RESOURCE
https://doi.org/10.1038/s41562-022-01347-1

nature human behaviour

Global survey on COVID-19 beliefs, behaviours and norms

Avinash Collis[1,11], Kiran Garimella[2,11], Alex Moehring[3,4,11], M. Amin Rahimian[5,11], Stella Babalola[6,7], Nina H. Gobat[8], Dominick Shattuck[6], Jeni Stolow[9], Sinan Aral[3,4,10] ✉ and Dean Eckles[3,4,10] ✉

Policy and communication responses to COVID-19 can benefit from better understanding of people's baseline and resulting beliefs, behaviours and norms. From July 2020 to March 2021, we fielded a global survey on these topics in 67 countries yielding over 2 million responses. This paper provides an overview of the motivation behind the survey design, details the sampling and weighting designed to make the results representative of populations of interest and presents some insights learned from the survey. Several studies have already used the survey data to analyse risk perception, attitudes towards mask wearing and other preventive behaviours, as well as trust in information sources across communities worldwide. This resource can open new areas of enquiry in public health, communication and economic policy by leveraging large-scale, rich survey datasets on beliefs, behaviours and norms during a global pandemic.

Non-pharmaceutical interventions in response to COVID-19 often depend for their effectiveness on the behavioural responses of the public. Even with a vaccine, uptake is not entirely in the control of experts and policy-makers. Rather than being a small factor, there is growing evidence that the prevention behaviours of people are dramatically influenced by many social and cultural factors[1–3]. Analyses of mobility data reveal that the movements of people are predicted and perhaps caused by their partisan affiliation[4], media consumption[5] and the behaviours of their social networks[6]. Thus, the epidemiological and economic effects of policies that close (or open) businesses and schools are substantially determined by people's beliefs. This is consistent with the recognition, at least among public health experts, that health communication is a core part of effective response to epidemics, ideally in concert with other policies and interventions. However, developing and deploying effective policies and communication strategies demands data about people's beliefs and how they have been affected by prior exposure to information from governments, peers and media—and these data are largely lacking, even as massive troves of medical and behavioural traces are used by researchers[3].

This motivated us to conduct a large-scale, international survey related to COVID-19 in 67 countries (Fig. 1 maps the countries included) to help policy-makers and researchers better monitor and understand people's knowledge, beliefs, behaviours, norms and risk perceptions across the world through a collaboration with Facebook and Johns Hopkins University and with input from experts at the World Health Organization and the Global Outbreak Alert and Response Network. The survey is organized into blocks on the basis of the question topics. Every survey begins with questions from the same five blocks: information exposure, knowledge, vaccine and healthcare and demographics. In 'snapshot' countries, all respondents are shown an information block and then three additional blocks that are randomly selected from the remaining blocks. In 'multiwave' countries, respondents are shown four randomly selected blocks. Precise questions and the codebook for the data can be found in the Supplementary Information. In constructing the survey instrument, we drew on input from a wide set of domain experts. The survey consisted of questions related to COVID-19 information exposure and trust in information sources, knowledge about the virus, community norms, prevention behaviours, beliefs about efficacy of measures, vaccine acceptance, risk perceptions and locus of control in addition to demographics. The survey data include weights that use the rich information Facebook has about its users to reduce bias from non-response and differential Facebook use among different subpopulations. This resource article presents the survey dataset and some example use cases of the data.

Results

We now provide some basic results about the survey sampling and weighting as well as assorted analyses using data from some of the modules of the survey, including vaccine acceptance over time, mismatch in COVID-19 perceptions and consumption and trust of various news sources. These are some examples of possible uses of the data. In the Discussion, we show some other examples from other papers using the same data and point to directions for future research using the data.

Characteristics of the sample. Figure 2 shows the sample size we obtained per country and the effective sample size (as measured by Supplementary Information equation (A.2)). Although, on average, we obtain 3,000 users per week, the effective sample size varies widely, Bangladesh being the lowest with an average of only

[1]McCombs School of Business, The University of Texas at Austin, Austin, TX, USA. [2]School of Communication and Information, Rutgers University, New Brunswick, NJ, USA. [3]Initiative on the Digital Economy, Massachusetts Institute of Technology, Cambridge, MA, USA. [4]Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA, USA. [5]Department of Industrial Engineering, University of Pittsburgh, Pittsburgh, PA, USA. [6]Center for Communications Programming, Johns Hopkins University, Baltimore, MD, USA. [7]Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA. [8]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK. [9]Global Outbreak Alert and Response Network, Geneva, Switzerland. [10]Institute for Data, Systems, and Society, Massachusetts Institute of Technology, Cambridge, MA, USA. [11]These authors contributed equally: Avinash Collis, Kiran Garimella, Alex Moehring, M. Amin Rahimian. ✉e-mail: sinan@mit.edu; eckles@mit.edu

