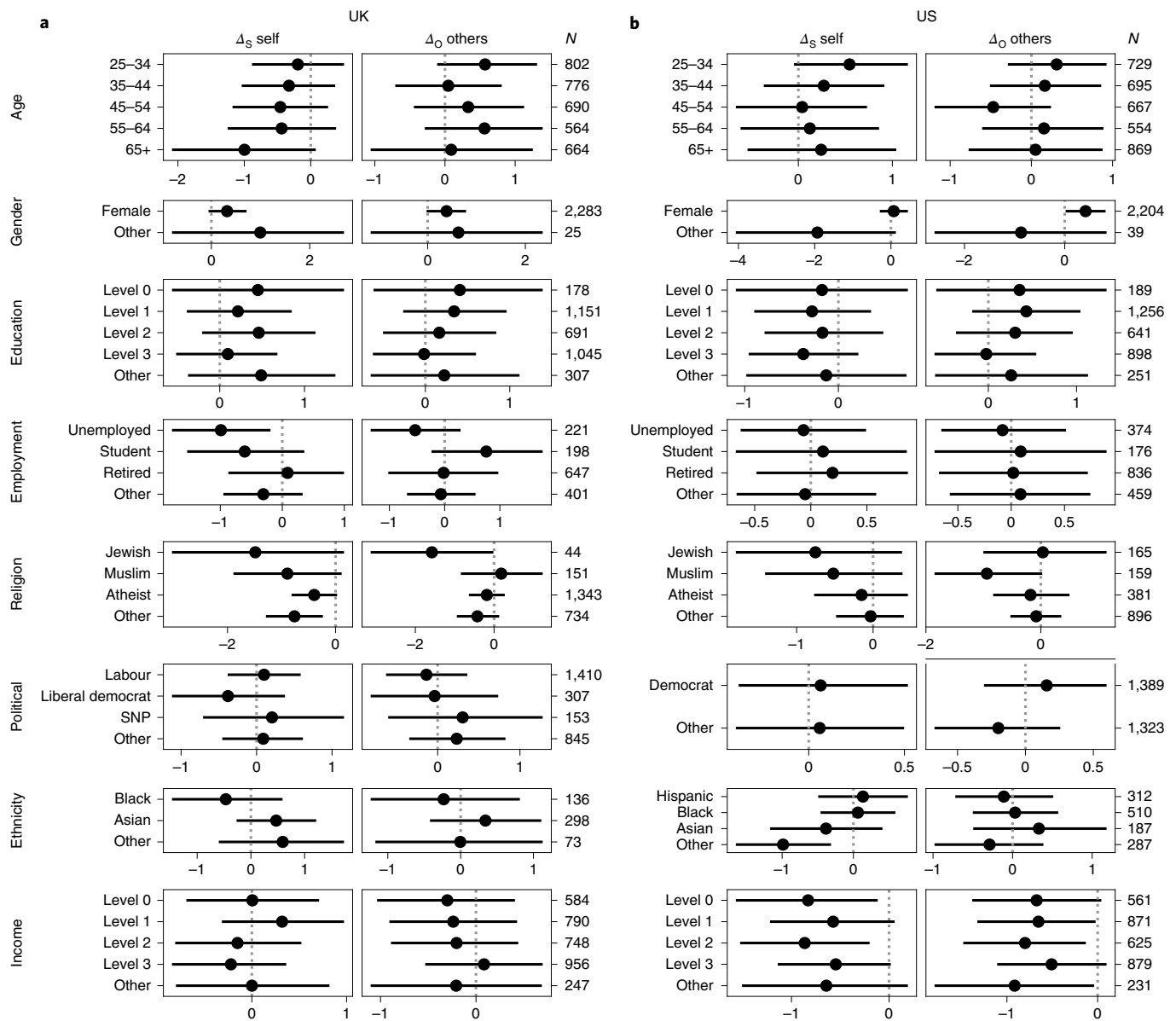
Finally, we investigated whether social media use and trust in sources of COVID-19 information differentially impacts vaccination intent. We remark that due to the similarity of HTEs obtained above for vaccination intent to protect oneself and to protect others, we report only analysis of vaccination intent to protect oneself here. We find no strong evidence to suggest that individuals in the UK or the USA who use social media more frequently are more likely

to lower their vaccination intent when exposed to misinformation compared with those in the control group (Supplementary Fig. 3 and Supplementary Table 5.) In the UK, individuals who trust celebrities for information about COVID-19 are more robust to COVID-19 misinformation than those who do not ( $\Delta_s = -1.31$  (95% PI  $-2.59$  to  $-0.03$ )), whereas in the USA, individuals who indicated trust in family or friends for such information are less robust than those who did not ( $\Delta_s = 0.52$  (95% PI 0.03 to 1.01)) (Supplementary Fig. 4 and Supplementary Table 6.)

**Correlational evidence of the appeal of scientific misinformation.**

After exposure to misinformation or factual information, respondents were asked to report whether, for each image: it raised their vaccination intent; they agreed with the information presented; they found the information to be trustworthy; they were likely to fact check; and they were likely to share the image with friends or followers (the full questionnaire and further details are provided in the Supplementary Materials). These post-exposure self-reported perceptions for all pieces of (mis)information are depicted in Fig. 4. Overall, it is apparent that in both countries, respondents were less likely to agree with, have trust in, fact check, share, or say that the information raised their vaccination intent when shown misinformation, as opposed to when they were shown factual information. Across both countries, around a quarter of respondents agreed with some of the misinformation or found it trustworthy, although the majority of respondents did not agree and did not find it trustworthy (Fig. 4a,b).

This study was not designed to investigate the causal impact of different kinds of vaccine messaging and cannot be used to draw causal inferences, as all participants rated self-perceptions after exposure. However, access to self-reported perceptions provides

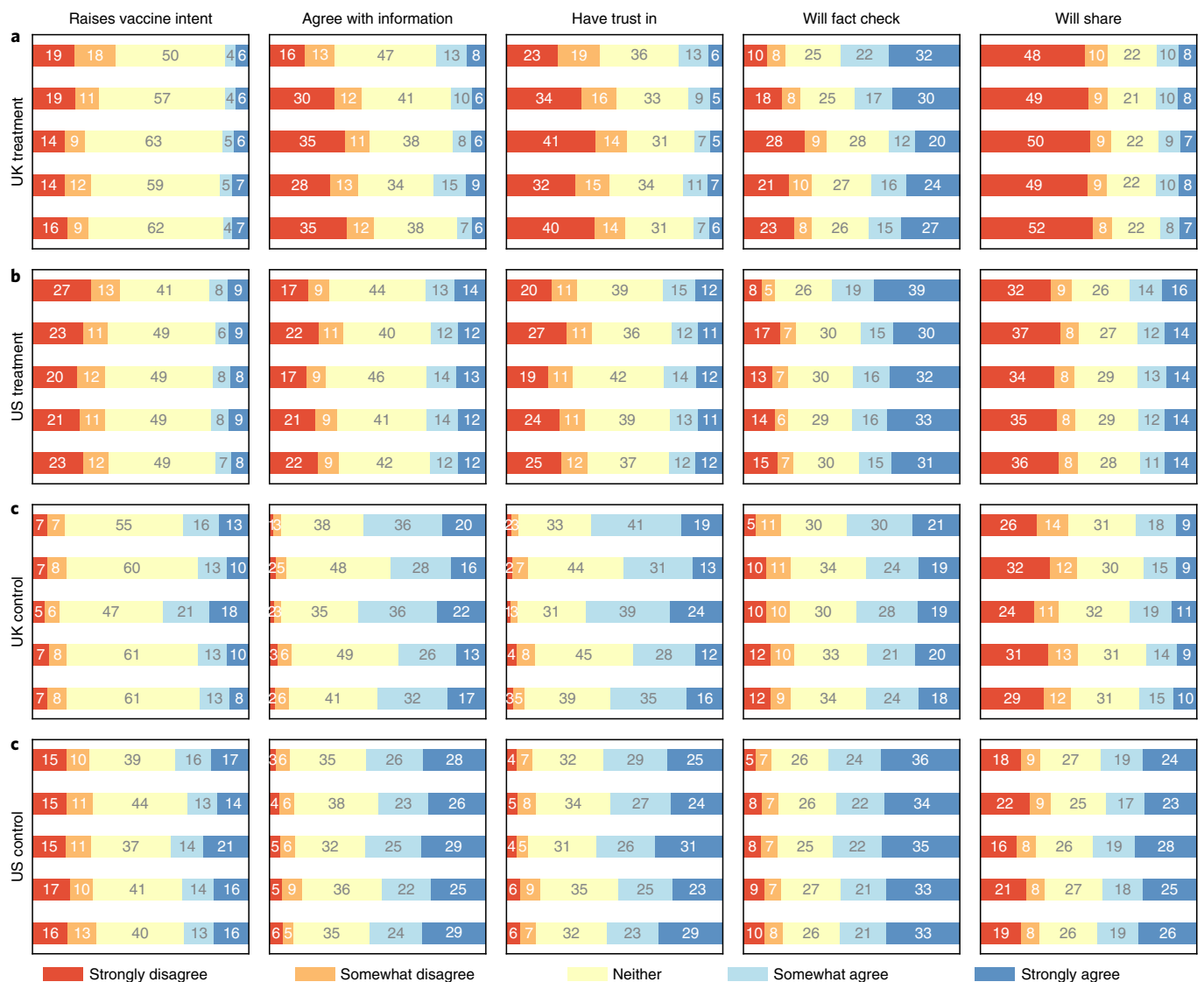


**Fig. 3 | Sociodemographic determinants of change in vaccination intent upon exposure to misinformation about COVID-19 vaccines, relative to factual information. a, b,** Contribution of sociodemographic characteristics to changes in intent to accept a vaccine to protect oneself (left column) and to protect others (right column) for the UK (**a**) and the USA (**b**). The reference category is male, 18–24 years of age, highest education, employed, Christian, Conservative (UK) or Republican (USA), white and highest income. Values indicate log cumulative odds ratios, such that a value above 0 indicates that the group is more likely to reject a COVID-19 vaccine than the reference group upon exposure to misinformation, relative to factual information, and a value below 0 indicates that they are less likely to reject the vaccine. Bars indicate 95% percentile intervals; numbers on the right indicate sample sizes of the corresponding demographic. Values are presented in Supplementary Tables 3 and 4.

correlational evidence of which pieces of information are associated with a greater (or lesser) decline in vaccination intent. For instance in the UK, image 1 (Supplementary Table 1)—which suggests that “scientists have expressed doubts [...] over the coronavirus vaccine [...] after all of the monkeys used in initial testing contracted coronavirus”—appears to be the misinformation piece that is associated with the largest decrease in vaccination intent (with 39% of respondents ‘disagreeing’ that the image raised their vaccination intent (Fig. 4)), whereas in the USA, it was image 1 (Supplementary Table 1)—which claimed “the new COVID-19 vaccine will literally alter your DNA”—that seems to induce the most impact (with 40% of respondents ‘disagreeing’ that the image raised their vaccination intent (Fig. 4)). In the control set—which was identical for respondents in both countries—the image that participants per-

ceived contributed the least to declines in vaccination intent was image 3 (Supplementary Table 1), in which the University of Oxford announced that their vaccine “produces a good immune response” and that the “teams @VaccineTrials and @OxfordVacGroup have found there were no safety concerns”—with 39% respondents in the UK and 35% in the USA ‘agreeing’ that the image raised their vaccination intent; see Fig. 4.

However, people’s self-reported changes in attitude—such as a ‘raise’ in vaccination intent—may mistakenly reflect their absolute levels of the attitude instead<sup>44</sup>—that is, level of vaccination intent. Therefore, to investigate the association of each individual image (information) with vaccination intent, weights (in the range of 0 to 1) were inferred for each image while regressing self-reported image perceptions against post-exposure vaccination intent (to protect



**Fig. 4 | Perceptive attitudes of respondents towards the information they were exposed to. a–d,** Bars indicate the breakdown of percentage of respondents providing a given response to each follow-up question to explore their perceptions of each image (information) they were exposed to. Respondents were asked whether each image raises their vaccination intent (column 1); contains information they agree with (column 2); contains information they find trustworthy (column 3); is likely to be fact-checked by them (column 4); and is something they will probably share with others (column 5). Rows represent images shown to the UK treatment group (a), US treatment group (b), UK control group (c) and US control group (d). Bars in each graph are ordered top to bottom from images 1 to 5. Those responding with ‘do not know’ were grouped with those saying ‘neither/nor’. The response scale for column 1 has been inverted—from ‘makes less inclined to vaccinate’ to ‘raises vaccine intent’—to facilitate direct comparison across all questions. The relevant questionnaire subsection is shown in the Supplementary Information.

oneself) and controlling for pre-exposure intent. This simultaneously reveals the predictive power of self-reported perceptions on actual change in vaccination intent and quantifies the association of each piece of (mis)information to the change in intent (Methods, ‘Estimating image impact’).

Since exactly five images were shown to each respondent, a weight above 0.2 would indicate a higher association with lowering vaccination intent than what would be expected at random, and a weight below 0.2 would indicate a lower association. This analysis confirms that the misinformation image with the largest (and statistically credible) association with loss in vaccination intent in the UK was indeed image 1 (Supplementary Table 1), with weight 0.42 (95% PI 0.28 to 0.56), while in the USA it was image 1 (Supplementary Table 1), with weight 0.41 (95% PI 0.25 to 0.58). Supplementary Table 7 presents a full description of these results.

While other images arguably used some scientific messaging (such as image 5 in Supplementary Table 1, “Big Pharma whistleblower: ‘97% of corona vaccine recipients will become infertile’”), the misinformation images identified as having the strongest association with decreased vaccination intent presented a direct link between the COVID-19 vaccine and adverse effects and cited articles and scientific imagery or links to articles purporting to be reputable to strengthen their claim. In the UK, this contrasted with more memetic imaging (for example, ‘striking images with text superimposed on top’<sup>42</sup>) which showed far weaker associations (images 3 and 4 in Supplementary Table 1).

**Discussion**

Using individual-level survey data collected from nationally representative samples of 4,000 and 4,001 respondents in the UK and the

USA, respectively, we reveal a number of key findings of importance to policymakers and stakeholders engaged in either public health communication or the design of vaccine-rollout programmes. We find that, as of September 2020, only 54.1% (95% PI 52.5 to 55.7) of the public in the UK and 42.5% (95% PI 41.0 to 44.1) in the USA would 'definitely' accept a COVID-19 vaccine to protect themselves.