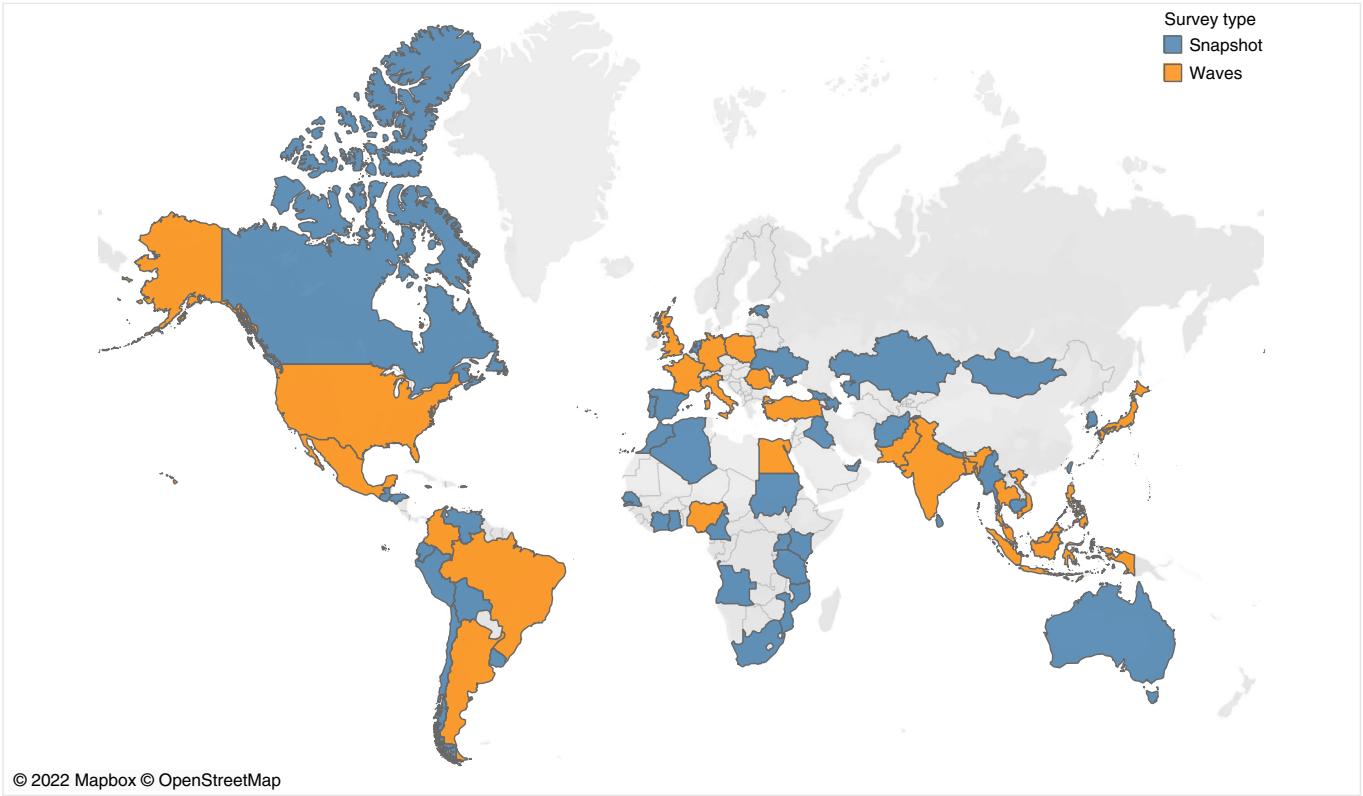

Fig. 1 | World map showing the countries represented in the survey. Twenty-three 'wave' countries were surveyed in 2-week waves from July 2020 until March 2021. Forty-four 'snapshot' countries were surveyed twice, once in July 2020 and another time in November 2020.

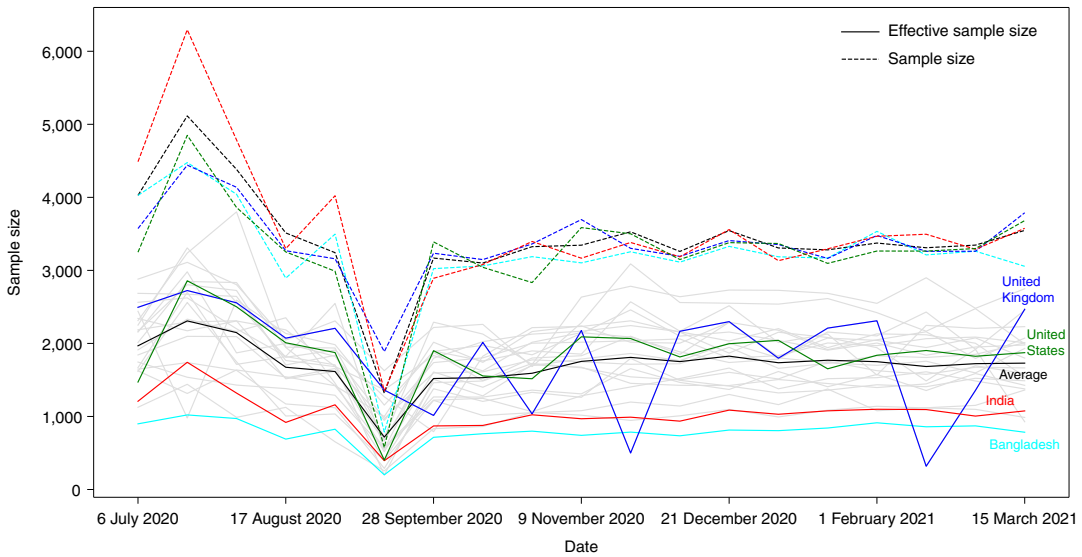

Fig. 2 | Sample sizes and effective sample sizes by country and survey wave. The survey consistently samples (except the wave starting on 14 September 2020) around 3,000 users every wave. However, the effective sample size varies more widely among the countries and within each country. The *x* axis indicates the start date of a 2-week period (a 'wave') of data collection. For clarity, four countries are highlighted; a full version of the plot for all 23 countries is shown in Supplementary Information Fig. A5.

791 users. Supplementary Information Tables A2 and A3 show the unweighted and weighted demographics of our sample, respectively. Supplementary Information Table A1 shows the two most popular languages used, by country.

Next, we plot the (inverse) conversion rate to the survey (how many users saw our survey prompt on their homepage) versus how many clicked and completed our survey. We can see from Fig. 3 that we needed, on average, 260 impressions for a single response. This is in line with the conversion for previous research using Facebook ads for surveys[7]. For most countries with good Facebook penetration (for example, in Europe), this number is around 50. For some countries (for example, Nigeria and India), the number was at least

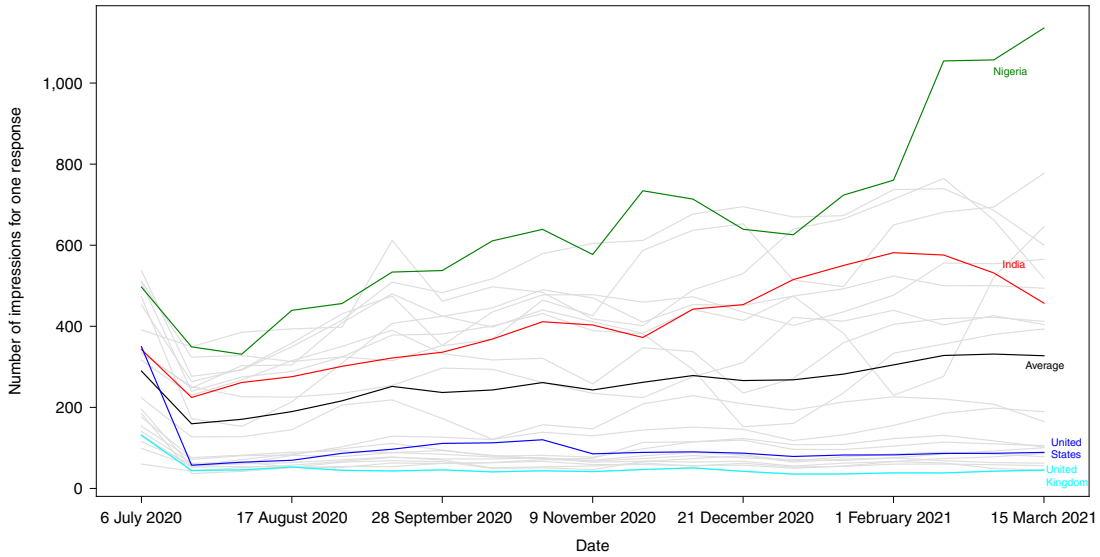

Fig. 3 | Number of impressions per response by country and survey wave. There is substantial heterogeneity in the conversion across countries. For clarity, four countries are highlighted; see Supplementary Information Fig. A6 for plots for all countries.