These values are lower than the proportion of vaccinated people required to achieve the anticipated herd immunity levels, suggesting that policymakers may need to convince those unsure about vaccinating to achieve these levels. Higher proportions of individuals in both countries would 'definitely' vaccinate to protect family, friends and at-risk groups, suggesting that effective altruistic messaging may be required to boost uptake. However, we also show that exposure to misinformation lowers individuals' intent to vaccinate to protect themselves and lowers their altruistic intent to vaccinate to protect others, which could complicate messaging campaigns focusing on altruistic behaviours. Campaigns may also have to compete with misinformation purporting to be based in science or medicine, which appears to be particularly damaging to vaccination intentions.

These findings are, however, unlikely to be representative of the effect of misinformation on uptake rates in real-world social media settings. Individuals are unlikely to experience misinformation in the same manner as implemented in this survey, and there will be differences in the volume and rate of misinformation people will be exposed to, depending on their online social media preferences and demographics. A demographic re-weighting would be required to obtain more robust estimates of anticipated COVID-19 vaccine rejection at sub-national or national levels. Misinformation may have already embedded itself in the public's consciousness, and studies have shown that brief exposure to misinformation can embed itself into long-term memory<sup>45</sup>. Policymakers may therefore find challenges ahead to 'undo' the impact it may have already had and to clearly communicate messages surrounding the safety, effectiveness, and importance of the vaccine.

Treatment with exposure to misinformation is found to differentially impact individuals' intent to vaccinate to protect themselves according to some sociodemographic factors. In the UK, the unemployed were more robust to exposure to misinformation compared with those who are employed (before March 2020). Unemployed individuals in the UK were recently found to be less undecided about whether to vaccinate than employed groups<sup>46</sup>. In the USA, 'other' ethnicities and lower-income groups are more robust to misinformation than those of white ethnicity. There is also evidence that exposure to misinformation makes those identifying as Jewish less likely to lower their vaccination intent to protect others compared with Christians in the UK. In the USA, females are more likely than males to lower their intent to vaccinate to protect others upon exposure to misinformation. Many recent studies in both the UK and the USA have highlighted females as less likely to vaccinate than males<sup>46–48</sup>.

We find no evidence that individuals who trust health authorities are any more or less likely to be impacted by misinformation (after controlling for their sociodemographic characteristics); however, trust in experts has been recently found to be associated with intent to pursue COVID-19 vaccine in the USA<sup>49</sup>. Interestingly, trust of celebrities in the UK is associated with more robustness to misinformation compared to controls, whereas trust in family and friends in the USA is associated with a susceptibility to misinformation compared to the control. This result aligns with a recent study that associates trust in non-expert sources with dismissal of misinformation relating to vaccine decision making<sup>50</sup>. Some recent work suggests that those who consume legacy media several times a day and online media less frequently exhibit lower COVID-19 vaccine hesitancy than those who consume less of both<sup>25</sup>. We also find no evidence that daily social media usage is associated with robustness

to effects of misinformation exposure on COVID-19 vaccination intent.

Although our study indicates the possible impact of COVID-19 misinformation campaigns on vaccination intent, this study does not replicate a real-world social media platform environment where information exposure is a complex combination of what is shown to a person by the platform's algorithms and what is shared by their friends or followers<sup>51</sup>. Online social network structures, governed by social homophily, can lead to selective exposure and creation of homogeneous echo chambers<sup>52,53</sup> and polarization of opinions, which may amplify (or dampen) the spread of misinformation among certain demographics. Previous work has shown that there is evidence of echo chambers on real social media platforms around information on vaccines, in general<sup>54,55</sup>. If such information silos also exist for COVID-19 vaccines, then they may lead to self-selection of misinformation or factual information, inducing individuals to become progressively more or less inclined to vaccinate. While our study does not directly quantify such social network effects, it emphasizes on the need to do so further. Furthermore, we find correlational evidence that misinformation identified by our participants after exposure as having the most impact on lowering their vaccination intent was made to have a scientific appeal, such as emphasizing on a direct link between a COVID-19 vaccine and adverse effects while using scientific imagery or links to strengthen their claims. However, our design does not allow causal inferences and we were limited in the type and volume of misinformation presented to respondents. Future research should examine the causal impact of different types of misinformation and identify whether there are other types of misinformation that may be far more impactful on vaccination intent. Therefore, our estimates for the losses in vaccination intent due to misinformation must be placed in the context of this study and the correlational evidence it provides, and caution must be exercised in generalizing these findings to a real-world setting, which may see larger or smaller decreases in vaccination intent depending on the wider context of influencing factors. Addressing the spread of misinformation will probably be a major component of a successful COVID-19 vaccination campaign, particularly given that misinformation on social media has been shown to spread faster than factually correct information<sup>56</sup> and that, even after a brief exposure, misinformation can result in long-term attitudinal and behavioural shifts<sup>45,57</sup> that pro-vaccination messaging may find hard to overcome<sup>57</sup>. With regards to COVID-19, misinformation has even been shown to lead to information avoidance and less systematic processing of COVID-19 information<sup>32</sup>; however, the amplification of questionable sources of COVID-19 misinformation is highly platform dependent, with some platforms amplifying questionable content less than reliable content<sup>58</sup>.

In conclusion, this study reveals that as of September 2020, in both the the UK and the USA, fewer people would 'definitely' take a vaccine than is required for herd immunity, and that misinformation could push these levels further away from herd immunity targets. This analysis provides a platform to help us test and understand how more effective public health communication strategies could be designed and on whom these strategies would have the most positive impact in countering COVID-19 vaccine misinformation.

## Methods

Ethical approval for this study was obtained by the London School of Hygiene and Tropical Medicine ethics committee on 15 June 2020 with reference 22647. A total of 8,001 respondents recruited via an online panel were surveyed by ORB (Gallup) International ([www.orb-international.com](http://www.orb-international.com)) between 7 and 14 September 2020. Respondent quotas for each country and each group (that is both treatment and control) were set according to national demographic distributions for gender, age, and sub-national region—the four census regions in the USA<sup>59</sup> and first level of nomenclature of territorial units in the UK<sup>60</sup>. Following randomized treatment assignment, 3,000 UK and 3,001 US respondents were exposed to images of recently circulating online misinformation related to COVID-19 and vaccines (treatment group) and 1,000 respondents in each country were shown images of



factual information about a COVID-19 vaccine to serve as a randomized control (control group). All respondents exposed to misinformation were debriefed after the survey; debriefing information can be found in the questionnaire included in Supplementary Information. Some respondent characteristics were recorded to reduce their number and facilitate comparison across the two countries. The recoding is provided in Supplementary Table 8 and a breakdown of respondents' characteristics is provided in Supplementary Table 9.

**Selection of images.** To elicit responses that can be most readily interpreted in light of the current state of online misinformation in both the UK and USA, the information shown to respondents—in the form of snippets of social media posts—should satisfy a number of criteria. It should: (1) be recent and relevant to a COVID-19 vaccine; (2) have a high engagement, either through user reach or other publicity, and thus represent information that respondents are not unlikely to be exposed to through social media use; (3) include posts shared by organizations or people with whom respondents are familiar (so that, for example, US and UK audiences are not shown information from people with whom they are unfamiliar); (4) form a distinct set, not replicating content or core messaging, enabling us to probe the most impactful types of misinformation. To this end, we followed a principled approach to select two sets of five images for the treatment and control groups, respectively, combining both quantitative and qualitative methods.

For the treatment set, we used a COVID-19 vaccine-specific Boolean search query—corona\* OR coronavirus OR covid\* OR 'wuhan virus' OR wuhanvirus OR 'Chinese virus' OR 'china virus' OR chinavirus OR 'nCoV\*' OR SARS-CoV\* AND vaccin\* AND (Gates OR 5G OR microchip OR 'New World Order' OR cabal OR globali\*)—to extract COVID-19 vaccine-related online information from 1 June, 2020 to 30 August, 2020 using Meltwater ([www.meltwater.com](http://www.meltwater.com)), an online social media listening platform. This Boolean search term was based on previous research that used similar search terms obtaining the highest levels of user engagement with COVID-19 media and social media articles containing misinformation. This search string returned over 700,000 social media posts that were initially filtered by user engagement and reach to provide the most widely shared and viewed posts. Two independent coders (S.J.P. and K.d.G.) screened top posts and excluded posts that failed criteria 1–4 above. Some posts had relatively low levels of engagement, but were included because they repeatedly appeared in different formats across different outlets and were thus deemed to be influential on social media. Reputable online sources of knowledge were consulted to determine which content was classified as misinformation—that is, information that is regarded false or misleading according to current expert knowledge. A set of five final posts were obtained for the UK and the USA, respectively. For instance, misinformation selected to be shown to the US sample included a post falsely claiming that a COVID-19 vaccine will alter DNA in humans, while that in the UK included a post falsely claiming that COVID-19 vaccine will cause 97% of recipients to become infertile.

In determining the 'control set', the aim was to expose people to factual COVID-19 vaccine information to serve as a control against the treatment exposure of misinformation, since exposure to any information can in principle cause respondents to change their vaccine inclination (that is control information controls for other elements of our survey), and respondents may misreport post-exposure vaccination intent due to recall bias or other between-conditions differences. Factual information was obtained by a coder (S.L.), also via Meltwater, using the same Boolean search term as above, but excluding the last clause containing misinformation-specific search keys, which returned over a hundred posts. Reputable online sources of knowledge were consulted to determine which content is classified as factual information—that is, information that is regarded correct as per current expert knowledge. A set of five final posts was obtained, common to both the UK and the USA. Information was often from authoritative sources (or otherwise referenced to authoritative sources) such as vaccine groups and scientific organizations. We ensured that these five posts were not overtly pro-vaccination and did not reference anti-vaccination campaigns or materials. For instance, information presented included an update on the current state of COVID-19 vaccine trials; the importance of a vaccine to get out of the COVID-19 pandemic; and how a candidate vaccine generates a good immune response. Supplementary Table 1 presents further details regarding the treatment and control image sets, including detailed explanations for classification of posts as misinformation or factual information.

**Estimating treatment effects.** In this study, the outcome of interest, vaccination intent, is measured on a four-level ordered scale. Using a classical approach in the potential outcomes framework<sup>61,62</sup> to determine treatment effects would either necessitate binarizing the outcome—which can lead to loss of information about vaccination intent—or making a strong assumption of linearity of the outcome scale. Therefore, a hierarchical Bayesian ordered logistic regression framework is used here to estimate the impact of (1) treatment of misinformation on change in vaccination intent relative to factual information, and (2) how these treatments differentially impact individuals by their sociodemographic characteristics (that is, HTEs). Full model details, including the statistics used to describe all effects, are detailed below. Throughout, the following notation is used: pre- ( $W$ ) and post-exposure ( $Y$ ) intents to accept a COVID-19 vaccine  $W, Y \in \{1, 2, 3, 4\}$

are ordered variables ('yes, definitely' (1); 'unsure, but leaning towards yes' (2); 'unsure, but leaning towards no' (3), and 'no, definitely not' (4)); treatment group  $G \in \{T, C\}$  where  $T$  denotes the treatment of exposure to misinformation and  $C$  is the control of exposure to factual information; and covariates (for example sociodemographics) are given by  $X$ .

Since vaccination intent is modelled as an ordered variable, one can expect the treatment to impact vaccination intent monotonically. To this end,  $W$  is modelled as a monotonic ordered predictor<sup>63,64</sup>. Using  $W$  as a predictor for  $Y$  has two advantages: it (1) controls for sampling discrepancies between the treatment and control groups, and (2) allows for the treatment to differentially affect those with different prior vaccination intents. We use ordered logistic regression<sup>65</sup> to model  $Y$  conditional on  $G, W$  and covariates  $X$ . Then for an individual respondent  $i$  we can write:

$$W(i) \sim \text{OrderedLogistic}(0, (\kappa_1, \kappa_2, \kappa_3)), \tag{1}$$

$$Y(i) | (G(i) = g, W(i), X(i)) \sim \text{OrderedLogistic} \left( \beta_W^g \sum_{j=1}^{W(i)-1} \delta_j^g + f(X(i); \beta_X^g), (\alpha_1^g, \alpha_2^g, \alpha_3^g) \right). \tag{2}$$

Where  $\beta_W^g \in \mathbb{R}, \delta_j^g \in \mathbb{R}_{\geq 0}$ , such that  $\sum_{j=1}^3 \delta_j^g = 1, -\infty < \alpha_1^g < \alpha_2^g < \alpha_3^g < \infty$  and  $-\infty < \kappa_1 < \kappa_2 < \kappa_3 < \infty$ . We use the ordered logistic distribution for  $k$  outcomes specified by  $Z \sim \text{OrderedLogistic}(\beta, (\alpha_1, \alpha_2, \dots, \alpha_{k-1}))$ , where  $P(Z \leq j) = \sigma(\alpha_j - \beta)$  and  $\sigma(x) = 1/(1 + e^{-x})$  is the standard logistic sigmoid function. We remark that this operationalizes a proportional-odds assumption<sup>65</sup>, wherein the difference in log of cumulative odds ratios between successive categories is independent of the slope  $\beta$ , that is,  $\forall j \in \{2, 3, \dots, k-1\}$ , we have:  $\log \left( \frac{P(Z \leq j)}{P(Z > j)} \right) - \log \left( \frac{P(Z \leq j-1)}{P(Z > j-1)} \right) = \alpha_j - \alpha_{j-1}$ .

This modelling framework allows us to model (1) the effect of treatment on vaccination intent and (2) the HTEs through the function  $f$ . In estimating (average) treatment effects (1),  $f=0$  (we still need to control for pre-exposure intent  $W$ ); whereas when estimating HTEs (2), we assume:

$$f(X(i)) = \sum_{d \in D} \beta_{d(i)}^g + \sum_{u \in U} \beta_{u(i)}^g,$$

where  $D = \{AGE, GEN, EDU, EMP, REL, POL, ETH, INC\}$  refers to the set of sociodemographic characteristics—of age (AGE), gender (GEN), highest education qualification received (EDU), (pre-pandemic) employment status (EMP), religion (REL), political affiliation (POL), ethnicity (ETH) and income (INC)—such that  $\forall d \in D, d(i)$  corresponds to the category to which  $i$  belongs and  $\beta_{d(i)}^g \in \mathbb{R}$  refers to the slope for that category. Specification of the set  $U$  allows us to investigate the HTEs for (1) sociodemographic characteristics ( $U = \{\}$ ); (2) social media use (while controlling for possible confounding effects of sociodemographics); and (3) sources of trust (while controlling for possible confounding effects of sociodemographics). We thus investigate HTEs for social media use and sources of trusted information about COVID-19 by specifying (1)  $U = \{SOCIAL\}$  (where  $\beta_{SOCIAL(i)}^g \in \mathbb{R}$  refers to the slope when  $SOCIAL(i)$  indicates the category of amount of daily social media usage for  $i$ ) and (2)  $U = \{TRUST_1, TRUST_2, \dots, TRUST_k\}$  for  $k$  different sources of information (where  $\beta_{TRUST_k(i)}^g \in \mathbb{R}$  refers to the slope when  $TRUST_k(i)$  is the category indicating whether  $i$  trusts the  $k$ th source of information: 1 for no and 2 for yes). We remark that the model for  $Y$  specified in equation (2) is equivalent to a traditional linear two-way interaction model for causal estimation under a binary treatment, composed with a logistic sigmoid function to model the cumulative distribution of the ordinal categorical outcome variable<sup>66</sup>.