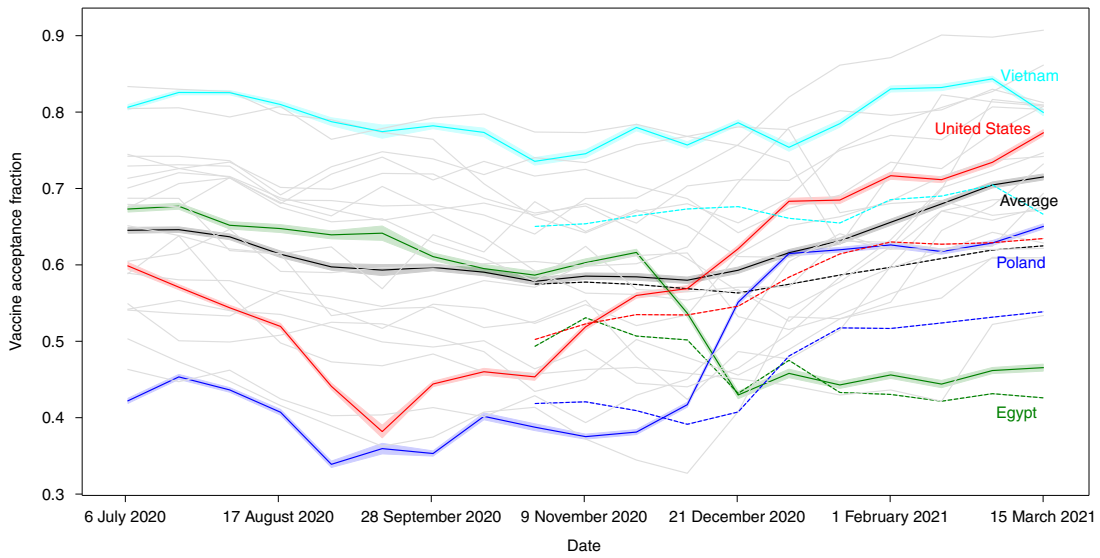

Fig. 4 | The fraction of respondents who say that they would take a vaccine or have taken the vaccine (vaccine acceptance fraction) over time. For clarity, only four countries are highlighted; see Supplementary Information Fig. A7 for individual country results. Bands are 95% confidence intervals.

an order of magnitude higher. This may reflect various differences, including perceived and actual costs of mobile data that would be used when completing the survey. Our survey weights are designed to reduce these biases in sampling.

Vaccine acceptance over time. We look at vaccine hesitancy and its trends over time. First, we computed the fraction of respondents who say that they would take a vaccine or have taken the vaccine (starting July 2020). Figure 4 shows the trends for the 23 wave countries over the duration of the survey (July 2020–March 2021). We observe a few clear trends. There is huge heterogeneity across countries, with Vietnam having a consistent vaccine acceptance of over 80% throughout the time period and countries like the United States and Poland experiencing an initial dip but improving in terms of acceptance later in the months before mass rollout of vaccines. Egypt, which would not see vaccines rolled out at scale for another 6 months, had a steady decline in vaccine acceptance during the same period[8]. On average, across the 23 wave countries, vaccine acceptance has varied in the range of 57% to 71% with slight improvements since late 2020. We notice these improvements across many countries where vaccines were being slowly rolled out, although making a causal connection between vaccine rollout and vaccine acceptance is beyond the scope of this paper.

Starting in wave 9 (end of October 2020), we also asked the following question about perceived vaccine norms: 'Out of 100 people in your community, how many do you think would take a COVID-19 vaccine if it were made available?'. The question helps us gauge perceptions of vaccine acceptance in the community. It is interesting to note that there is a significant difference between individual beliefs ('acceptance') and beliefs about others ('norms'). There is at least a 10% gap between them consistently. Respondents think that at least an additional 10% of the population would not take the vaccine.

Figure 5 shows the proportion of responses to the vaccine acceptance question for the four countries. The figure shows the importance of the 'Don't know' response or people who are yet undecided