Regularizing hierarchical priors are placed on all primary model parameters to aid model identifiability and prevent detection of spurious treatment effects:  $\alpha_j^g \sim \text{Normal}(\alpha_j, 1), \beta_z^g \sim \text{Normal}(\beta_z, 1), \forall z \in D \cup U, \beta_w^g \sim \text{Normal}(\beta_w, 1)$  and  $\kappa_j, \alpha_j, \beta_z, \beta_w \sim \text{Normal}(0, 1)$ . Non-informative hierarchical priors were placed on  $\delta_j^g: (\delta_1^g, \delta_2^g, \delta_3^g) \sim \text{Dirichlet}((\delta_1, \delta_2, \delta_3))$  and  $\delta_j \sim \text{Exponential}(1)$ .

**Statistics for measuring treatment effects.** We are interested in measuring the causal effect of exposure to misinformation ( $G=T$ ), relative to the control of exposure to factual information ( $G=C$ ), on (post-exposure) vaccination intent  $Y$ . When computing average treatment effects, it would be conventional to calculate a difference in conditional expectations, that is  $E[Y|T] - E[Y|C]$ . However, as  $Y$  is an ordered categorical, a conditional expectation has no meaningful interpretation. Therefore, we can compute a conditional probability mass function,  $P(Y|G)$ , and define a corresponding statistic for treatment effect<sup>67</sup> on vaccination intent as:

$$\Delta(Y) \equiv P(Y = y|G = T) - P(Y = y|G = C). \tag{3}$$

Since treatment may also depend on individuals' pre-exposure vaccination intent  $W$ , we also compute the statistic:

$$\Delta_W(y; w) \equiv P(Y = y|G = T, W = w) - P(Y = y|G = C, W = w), \tag{4}$$

for  $\forall (y, w) \in \{1, 2, 3, 4\} \times \{1, 2, 3, 4\}$ . Using the ordered logistic regression model specified in equations (1) and (2), these two

statistics are given by  $\Delta(y) = \sum_{w=1}^4 \mu(w)(\rho(T; y, w) - \rho(C; y, w))$  and  $\Delta_W(y; w) = \rho(T; y, w) - \rho(C; y, w)$ , where

$$\rho(g; y, w) \equiv P(Y = y | W = w, G = g) = \sigma\left(\alpha_y^g - \beta_W^g \sum_{j=1}^{w-1} \delta_j^g\right) - \sigma\left(\alpha_y^g - \beta_W^g \sum_{j=1}^{w-1} \delta_j^g\right) \text{ and } \mu(w) \equiv P(W = w) = \sigma(\kappa_w) - \sigma(\kappa_{w-1}).$$

The interpretation of equations (3) and (4) is as follows. (3): If  $\Delta(y) > 0$  ( $\Delta(y) < 0$ ), then the treatment induces an average individual with vaccination intent  $y$  to not change their vaccination intent (to change their vaccination intent)—relative to control. Alternately,  $100 \times \Delta(y)$  indicates the percentage point change in the number of people with intent  $y$  after exposure to misinformation, relative to factual information. (4): If  $\Delta_W(y; w) > 0$  ( $\Delta_W(y; w) < 0$ ), then the treatment induces an average individual to (not) change their vaccination intent from  $w$  to  $y$ —relative to control. Alternately,  $100 \times \Delta_W(y; w)$  indicates the percentage point rise or drop in the number of people with prior intent  $w$  who change it to  $y$  after exposure to misinformation, relative to factual information.

Statistics for measuring HTEs treatment effects may depend on sociodemographic groups: misinformation or factual information may cause some sociodemographic groups to be more or less likely to vaccinate than others. Following the conditional probability mass function framework, these HTEs would correspond to computing the following conditional statistic:

$$P(Y = y | G = T, X = x) - P(Y = y | G = C, X = x). \quad (5)$$

Because we consider many covariates, in the interest of being concise we cannot estimate conditional treatment effects for every multivariate combination of covariates. However, some progress can be made by considering the following modifications. Firstly, we can compute a different statistic that still permits a form expressed as the linear difference of a function over treatments and controls separately. In particular, since vaccination intent is ordered, we can define a statistic conveniently in terms of the conditional cumulative distribution function. More precisely, consider the negative logarithm of cumulative odds ratio  $\theta(g, x; y) \equiv -\log\left(\frac{P(Y \leq y | G=g, X=x)}{P(Y > y | G=g, X=x)}\right)$ , which indicates how likely an individual  $x$  is to have a vaccination intent up to level  $y$  after exposure to misinformation ( $G = T$ ) or factual information ( $G = C$ ). The larger this statistic is for given  $y$ , the less likely the individual  $x$  in group  $g$  is to have a high vaccination intent. Given the ordered logistic model specification in equations (1) and (2), and by considering this estimand on the latent scale of log of cumulative odds ratio, one can use this latent continuous variable  $\theta$  as the de-facto outcome instead of  $Y$  (ref. <sup>68</sup>). Estimands on the latent scale are more difficult to interpret due to non-identifiability of the function mapping  $\theta$  to  $Y$ , but the use of regularizing priors in the model makes the function identifiable. Secondly, since we are interested in finding whether a group is more or less susceptible to treatment effects, we can do so by picking a reference group  $x_0$ , that is, we compute a difference of conditional treatment effects when  $X = x$  relative to when  $X = x_0$ . Therefore, we are interested in a statistic which is the difference in log cumulative odds ratios (or simply log odds ratios),  $\eta(g, x; x_0) \equiv \theta(g, x; y) - \theta(g, x_0; y)$ . This leads to the following (relative) measure for heterogeneous effects:

$$\Delta_X(x; x_0) \equiv \eta(T, x; x_0) - \eta(C, x; x_0). \quad (6)$$

Given the model definition in equations (1) and (2), the statistic  $\eta$  is given by  $\eta(g, x; x_0) = f(x; \beta_X^g) - f(x_0; \beta_X^g)$ , which is simply given by a difference in the log cumulative odds ratios. For example, the statistic of difference in log odds for gender is  $\eta(g, \text{Female}; \text{Male}) = \beta_{\text{Female}}^g - \beta_{\text{Male}}^g$  and the corresponding heterogeneous effect is given by  $\Delta_X(\text{Female}; \text{Male}) = (\beta_{\text{Female}}^T - \beta_{\text{Female}}^C) - (\beta_{\text{Male}}^T - \beta_{\text{Male}}^C)$ . In this model, parameters exist for every sociodemographic group—with regularizing priors ensuring identifiability—allowing for posterior contrast distributions of  $\eta$  and  $\Delta$  with regards to any reference group  $x_0$ . In our analysis, for categorical characteristics we pick the most populated—‘employed’ for employment, ‘Christian’ for religion, ‘white’ for ethnicity—or second-to-most populated group—‘male’ for gender, ‘Conservative’ (UK) or ‘Republican’ (USA) for political affiliation—as the reference group, which allows for a natural comparison of minority groups against the majority. For ordinal characteristics, we pick one of the end groups as the reference group—‘18–24’ (lowest) for age, ‘level 4’ (highest) for income, ‘level 4’ (highest) for education, ‘none’ (lowest) for social media use—which allows for a natural comparison to the extrema of the characteristic. For binary characteristics, we pick the null group as the reference—indicating no trust in a source of COVID-19 information.