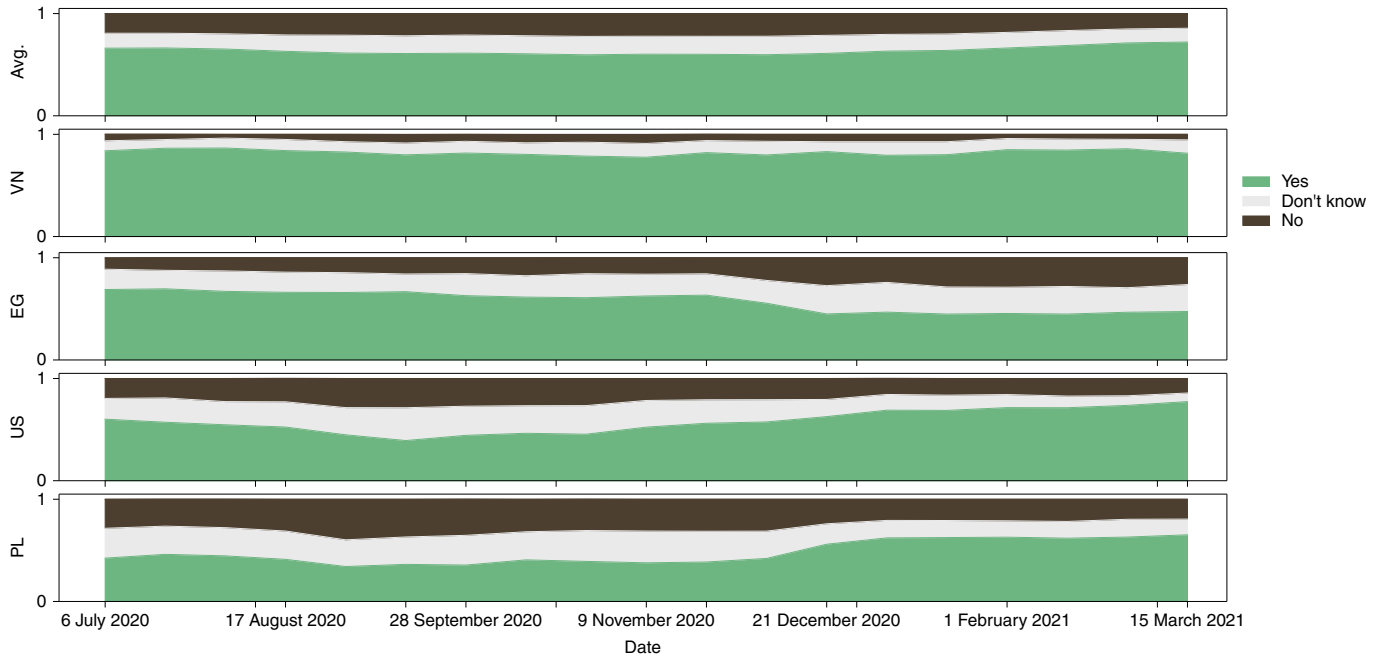

Fig. 5 | Distribution of responses for the vaccine acceptance question for four countries along with the average across our sample. Avg, average across all 23 wave countries; VN, Vietnam; EG, Egypt; US, United States; PL, Poland. We can observe the wide variance across countries about the role of vaccine-hesitant users (respondents saying 'No') and unsure users (respondents saying 'Don't know'). Supplementary Information Fig. A8 shows the result for all the countries in our sample.

on the vaccine. Consider the case of the United States, where the proportion of unsure users declined over time, while the proportion of users saying they would take a vaccine increased. Similarly, in Egypt, the proportion of users who oppose the vaccine as well as those who are unsure has increased in the last few months which is a good case study for policy intervention. Overall, on average, across the 23 countries in our dataset, vaccine acceptance varied between 59% and 72% between July 2020 to March 2021. The proportion of users who are not sure has ranged between 13% and 18% during that same period.

Next, we plotted the correlation between acceptance and norms for the 44 snapshot countries (for wave 9, in which the norms question was asked). We observe similar trends in Fig. 6. In all the 44 countries, respondents think others are much less likely to get the vaccine than they are themselves. We highlighted four countries to indicate the heterogeneity across countries. Note that self-reported intentions to vaccinate might differ from actual vaccine uptake[9].

Mismatch in COVID-19 perceptions. We asked two questions about the perception of seriousness of COVID-19 and perceptions among the community: community_action_importance—'How important is it for you to take actions to prevent the spread of COVID-19 in your community?' (possible answers—extremely important, very important, moderately important, slightly important, not important at all); and community_action_norms—'How important do other people in your community think it is to take actions to prevent the spread of COVID-19?'. If respondents themselves think taking action against COVID-19 to be extremely important, but think others do not take it seriously (or vice versa), they might adapt their behaviour to take steps that would not be necessary. Figure 7 shows the mismatch in beliefs for two countries: the United States and Japan. The figure shows a heat map of the mismatch. The plots are normalized by row (one's own beliefs) and each cell indicates the conditional probability of beliefs about others (columns) given one's own beliefs (rows sum to 1). We see that there

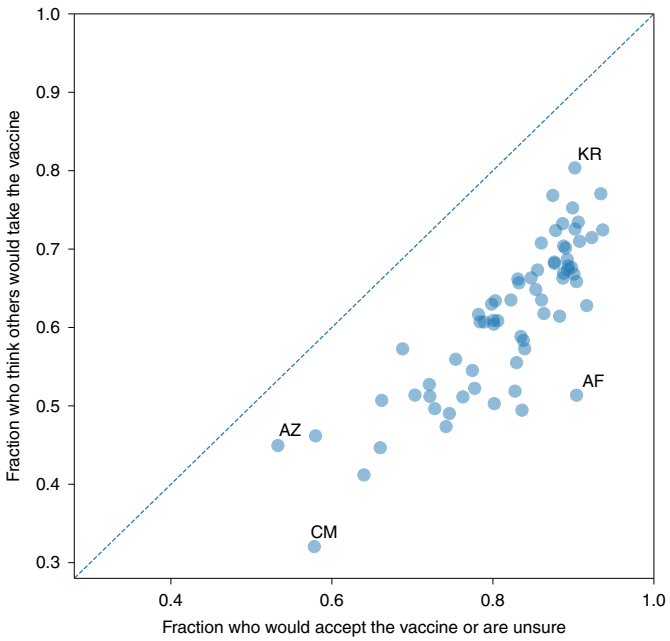

Fig. 6 | Vaccine acceptance versus vaccine norms for all 67 countries in our sample. Each point is the weighted fraction for that country. Across all countries, respondents think that far fewer people in their country say they would take the vaccine. AZ, Azerbaijan; CM, Cameroon; AF, Afghanistan; KR, South Korea.

is a clear difference in the distributions across the United States and Japan, with most people in Japan having a congruent view, compared to the wide range of disagreement in the United States. The two countries were chosen to show an example of how divergent the beliefs about others could be in different cultures.

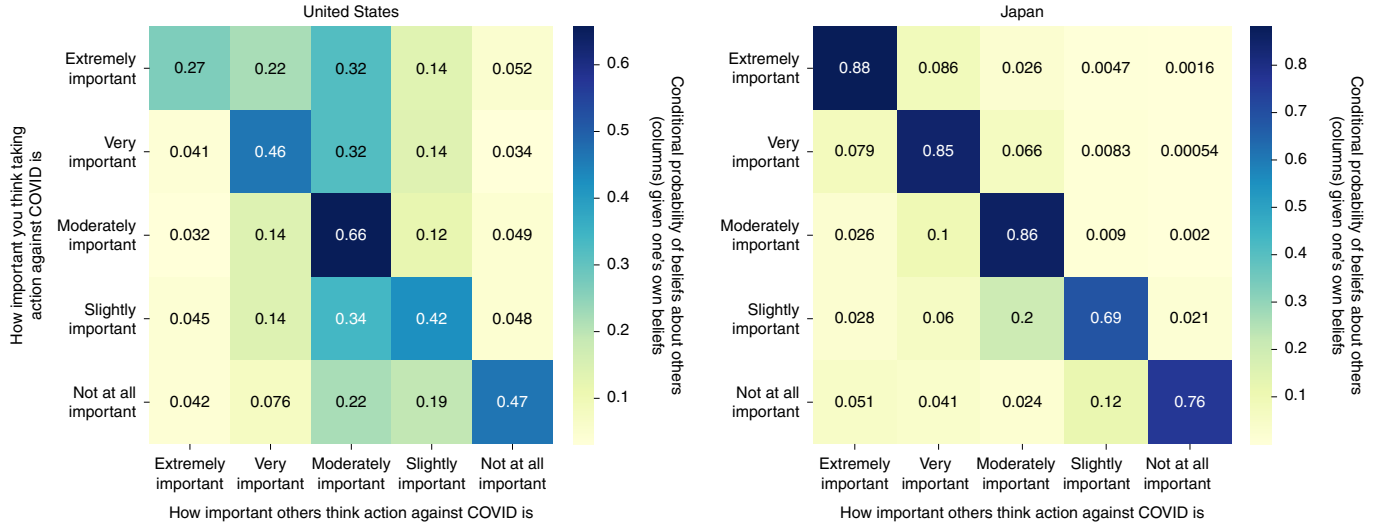

Fig. 7 | Beliefs about the importance of taking action against COVID versus beliefs about the beliefs of others in the United States and Japan. The *y* axis indicates how serious the respondent considers COVID-19 to be and the *x* axis indicates how serious they think others think COVID-19 is. The number in each cell represents the conditional probability of beliefs about others (columns) given one's own beliefs (rows, which sum to 1). For instance, for the United States (plot on the left), the top left cell value of 0.27 indicates that 27% of the users who think taking action against COVID-19 is 'extremely important' also think that others think it is 'extremely important'.

News sources and medium: consumption versus trust. Finally, we asked for the sources/ mediums users consumed COVID-19 related information from and their trust in these sources (pages 20 and 21 in the Supplementary Information list these survey questions). Figure 8 shows the trends for consumption and trust for five sources: online news, radio, television, local health workers and politicians. In a pandemic, it is important to have widely trusted sources provide information that is widely consumed[10–12]. However, as we can see from the figure, most sources do not satisfy this criterion. Some interesting trends emerge: politicians are the least trusted, and in most countries the least used, source of information. Television has high consumption but trust in television varies widely among the countries in our sample. Local health workers are typically well trusted but they are not a source of information for most countries.

Discussion

The paper describes a global, longitudinal survey on COVID-19 behaviours, beliefs and norms. We present three examples of potential use cases for the dataset: (1) vaccine acceptance and norms, (2) mismatch between own beliefs and beliefs about others and (3) trust in versus consumption of various news sources. Some of the trends observed here, particularly at a global scale and including countries in the global south, are valuable for understanding behavioural and social drivers of vaccination[13] and would not have been made available to the research community otherwise. Identifying what people think and feel and the social processes, such as norms[14], that influence their thinking will help researchers identify motivations behind critical health behaviours. Such a strategy is, for instance, extensively used by WHO for measuring behavioural and social drivers of vaccine hesitancy[15]. Overall, this paper provides a valuable resource which should serve as a foundation for future research and give rise to new questions in understanding the COVID-19 pandemic and developing policy solutions around it. For instance, our findings on heterogeneity in vaccine trends across countries (Fig. 4) or the mismatch in perceptions across countries (Fig. 7) are new and may not be explained by existing literature. Combining our data with historic and cultural trends could help identify new insights on the role of country-specific variables in explaining the results[16,17]. Some of the temporal variations in vaccine acceptance (for example, in countries such as the United States, Poland and Egypt, highlighted in Fig. 4) remain unexplained and open venues for future research into factors behind vaccine acceptance trends.

Our survey data can directly inform policies on the national and global stage. For example, others in their study of political messaging and attitudes towards vaccination in Latin America[18] use our surveys to assess the relationship between vaccine acceptance, political vaccination campaigns and political trust. Another study of our survey responses for South Asian countries identified gender, age, knowing someone who tested positive for COVID-19 and perceived effectiveness of mask wearing as significant determinants of COVID-19 vaccine hesitancy, arguing for targeted vaccine education and communication campaigns[19]. Others[20] analysed responses among ten snapshot countries in sub-Saharan Africa in the two survey rounds that happened in July and November 2020 (Fig. 1). They use the 'yes' and 'no' answers to the survey question about handwashing in the past week as their primary outcome. Using a multivariate logistic regression, they identify the main determinants of handwashing that are classified sociodemographic (age, gender, education and rural or urban residence) and ideational (perceived personal health, beliefs about handwashing, knowing someone diagnosed with COVID-19 and perceived norms), adjusting for country-level fixed effects. The authors document clear regional and country-level variations in handwashing, pointing to settings with the greatest opportunity for improvement. Similarly, the significant country-level heterogeneity of our survey measures and, in particular, the vaccine trends, have served as motivation or explanatory factors in other research studies that target local populations; for example, in Spain[21] or Australia[22].

Several other studies have used COVID-19 beliefs, behaviours and norms survey data to analyse risk perception, attitudes towards mask wearing and other preventive behaviours, as well as trust in information sources across communities worldwide. A previous study[23] uses the survey data to identify significant predictors of risk perception in older adults and its association with their preventive behaviours and medical avoidance. They find accurate knowledge to be a crucial factor in disentangling this association. Joining the survey data with COVID-19 cases and death counts worldwide, another study[24] shows that mask wearing and attitudes towards

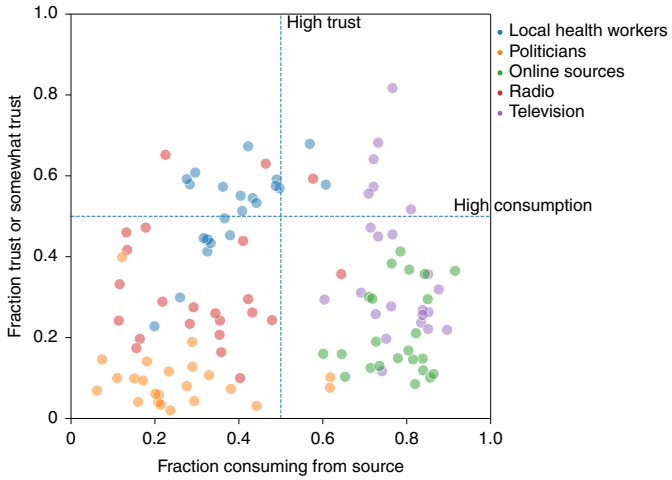

Fig. 8 | Trust versus consumption of news sources and media. Each point is a country and source pair for all wave countries, averaging over all waves of the survey.

masks are associated with fewer cases and deaths across different countries, controlling for socioeconomic factors such as population density, human development and mobility. Another analysis[17] of the survey data reveals that mask usage is higher in countries with more collectivistic (versus individualistic) cultures after controlling for a host of variables such as COVID-19 severity, government policy, population density, GDP per capita and demographics. Others[25] analyse our survey responses to construct various measures of vaccine intention, perceived invincibility and prosocial concerns at the individual level and study their relationships, controlling for perceived personal health and demographic attributes measured in the survey, as well as estimates of country-level cultural collectivism from other studies. They show that perceived invincibility has an overall negative effect on both prosocial concerns and vaccine intentions. These effects are particularly pronounced in counties with low cultural collectivism and shown to be robust across age cohort and gender. This ability to investigate individual health-related behaviours by controlling for country-level variables, such as cultural collectivism, shows the unique contribution of the present resource to the research community. Such investigations would not have been possible without the global COVID-19 beliefs, behaviours and norms survey data.