The interpretation of the HTE (equation (6)) is as follows. If  $\Delta_X(x; x_0) > 0$  ( $\Delta_X(x; x_0) < 0$ ), then the treatment makes an individual of group  $x$  more (less) likely to move from lower vaccine hesitancy to a higher one when compared to an individual of group  $x_0$ —relative to control. The interpretation of equation (5) is as follows: if  $\eta(g, x; x_0) > 0$  ( $\eta(g, x; x_0) < 0$ ), then the exposure within treatment group  $g$  makes an individual of group  $x$  more (less) likely to move from lower vaccine hesitancy to a higher one when compared to an individual of group  $x_0$ .

**Estimating image impact.** To study which images, corresponding to misinformation or factual information, are perceived by participants to induce a larger drop in vaccination intent upon exposure, we make use of ratings given by the respondents

to each of the 5 images presented along 5 different perception metrics as features to learn how each image metric and each image itself contributes to the measured drop in vaccination intent. As before, let  $W$  denote pre-exposure intent,  $Y$  is the post-exposure intent and  $G$  is the treatment group. Furthermore, let  $X(i)$  refer to the  $5 \times 5$  matrix such that  $X_{jk}(i)$  refers to the  $i$ th individual's rating on the  $j$ th image metric for the  $k$ th image. Then, the model definition here is very similar to when pursuing HTEs analysis, except the function of covariates now corresponds to an aggregation of ratings across images and image metrics:

$$Y(i) | (G(i) = g, W(i), X(i)) \sim \text{OrderedLogistic} \left( \beta_W^g \sum_{j=1}^{W(i)-1} \delta_j^g + \sum_{j=1}^5 \sum_{k=1}^5 \beta_{jk}^g X_{jk}(i) \gamma_k^g, (\alpha_1^g, \alpha_2^g, \alpha_3^g) \right), \quad (7)$$

where  $\beta_W^g, \beta_{jk}^g \in \mathbb{R}, \delta_j^g \in \mathbb{R}_{\geq 0}$  such that  $\sum_{j=1}^3 \delta_j^g = 1, \gamma_k^g \in \mathbb{R}_{\geq 0}$  such that  $\sum_{k=1}^5 \gamma_k^g = 1$ , and  $-\infty < \alpha_1^g < \alpha_2^g < \alpha_3^g < \infty$ . As noted above,  $X_{jk}(i)$  indicates the Likert response of the  $i$ th individual's rating on the  $j$ th image metric for the  $k$ th image. Here, we assume a signed response  $X_{jk}(i) \in \{-2, -1, 0, 1, 2\}$  corresponding to the negative and positive ratings of a five-level Likert scale—those reporting ‘do not know’ were included in the response category of 0. This allows us to gauge both (1) which images have the most impact (from  $\gamma_k^g$ ) and (2) which image metrics or features have the most impact (from  $\beta_{jk}^g$ ).

The image metrics considered are, in order, whether (1) the respondents perceived the image to have made them less inclined to vaccinate, (2) they agreed with the image, (3) they found the image trustworthy, (4) they were likely to fact check the information shown in the image, and (5) they were likely to share the image. Regularizing priors are placed on all primary model parameters:  $\alpha_1^g, \beta_{jk}^g, \beta_W^g \sim \text{Normal}(0, 1)$ . Non-informative priors are placed on  $\gamma_k^g, \delta_j^g$ :  $(\gamma_1^g, \gamma_2^g, \gamma_3^g, \gamma_4^g, \gamma_5^g) \sim \text{Dirichlet}((1, 1, 1, 1, 1))$  and  $(\delta_1^g, \delta_2^g, \delta_3^g) \sim \text{Dirichlet}((1, 1, 1))$ . The statistics reported in Supplementary Table 7 refer to  $\beta_{jk}^g$  and  $\gamma_k^g$ . If  $\beta_{jk}^g > 0$  ( $\beta_{jk}^g < 0$ ), then a higher (lower) rating on the  $j$ th metric is more predictive of a drop in vaccination intent in treatment group  $g$ , after exposure. Since five images were shown to each respondent,  $\gamma_k^g > 0.2$  ( $\gamma_k^g < 0.2$ ) indicates that the  $k$ th image contributes more (less) to the drop in vaccination intent in treatment group  $g$ , after exposure, than that would be expected at random.

**Statistical inference.** Model inference was performed by Hamiltonian Monte Carlo with the NUTS sampler using PyStan<sup>69</sup>, the Python implementation of Stan. Samples from the posterior distribution of the model parameters were collected from 4 chains and 2,000 iterations (that is, 4,000 samples excluding warm-up) after ensuring model convergence, with the potential scale reduction factor satisfying  $\hat{R} \leq 1.02$  for all model parameters, while ensuring that the smallest effective sample size for all model parameters is greater than 500 (refs. <sup>70,71</sup>) (Supplementary Table 10). The target average proposal acceptance probability for the NUTS sampler was set to 0.9 and increased to 0.99 to remove any divergent transitions if they were encountered. The maximum tree depth for the sampler was set to 10 but increased to 15 if the limit was reached for any model. Relevant statistics for parameters of interest (coefficients, contrasts, log odds ratios, percentages and weights) were extracted from the samples, and all results report the mean estimate—the effect size—alongside 95% PIs (that is, values at 2.5% and 97.5% percentiles) to indicate credible values of the statistic.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The data used in this study are available at <https://github.com/sloomba/covid19-misinfo/>. A copy of the materials used in this study, as displayed to respondents, can be obtained from the authors upon request. Source data are provided with this paper.

## Code availability

The code developed for this study is available at <https://github.com/sloomba/covid19-misinfo/>.

Received: 19 October 2020; Accepted: 19 January 2021;

Published online: 5 February 2021

## References

- Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C. & Garry, R. F. The proximal origin of SARS-CoV-2. *Nat. Med.* **26**, 450–452 (2020).
- Zhou, P. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
- World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19—11 March 2020 <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020> (2020).