Yet another study[26] uses the randomized order of the survey questions to show that highlighting accurate information about vaccine norms increases vaccine acceptance. Several layers of randomization throughout the survey provide a ripe ground to explore priming, anchoring and information treatment effects on different demographics in a representative global sample (for example, respondents are randomized to see questions about their risk perception and perceived control over health outcomes which affects their answers to follow-up questions about their adherence to preventive measures in ways that can inform public health communication). The longitudinal data are collected over a period of global pandemic emergency that coincided with high-profile events, providing natural experimental opportunities on national and international scales (for example, the US presidential election, epidemic peaks and emergency use approvals of vaccines in different countries). In addition to in-depth demographic, psychographic and sociometric measurements of health-related behaviours as well as media and news consumption (some of which were show-cased in the Results), the survey resource also has questions about work and travel (full survey instrument given in Supplementary Information D). We expect the confluence of these factors will open new areas of enquiry in public health, communication and economic policy and we are optimistic that future researchers will leverage these large-scale, rich survey data on beliefs, behaviours and norms during the COVID-19 pandemic in innovative ways.

Methods

The survey's purpose was to guide policy and research around individual responses to COVID-19 beyond symptoms and the most closely associated behaviours. The Committee on the Use of Humans as Experimental Subjects at the Massachusetts Institute of Technology (MIT) approved the survey as exempt (project no. E-2294) and informed consent was obtained from all participants. The survey ran from July 2020 until March 2021. It was translated into 51 languages and fielded in 67 countries, yielding over 2 million responses. The full survey instrument is provided in Supplementary Information D. The survey data dictionary is provided in Supplementary Information B and the log of changes to the survey over the course of its duration are provided in Supplementary Information C.

Survey instrument design. There were multiple goals for this survey and associated topics for each goal that formed individual modules of questions. The users of this survey include academic researchers, governments and non-governmental organizations. As the pandemic was occurring during the lifetime of this survey, one of our main goals was to provide ongoing tracking of key measures of knowledge about COVID-19 and how to prevent its spread, which can inform targeting and evaluation of public health campaigns. For researchers, the goal behind the survey was to provide them with a rich dataset spanning multiple countries to conduct more in-depth research. We gave examples of research papers applying this dataset in the Discussion.

More specifically, we wanted to provide data to help achieve the following goals:

- Understand which preventive behaviours are most/least understood and practiced by region/country and how this changes over time
- Identify countries/regions with low knowledge of given preventive behaviours and understand how and why this differs from adjacent countries/regions
- Identify differences in self-reported preventive behaviours associated with differences in psychosocial behavioural determinants
- Identify countries/regions with the biggest gap between knowledge and practices and understand how and why this differs from adjacent countries/regions
- Understand how COVID-19 related policies impact knowledge, attitudes and behaviours by geography

These survey goals led us to build different modules within the survey including (see Supplementary Information D for the full survey instrument):

- Basic demographics and localization
- Knowledge about COVID
- Knowledge about preventive measures
- Current behaviours for prevention
- Beliefs about norms
- Hypotheticals about relaxed restrictions
- Exposure to various sources of information

Sampling and weighting. The survey was fielded in two different ways. First, in countries with a sufficient pool of Facebook users to sample, we fielded a multiwave survey that ran continuously in multiple 2-week waves from July 2020 until March 2021. In each wave, Facebook aimed to deliver 3,000 respondents to our survey. In countries with a more limited survey pool, we fielded a snapshot survey where Facebook aimed to deliver 3,000 respondents over a 2-week period; this was done twice, first in July 2020 and then in November 2020. The list of countries is selected on the basis of survey viability (which is determined by the population of Facebook users in that country), regional representation and feedback from survey partners at the World Health Organization and the Global Outbreak Alert and Response Network. See Fig. 1 for a map showing the countries.

The Facebook team uses non-response modelling and poststratification techniques from survey statistics to design the following components[27,28]:

(1) Sampling—deciding who to present with the invitation to participate in the survey
(2) Weighting—providing a weight per user so that respondents better represent the target population as a whole

Using the total survey error framework, the goals of the sampling and weighting steps are to minimize the representation error due to the coverage, sampling variability and non-response biases[28]. Each sampled user was presented with an online consent form to take part in the MIT survey (the survey instrument in the Supplementary Information gives the text of the consent form). The weights are constructed on the basis of behavioural covariates such as frequency and duration of user activity as well as self-reported demographics. These attributes are used by Facebook in their routine surveys and no new data were collected by Facebook for this purpose. The survey weights do not come from individual survey responses, which Facebook does not collect or have access to but rather come

from internal Facebook data. The exact nature of the client data and algorithms are proprietary to Facebook. The data are protected by Facebook and its original collection is not subject to the consent form but is covered by Facebook's terms of service. MIT has no access to these data per the terms of the data use agreement between Facebook and MIT. Survey respondents consented to MIT receiving their survey weights from Facebook.

The MIT team supplied binary survey completion flags (binary indicators of whether or not each respondent has completed the survey) along with a respondent identifier (a random number associated with each survey respondent) back to the Facebook team. No other data about individual respondents were sent by MIT to Facebook. We provide the completion flags for the following two analytical samples:

(1) Respondents who have completed the basic knowledge and demographics parts of the survey. This part consists of a briefing followed by questions about information exposure, availability of treatments and vaccines and contact with healthcare workers, as well as gender, age, education, overall health, country and, in the case of the United States and India, state as well. We call this the 'demographic completion type'.

(2) Respondents who have reached the end of the entire survey, viewing (and typically answering) additional questions about information sources; information needs; their knowledge about high-risk populations, methods of transmission and disease symptoms; norms and beliefs about distancing, mask wearing and other preventive measures; risk perception and locus of control; work, travel and intentions to visit various locations, followed by a debrief. We call this the 'full survey completion type', although note that there can still be missing data due to non-response to individual questions and random assignment to different survey blocks.