4. Jackson, L. A. et al. An mRNA vaccine against SARS-CoV-2—preliminary report. *N. Engl. J. Med.* **383**, 1920–1931 (2020).
5. Mulligan, M. J. et al. Phase I/II study of COVID-19 RNA vaccine BNT162b1 in adults. *Nature* **586**, 589–593 (2020).
6. Folegatti, P. M. et al. Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2: a preliminary report of a phase 1/2, single-blind, randomised controlled trial. *Lancet* **396**, 467–478 (2020).
7. Pfizer and BioNTech conclude phase 3 study of COVID-19 vaccine candidate, meeting all primary efficacy endpoints <https://www.pfizer.com/news/press-release/press-release-detail/pfizer-and-biontech-conclude-phase-3-study-covid-19-vaccine> (2020).
8. US National Institute of Allergy and Infectious Diseases. Promising interim results from clinical trial of NIH-Moderna COVID-19 vaccine <https://www.nih.gov/news-events/news-releases/promising-interim-results-clinical-trial-nih-moderna-covid-19-vaccine> (2020).
9. Voysey, M. et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *Lancet* [https://doi.org/10.1016/S0140-6736\(20\)32661-1](https://doi.org/10.1016/S0140-6736(20)32661-1) (2020).
10. US Food and Drug Administration. FDA takes key action in fight against COVID-19 By issuing emergency use authorization for first COVID-19 vaccine <https://www.fda.gov/news-events/press-announcements/fda-takes-key-action-fight-against-covid-19-issuing-emergency-use-authorization-first-covid-19> (2020).
11. US Food and Drug Administration. FDA takes additional action in fight against COVID-19 by issuing emergency use authorization for second COVID-19 vaccine <https://www.fda.gov/news-events/press-announcements/fda-takes-additional-action-fight-against-covid-19-issuing-emergency-use-authorization-second-covid-19> (2020).
12. UK Medicines and Healthcare products Regulatory Agency. UK medicines regulator gives approval for first UK COVID-19 vaccine <https://www.gov.uk/government/news/uk-medicines-regulator-gives-approval-for-first-uk-covid-19-vaccine> (2020).
13. UK Department of Health and Social Care. Oxford University/AstraZeneca vaccine authorised by UK medicines regulator <https://www.gov.uk/government/news/oxford-universityastrazeneca-vaccine-authorised-by-uk-medicines-regulator> (2020).
14. European Medicines Agency. EMA recommends first COVID-19 vaccine for authorisation in the EU <https://www.ema.europa.eu/en/news/ema-recommends-first-covid-19-vaccine-authorisation-eu> (2020).
15. Kwok, K. O., Lai, F., Wei, W. I., Wong, S. Y. S. & Tang, J. W. T. Herd immunity—estimating the level required to halt the COVID-19 epidemics in affected countries. *J. Infect.* **80**, e32–e33 (2020).
16. Sanche, S. et al. High contagiousness and rapid spread of severe acute respiratory syndrome coronavirus 2. *Emerg. Infect. Dis.* **26**, 1470–1477 (2020).
17. de Figueiredo, A., Simas, C., Karafillakis, E., Paterson, P. & Larson, H. J. Mapping global trends in vaccine confidence and investigating barriers to vaccine uptake: a large-scale retrospective temporal modelling study. *Lancet* **396**, 898–908 (2020).
18. Bellaby, P. Communication and miscommunication of risk: understanding UK parents' attitudes to combined MMR vaccination. *BMJ* **327**, 725–728 (2003).
19. Burki, T. Vaccine misinformation and social media. *Lancet Digit. Health* **1**, E258–E259 (2019).
20. Lo, N. C. & Hotez, P. J. Public health and economic consequences of vaccine hesitancy for measles in the United States. *JAMA Pediatr.* **171**, 887–892 (2017).
21. Megget, K. Even COVID-19 can't kill the anti-vaccination movement. *BMJ* **369**, m2184 (2020).
22. Geldsetzer, P. Knowledge and perceptions of COVID-19 among the general public in the United States and the United Kingdom: a cross-sectional online survey. *Ann. Intern. Med.* **173**, 157–160 (2020).
23. Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G. & Rand, D. G. Fighting COVID-19 misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention. *Psychol. Sci.* **31**, 770–780 (2020).
24. Salathé, M. & Bonhoeffer, S. The effect of opinion clustering on disease outbreaks. *J. R. Soc. Interface* **5**, 1505–1508 (2008).
25. McAndrew, S. & Allington, D. Mode and frequency of COVID-19 information updates, political values, and future covid-19 vaccine attitudes. Preprint <https://doi.org/10.31234/osf.io/j7srx> (2020).
26. Tyson, A., Johnson, C. & Funk, C. U.S. public now divided over whether to get COVID-19 vaccine. *Pew Research Center* <https://www.pewresearch.org/science/2020/09/17/u-s-public-now-divided-over-whether-to-get-covid-19-vaccine/> (2020).
27. Peretti-Watel, P. et al. A future vaccination campaign against COVID-19 at risk of vaccine hesitancy and politicisation. *Lancet Infect. Dis.* **20**, 769–770 (2020).
28. Galvão, J. COVID-19: the deadly threat of misinformation. *Lancet Infect. Dis.* [https://doi.org/10.1016/S1473-3099\(20\)30721-0](https://doi.org/10.1016/S1473-3099(20)30721-0) (2020).
29. Altmann, D. M., Douek, D. C. & Boyton, R. J. What policy makers need to know about COVID-19 protective immunity. *Lancet* **395**, 1527–1529 (2020).
30. Zarocostas, J. How to fight an infodemic. *Lancet* **395**, 676 (2020).
31. Islam, M. S. et al. COVID-19-related infodemic and its impact on public health: a global social media analysis. *Am. J. Trop. Med. Hyg.* **103**, 1621–1629 (2020).
32. Kim, H. K., Ahn, J., Atkinson, L. & Kahlor, L. A. Effects of COVID-19 misinformation on information seeking, avoidance, and processing: a multicountry comparative study. *Sci Commun* **42**, <https://doi.org/10.1177/1075547020959670> (2020).
33. Kreps, S. E. & Kriner, D. Medical misinformation in the COVID-19 pandemic. *SSRN* <https://doi.org/10.2139/ssrn.3624510> (2020).
34. Murphy, J., Vallières, F. & Bentall, R.P. et al. Psychological characteristics associated with COVID-19 vaccine hesitancy and resistance in Ireland and the United Kingdom. *Nat. Commun.* **12**, 29 (2021).
35. Roozenbeek, J. et al. Susceptibility to misinformation about COVID-19 around the world. *R. Soc. Open Sci.* **7**, 201199 (2020).
36. Romer, D. & Jamieson, K. H. Conspiracy theories as barriers to controlling the spread of COVID-19 in the U.S. *Soc. Sci. Med.* **263**, 113356 (2020).
37. Ernala, S. K., Burke, M., Leavitt, A. & Ellison, N. B. How well do people report time spent on Facebook? In *Proc. 2020 CHI Conference on Human Factors in Computing Systems* (ACM, 2020); <https://doi.org/10.1145/3313831.3376435>
38. National Health Service England. Protect your loved ones <https://www.nhs.uk/apps-library/nhs-covid-19/> (2020).
39. National Health Service England. We must keep on protecting each other <https://coronavirusresources.phe.gov.uk/Hands-Face-Space-/resources/co-branded-posters/> (2020).
40. US Centers for Disease Control and Prevention. *Do it for yourself and your friends.* <https://www.cdc.gov/coronavirus/2019-ncov/downloads/Young-Mitigation-recommendations-resources-toolkit-COL.pdf> (2020).
41. US Centers for Disease Control and Prevention. *Wear a mask to protect you and your friends.* <https://www.cdc.gov/coronavirus/2019-ncov/downloads/Young-Mitigation-recommendations-resources-toolkit.pdf> (2020).
42. Lazer, D. M. J. et al. The science of fake news. *Science* **359**, 1094–1096 (2018).
43. Vraga, E. K. & Bode, L. Defining misinformation and understanding its bounded nature: Using expertise and evidence for describing misinformation. *Political Commun.* **37**, 136–144 (2020).
44. Graham, M. & Coppock, A. Asking about attitude change. *Public Opinion Quarterly* (in the press).
45. Zhu, B. et al. Brief exposure to misinformation can lead to long-term false memories. *Appl. Cogn. Psychol.* **26**, 301–307 (2012).
46. Paul, E., Steptoe, A. & Fancourt, D. Attitudes towards vaccines and intention to vaccinate against COVID-19: implications for public health communications. *Lancet Reg. Health Eur.* <https://doi.org/10.1016/j.lanepe.2020.100012> (2020).
47. Malik, A. A., McFadden, S. M., Elharake, J. & Omer, S. B. Determinants of COVID-19 vaccine acceptance in the US. *EClinicalMedicine* **26**, 100495 (2020).
48. de Figueiredo, A. Sub-national forecasts of COVID-19 vaccine acceptance across the UK: a large-scale cross-sectional spatial modelling study. Preprint at *medRxiv* <https://doi.org/10.1101/2020.12.17.20248382> (2020).
49. Callaghan, T. et al. Correlates and disparities of COVID-19 vaccine hesitancy. *SSRN* <https://doi.org/10.2139/ssrn.3667971> (2020).
50. Motta, M., Callaghan, T. & Sylvester, S. Knowing less but presuming more: Dunning-Kruger effects and the endorsement of anti-vaccine policy attitudes. *Soc. Sci. Med.* **211**, 274–281 (2018).
51. Bakshy, E., Messing, S. & Adamic, L. A. Exposure to ideologically diverse news and opinion on Facebook. *Science* **348**, 1130–1132 (2015).
52. Del Vicario, M. et al. The spreading of misinformation online. *Proc. Natl Acad. Sci. USA* **113**, 554–559 (2016).
53. Del Vicario, M., Quattrociocchi, F., Scala, S. & Zollo, F. Polarization and fake news: early warning of potential misinformation targets. *ACM Trans. Web* **13**, 10 (2019).
54. Schmidt, A. L., Zollo, F., Scala, A., Betsch, C. & Quattrociocchi, W. Polarization of the vaccination debate on Facebook. *Vaccine* **36**, 3606–3612 (2018).
55. Cossard, A. et al. Falling into the echo chamber: the Italian vaccination debate on Twitter. In *Proc. 14th International AAAI Conference on Web and Social Media* 130–140 (AAAI, 2020).
56. Vosoughi, S., Roy, D. & Aral, S. The spread of true and false news online. *Science* **359**, 1146–1151 (2018).
57. Pluviano, S., Watt, C. & della Sala, S. Misinformation lingers in memory: failure of three pro-vaccination strategies. *PLOS ONE* **12**, e0181640 (2017).
58. Cinelli, M. et al. The COVID-19 social media infodemic. *Sci. Rep.* **10**, 16598 (2020).
59. US Census Bureau. Census regions and divisions of the United States. [https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us\\_regdiv.pdf](https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf) (2010).