Subsequently, the Facebook team computed and returned sets of survey weights to the MIT team, one set for each analytical sample. No other data about respondents were sent by Facebook to MIT besides a respondent identifier (a random number associated with each survey respondent), their language preference, these survey weights and an indicator of whether these survey weights were clipped (Supplementary Information A). The weights are meant to be used in Háajek estimators (normalized importance sampling estimators) for measuring population means. Specifically, let Y_i be an outcome variable of interest measured for the respondent i whose weight is w_i. The Háajek estimator, $\hat{Y}$, for the population mean of the outcome, $\bar{Y}$, is given by:

$$\hat{Y} = \frac{\sum_{i=1}^{n} w_i Y_i}{\sum_{i=1}^{n} w_i}, \quad (1)$$

This is the default in most statistical software for computing a weighted mean. Subsequently, if interested in population totals, analysts should use $N\hat{Y}$ as an estimator of the total outcome level where N is the population size. That is, analysts should not use the weights in an unnormalized way, as in a Horvitz–Thompson estimator (an unnormalized importance sampling estimator), as, while the weights are approximately on the level of each country's adult population, the clipping and other adjustments to the weights make them unsuitable for direct estimation of total outcome levels without normalization. More generally, users can use these weights in other related estimators that appropriately normalize the weights[29]. Survey weights are critical to maintaining statistical representativeness and especially important for large samples[9]. Supplementary Information A includes a detailed description of the survey weights design and various consistency checks for representativeness of the weighted survey sample.

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

Data availability

Aggregate data can be found at https://covidsurvey.mit.edu/ and researchers can request access to respondent-level responses (microdata) at https://dataforgood.facebook.com/dfg/docs/preventive-health-survey-request-for-data-access.

Code availability

Analysis code to reproduce figures in the manuscript are available at https://github.com/gvrkiran/Global-Survey-on-COVID-19-Beliefs-Behaviors-and-Norms.

Received: 15 September 2021; Accepted: 1 April 2022;
Published online: 23 May 2022

References

1. Ferguson, N. Capturing human behaviour. *Nature* **446**, 733–733 (2007).
2. West, R., Michie, S., Rubin, G. J. & Amlôt, R. Applying principles of behaviour change to reduce SARS-CoV-2 transmission. *Nat. Hum. Behav.* **4**, 451–459 (2020).
3. Van Bavel, J. J. et al. Using social and behavioural science to support COVID-19 pandemic response. *Nat. Hum. Behav.* **4**, 460–471 (2020).
4. Allcott, H. et al. Polarization and public health: partisan differences in social distancing during the coronavirus pandemic. *J. Public Econ.* **191**, 104254 (2020).
5. Bursztyn, L., Rao, A., Roth, C. P. & Yanagizawa-Drott, D. H. *Misinformation During a Pandemic* No. w27417 (National Bureau of Economic Research, 2020).
6. Bailey, M. et al. *Social Networks Shape Beliefs and Behavior: Evidence From Social Distancing During the Covid-19 Pandemic* No. w28234 (National Bureau of Economic Research, 2020).
7. Allcott, H., Gentzkow, M. & Song, L. *Digital Addiction* No. w28936 (National Bureau of Economic Research, 2021).
8. Mathieu, E. et al. A global database of COVID-19 vaccinations. *Nat. Hum. Behav.* **5**, 947–953 (2021).
9. Bradley, V. C. et al. Unrepresentative big surveys significantly overestimated US vaccine uptake. *Nature* **600**, 695–700 (2021).
10. Vaughan, E. & Tinker, T. Effective health risk communication about pandemic influenza for vulnerable populations. *Am. J. Public Health* **99**, S324–S332 (2009).
11. Leonard, M. B., Pursley, D. M., Robinson, L. A., Abman, S. H. & Davis, J. M. The importance of trustworthiness: lessons from the COVID-19 pandemic. *Ped. Res.***91**, 482–485 (2022).
12. Merkley, E. & Loewen, P. J. Anti-intellectualism and the mass public's response to the COVID-19 pandemic. *Nat. Hum. Behav.* **5**, 706–715 (2021).
13. Brewer, N. T., Chapman, G. B., Rothman, A. J., Leask, J. & Kempe, A. Increasing vaccination: putting psychological science into action. *Psychol. Sci. Public Interest* **18**, 149–207 (2017).
14. Hensel, L. et al. Global behaviors, perceptions, and the emergence of social norms at the onset of the COVID-19 pandemic. *J. Econ. Behav. Organ.* **193**, 473–496 (2022).
15. *Development of Tools to Measure Behavioural and Social drivers (BeSD) of Vaccination* (World Health Organization, 2020).
16. Chen, C., Frey, C. B. & Presidente, G. Culture and contagion: individualism and compliance with COVID-19 policy. *J. Econ. Behav. Organ.* **190**, 191–200 (2021).
17. Lu, J. G., Jin, P. & English, A. S. Collectivism predicts mask use during COVID-19. *Proc. Natl Acad. Sci. USA* **118**, e2021793118 (2021).
18. Piazza, K. S. & Schwier, A. Ready, set, vaccine: the path to COVID-19 recovery in Latin America. *Rev. Latinoam. Opin. Pública* **10**, 179–190 (2021).
19. Hartigan-Go, K. Y., Mendoza, R. U., Ong, M. M. A. & Yap, J. K. COVID-19 vaccine hesitancy in ASEAN: insights from a multi-wave survey database from July 2020 to March 2021. *Acta Med. Philipp.* https://doi.org/10.47895/amp.vi0.3679 (2020).
20. Olapeju, B. et al. Trends in handwashing behaviours for COVID-19 prevention: longitudinal evidence from online surveys in 10 sub-Saharan African countries. *PLOS Glob. Public Health* **1**, e0000049 (2021).
21. Matute, J., Palau-Saumell, R., Meyer, J., Derqui, B. & Jiménez-Asenjo, N. Are you getting it? Integrating theories to explain intentions to get vaccinated against COVID-19 in Spain. *J. Risk Res.* https://doi.org/10.1080/13669877.2021.1958044 (2021).
22. Carter, J., Rutherford, S. & Borkoles, E. COVID-19 vaccine uptake among younger women in rural Australia. *Vaccines* **10**, 26 (2022).
23. Lu, P., Kong, D. & Shelley, M. Risk perception, preventive behavior, and medical care avoidance among American older adults during the COVID-19 pandemic. *J. Aging Health* https://doi.org/10.1177/08982643211002084 (2021).
24. Adjodah, D. et al. Association between COVID-19 outcomes and mask mandates, adherence, and attitudes. *PLoS ONE* **16**, e0252315 (2021).
25. Leonhardt, J. M., Ridinger, G., Rong, Y. & Talaei-Khoe, A. Invincibility threatens vaccination intentions during a pandemic. *PLoS ONE* **16**, e0258432 (2021).
26. Moehring, A. et al. *Surfacing Norms to Increase Vaccine Acceptance* (SSRN, 2021); https://doi.org/10.2139/ssrn.3782082
27. Barkay, N. et al. Weights and methodology brief for the COVID-19 symptom survey by University of Maryland and Carnegie Mellon University, in partnership with Facebook. Preprint at https://arxiv.org/abs/2009.14675v2 (2020).
28. Groves, R. M. & Lyberg, L. Total survey error: past, present, and future. *Public Opin. Q.* **74**, 849–879 (2010).
29. Särndal, C. E., Swensson, B. & Wretman, J. *Model Assisted Survey Sampling* (Springer Science & Business Media, 2003).