60. An Overview of the 3 NUTS and 2 LAU Layers in the UK (UK Office for National Statistics, accessed September 2020); <https://www.ons.gov.uk/methodology/geography/ukgeographies/eurostat>
61. Rubin, D. B. Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* **66**, 688–701 (1974).
62. Holland, P. W. Statistics and causal inference. *J. Am. Stat. Assoc.* **81**, 945–960 (1986).
63. Bürkner, P. & Charpentier, E. Modelling monotonic effects of ordinal predictors in Bayesian regression models. *Br. J. Math. Stat. Psychol.* **73**, 42–451 (2020).
64. McElreath, R. *Statistical Rethinking: A Bayesian Course with Examples in R and STAN* (CRC Press, 2020).
65. McCullagh, P. Regression models for ordinal data. *J. R. Stat. Soc.* **42**, 109–142 (1980).
66. Wang, C., Dominici, F., Parmigiani, G. & Zigler, C. M. Accounting for uncertainty in confounder and effect modifier selection when estimating average causal effects in generalized linear models. *Biometrics* **71**, 654–6565 (2015).
67. Boes, S. Nonparametric analysis of treatment effects in ordered response models. *Empir. Econ.* **44**, 81–109 (2013).
68. Volfovsky, A., Airoidi, E. M. & Rubin, D. B. Causal inference for ordinal outcomes. Preprint at *arXiv* <https://arxiv.org/abs/1501.01234> (2015).
69. PyStan: the Python Interface to Stan Version 2.17.1.0 (Stan Development Team, 2019).
70. Gelman, A. & Rubin, D. B. Inference from iterative simulation using multiple sequences. *Stat. Sci.* **7**, 457–472 (1992).
71. Vehtari, A., Gelman, A., Simpson, D., Carpenter, B. & Bürkner, P.-C. Rank-normalization, folding, and localization: an improved  $\hat{R}$  for assessing convergence of MCMC. *Bayesian Anal.* <https://doi.org/10.1214/20-BA1221> (2020).

### Acknowledgements

This project was funded by the United Nations' Verified Initiative. The funders had no role in data collection, questionnaire design, data analysis, data interpretation or writing of this report. The corresponding authors had full access to all the data in the

study and had final responsibility for the decision to submit for publication. The Vaccine Confidence Project acknowledges its long partnership and collaboration with ORB (Gallup) International who collected all data for this study.

### Author contributions

All authors contributed to questionnaire design, with A.d.F. obtaining ethical approval for the study via the London School of Hygiene and Tropical Medicine ethics committee. S.L. and A.d.F. designed the statistical analyses. S.L. performed all model calculations and inference and created all figures for publication. S.J.P., K.d.G. and S.L. performed image selection. A.d.F. and S.L. wrote the final manuscript with input from all authors. All authors contributed to the interpretation of the results. A.d.F. and H.J.L. supervised the project.

### Competing interests

H.J.L. and A.d.F. are involved in Vaccine Confidence Project collaborative grants with GlaxoSmithKline, Janssen and Merck. S.J.P. and H.J.L. are involved with Vaccine Confidence Project collaborative grants with Janssen. H.J.L. has also received honoraria as a member of the Merck Vaccine Confidence Advisory Board and GlaxoSmithKline advisory roundtables. H.J.L. is a member of the Merck Vaccine Confidence Advisory Board.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41562-021-01056-1>.

**Correspondence and requests for materials** should be addressed to A.d.F. or H.J.L.

**Peer review information** *Nature Human Behaviour* thanks Paul Bürkner, Matt Motta and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Stavroula Kousta.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2021

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- |     |           |
|-----|-----------|
| n/a | Confirmed |
|-----|-----------|
- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
  - A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
  - The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
  - A description of all covariates tested
  - A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
  - A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
  - For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
  - For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
  - For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
  - Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

n/a

Data analysis

Python version 3.7.3 was used for all analysis with the following libraries:

```
-- pystan
-- pandas
-- numpy
-- matplotlib
```

The custom code for this project is available at <https://github.com/sloomba/covid19-misinfo/>.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data used in this study are available at <https://github.com/sloomba/covid19-misinfo/>.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences  Behavioural & social sciences  Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Quantitative cross-sectional study
Research sample	Nationally representative (by age, sex, subnational region) samples are obtained via online panels of participants in the UK and US. 4,000 respondents in both countries are surveyed to quantify national-level summaries of individuals' vaccine intent and susceptibility of misinformation. All respondents surveyed were 18 or over and informed consent was obtained by ORB (Gallup) International.
Sampling strategy	Survey participants were recruited via online survey panels by ORB (Gallup) International using fixed quotas at the national level for sex, age, and sub-national region.
Data collection	n/a
Timing	The survey was in field between 7 and 14 September, 2020.
Data exclusions	No data was excluded from the analyses.
Non-participation	Individuals who did not wish to participate simply did not complete the survey. We (the authors) do not have access to this information as ORB (Gallup) International provided a quota-matched sample of respondents.
Randomization	Of the 4,000 respondents in each country, 75% were exposed to misinformation, while 25% were exposed to factual information.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

### Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	See above
Recruitment	Participants were recruited via an online panel of respondents (see above). There may be technological literacy biases associated with this sample, however, we do control for highest level of education.
Ethics oversight	Approval for this study was obtained by the LSHTM ethics committee on 15 June 2020 with reference 22130.

Note that full information on the approval of the study protocol must also be provided in the manuscript.