Acknowledgements

This work was funded in part by a grant from Facebook to the MIT Initiative on the Digital Economy. The funder had no role in the decision to publish or preparation of the manuscript. This survey was a collaborative effort involving contributions from individuals at multiple institutions, especially MIT, Johns Hopkins University, the Global Outbreak Alert and Response Network (GOARN), the World Health Organization and Facebook. At Facebook, key contributors were E. Kim, K. Mulcahy, P. Raja, S. Sasser, C. Velasco and T. Wynter. At MIT, we thank A. Ruel and S. Watts for managing other institutions gaining access to the microdata. We thank the millions of respondents to this survey worldwide.

Author contributions

A.C., K.G., A.M. and M.A.R. led the project with supervision from D.E. and S.A. Authors A.C., K.G., A.M., M.A.R., D.E., S.A., S.B., N.H.G., D.S. and J.S. contributed to designing the survey. A.C. and A.M. led the implementation of the survey design. A.M. managed data processing and exchange. K.G. led the analysis with inputs from A.C., A.M., M.A.R. and D.E. The paper was written by A.C., K.G., A.M., M.A.R. and D.E.

Competing interests

Facebook has sponsored a conference organized by S.A. and D.E. D.E. is a consultant to Twitter. A.C. and D.E. have received funding for other research from Facebook. M.A.R. serves on the advisory committee of a vaccine confidence fund created by Facebook and Merck. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41562-022-01347-1.

Correspondence and requests for materials should be addressed to Sinan Aral or Dean Eckles.

Peer review information *Nature Human Behaviour* thanks Jeffrey Lazarus, Anne Templeton and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

nature portfolio

Corresponding author(s): Avinash Collis

Last updated by author(s): Mar 21, 2022

Reporting Summary

Nature Portfolio 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 Portfolio 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	The survey instrument is provided in the appendix. The Qualtrics survey preview link is here: https://mit.co1.qualtrics.com/jfe/preview/SV_9ZxxxOQGLnp29yR?Q_CHL=preview&Q_SurveyVersionID=current&fbclid=IwAR1gG1ZJNtBVhadC8maaZftWgYeKo6xLN_lfaErJqDIAgA14XjHiqd7dmCE
Data analysis	All analysis done in python was done using python 3.8 with the following packages numpy (1.21.2), pandas (1.3.0), patsy (0.5.1), scipy (1.6.2), stargazer (0.0.5), statsmodels (0.12.2). The multilevel modeling analysis was run using R version 3.5.1 and additional auxiliary analysis was run using R 4.0.21. Analysis code to reproduce figures in the manuscript are available at: https://github.com/gvrkiran/Global-Survey-on-COVID-19-Beliefs-Behaviors-and-Norms

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 Portfolio 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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Aggregate data can be found at this link: https://covidsurvey.mit.edu/, and researchers can request access to respondent-level responses (microdata) by requesting access at this link: https://dataforgood.fb.com/docs/preventative-health-survey-request-for-data-access/

All academic and non-profit researchers can request access to the microdata.

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

Behavioural & social sciences study design

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

Study description	This resource describes quantitative data from a global survey on COVID-19 beliefs, behaviors and norms, conducted in 67 countries yielding 2.0 million responses.
Research sample	The sample was recruited in collaboration with Facebook and consists of users of Facebook. Users were recruited through a message on their news-feed. The survey was translated to 51 languages and fielded in 67 countries yielding over 2.0 million responses. This is one of the first international social science surveys of this magnitude. Our collaboration with Facebook has enabled both this wide distribution, but also state-of-the-art, privacy-preserving adjustment for non-response leveraging demographic and behavioral variables observed by Facebook. No new data was collected by Facebook for the purpose of this survey. The variables used by Facebook to construct weights are already used in their routine surveys. The weights are designed to make the sample representative of populations of interest.
Sampling strategy	The survey was fielded in two different ways. First, in countries with a sufficient pool of users to sample, we fielded a multi-wave survey that ran continuously in multiple two-week waves from July 2020 until March, 2021. In each wave, Facebook aimed to deliver 3,000 respondents to our survey. In countries with a more limited survey pool, we fielded a snapshot survey where Facebook aimed to deliver 3,000 respondents over a two week period; this was done twice, first in July, 2020 and then in November, 2020.
Data collection	Survey was conducted on Qualtrics. The Qualtrics survey preview link is here: https://mit.co1.qualtrics.com/jfe/preview/SV_9ZxxxOQGLnp29yR?Q_CHL=preview&Q_SurveyVersionID=current&fbclid=IwAR1gG1ZJNtBVhadC8maaZftWgYeKo6xLN_IfaErJqDIAgA14XjHiqd7dmCE Users were blinded to their assigned experimental condition.
Timing	Wave countries: every two weeks from July 2020 until March 2021 Snapshot countries: Twice, in July 2020 and November 2020.
Data exclusions	No data were excluded from the analyses.
Non-participation	On average, we needed around 200 impressions for a single response. This is in line with the conversion for previous research using Facebook ads for surveys (Allcott et al., 2021). The survey data includes weights that use the rich information Facebook has about its users to reduce bias from non-response and differential Facebook use among different subpopulations.
Randomization	For the norms intervention described in Moehring et al. (2021), we provided the treatment at random times (either before or after the outcome was measured) and the treatment contained information about a randomly chosen preventative behavior. All subjects who were eligible for the information eventually saw the information if they completed the survey. Reference: Moehring, A., Collis, A., Garimella, K., Rahimian, M. A., Aral, S., & Eckles, D. (2021). Surfacing norms to increase vaccine acceptance. Available at SSRN 3782082.

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	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	The distribution of demographic characteristics are displayed in the manuscript. The sample was recruited to be representative of the adult population in a country after incorporating survey weights.
Recruitment	Subjects were recruited by Facebook and sent to our platform. In the manuscript we demonstrate substantial sampling and non-response bias that are in large part corrected when survey weights are incorporated. Randomization happens after recruitment, so these biases do not threaten internal validity of the study. Moreover, we show that results are robust to whether or not we adjust using the survey weights.
Ethics oversight	The MIT Committee on the Use of Humans as Experimental Subjects approved both the original survey (protocol E-2294) and the randomized experiment (protocol E-2674) as exempt studies.

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