

Who and why: a study of British distrust in the government and media during the COVID-19 pandemic

Maria Balaet (✉ m.balaet17@imperial.ac.uk)

Imperial College London

Danielle Kurtin

Imperial College London

Dragos Gruia

Imperial College London

Annalaura Lerede

Imperial College London

Darije Custovic

Imperial College London

William Trender

Imperial College London <https://orcid.org/0000-0003-3947-5532>

Amy Jolly

University College London

Peter Hellyer

King's College London

Adam Hampshire

Imperial College London

Article

Keywords:

Posted Date: April 18th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1564442/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

The COVID-19 pandemic represents a unique context for studying the spread of conspiratorial beliefs within the general population and their role in mediating compliance with government guidance. Here, we apply multivariate and machine learning methods to analyse data from tens of thousands of members of the British public at 6-monthly timepoints during the COVID-19 pandemic. We report that distrust and conspiratorial beliefs significantly predict non-compliant behaviours and covary with sociodemographic variables, being most prevalent for disadvantaged and minority groups. Free text analyses reveal that perceptions of corruption, cronyism, disputing lockdown motivations and the potential of COVID-19 being a lab-leaked bioweapon were common topics motivating non-compliance, with the prevalence and distribution of such beliefs evolving as the pandemic progressed. We propose that survey data could be analysed this way to identify current topics of distrust and map them to demographic variables, enabling the most relevant arguments to be tailored for each individual.

Introduction

The COVID-19 pandemic has been characterised by a need for cooperation between citizens and bodies of authority, in order to contain the virus and minimise health effects on the population. Achieving high compliance with the measures that governments put in place to mitigate infection and hospitalisation rates has posed a non-trivial challenge. For example, the U.K. rapidly achieved high rates of vaccination uptake, yet a substantial minority still remain unvaccinated and at risk as COVID-19 infections are still occurring. Also, other guidelines such as the mandate for wearing masks, social distancing in public spaces or using the NHS app were adhered to by many but by no means all. This raises the question as to why people would not voluntarily follow government guidance and safety measures. What beliefs motivate non-compliance and who is it that holds them?

A number of factors have been proposed to contribute to this lack of trust, including mixed messages from scientists, politicians and media regarding the optimal and appropriate level of control measures (1), double standards in adherence to rules of government vs. general public, and a perception of corruption in the awarding of lucrative contracts. In parallel, social media platforms have acted as hubs for misinformation related to the pandemic, fueling public unrest (2) and a lack of compliance (3). Of particular relevance, diverse conspiracy theories have emerged questioning the motives underpinning governmental response to the virus, the virus' origins, or even its existence (4).

Indeed, previous studies paint a bleak and often contradictory picture of distrust; for example, a survey of a representative sample of the UK's population in April 2020 showed that 62% of respondents considered the government's reaction to COVID-19 to be too slow, 31% did not trust the government to control the spread of COVID-19, and 24% did not believe the government had been truthful about COVID-19 (5). Low trust in government has been associated with higher levels of conspiratorial beliefs, unwillingness to engage in health protective behaviours, and reluctance to vaccinate against the virus (6–9).

Understanding peoples' beliefs and how they covary in the population with sociodemographic, lifestyle and health related factors could guide better targeting of attempts to persuade people towards improved compliance. To date, there have been limited attempts to link different aspects of compliance to select aspects of personality (10), political orientation (11), and demographics (12–14); but there have been none looking at the influence of lifestyle factors or neuropsychiatric status, nor that determine which among the many such variables has the most predictive value, or is associated with what specific belief.

What the most common classes of belief have been that motivate non-compliance is a critical puzzle piece. A major limitation of almost all previously published studies on this topic is that they rely on social media to quantify the range of opinions about the pandemic and its handling by bodies of authority, which may fuel misinformation (2) but also has variable engagement across the population, or traditional survey methods (6, 15–17), in which responses are given by selecting from pre-specified answers to pre-formulated questionnaire scales; for example, reflecting how strongly a respondent agrees with a particular statement. While fixed-answer approaches are efficient in some contexts – e.g., when determining rates of compliance to specific measures – they pose major limitations when seeking to ascertain the beliefs that motivate individuals to support or reject a given measure because they are constrained by the insights of the surveyor.

Identifying the most common beliefs that underlie distrust can be achieved by analyzing free text responses to questions (18), but this requires surveying large samples of the general population. High numbers of observations also are required to determine how different aspects of compliance or belief covary with each other within the general population, and how they relate to multiple potentially relevant levels of psychological, sociodemographic and lifestyle factors. To add further complication, if (as is likely) it is marginalised individuals who are most likely to hold such beliefs, then this presents a challenge because they tend to be the least likely to respond to standard surveying methods. While there have been studies using data from social media platforms to map the public opinion about pandemic (19, 20), they are marked by limitations such as poorly characterised subjects demographics due to social media making it easy to hide or change aspects of one's identity, spam messages leading to an uneven representation of beliefs, and an inability to target specific pandemic related questions in a controlled manner.

Here, we use a novel approach to map the common beliefs that underlie distrust and non-compliance with COVID-19 safety measures. We apply multivariate and machine learning methods to large-scale citizen science data, capturing a large and inclusive cross-section of the UK general public at three distinct stages of the COVID-19 pandemic. First, we quantify the distribution of answers in our sample using questions probed by three general themes: people's compliance with pandemic-related restrictions; their sources of news and information; and their beliefs about the nature and origin of the virus and the government and media's responses to its spread. We create a composite compliance score and map the probability of non-compliance across the population and use logistic binomial and multinomial regression to identify the profiles of people who are most likely to hold divergent beliefs or to endorse non-compliance. Critically, we supplemented fixed answer questions with free-text responses that allow

those who express distrust or lack of compliance to explain their reasons for doing so. We apply one of the most established topic modelling methods, *Latent Dirichlet Allocation* (LDA), to discern in a data-driven manner the reasons that motivate distrust, and by association, non-compliance. Finally, we map topic prevalence across relevant population variables and examine changes in their prevalence across three timepoints within the pandemic between December 2020 and January 2022. We conclude by highlighting topics that the government and media might address with the general public in order to improve trust, and consequently compliance in the case of crisis situations (such as but not limited to the COVID-19 pandemic) where cooperation between citizens and bodies of authority is needed.

Results

Participants were recruited from those who signed up for the Great British Intelligence Test (21) in response to emails sent in December 2020 and June 2021 and January 2022. We analysed data from 20,922 individuals in December 2020, from 12,796 individuals in June 2021 and from 14,090 individuals in January 2022 using the Cognitron online assessment platform (22–24). 2,797 people completed all three timepoints. Demographics of participants at all timepoints are presented in Supplementary Materials.

1. Characterising behaviours and trust: an analysis of fixed response questionnaire data

We assessed the distribution of responses to three types of pre-formulated questions. The first type probed whether or not respondents were distrustful of the government's approach and the media's response to the pandemic, vaccines, and the mainstream narrative of the origins of COVID-19 (Figure 1A-E). The second type probed whether respondents had complied with suggested measures to reduce the transmission of COVID-19 (Figure 1F-I). The third type probed respondents' primary sources of information regarding the pandemic (Figure 1J).

1.1 The distribution of trust and compliance and information sources

As expected, although the majority of the population endorsed trust in the government, media, vaccines, and the standard narrative for the origins of COVID-19 (Figure 1A-E) there were substantial minorities who endorsed distrust and a lack of compliance. For example, in December 2020, 22.7% of respondents reported assuming that there were ulterior motives behind the government's response to the pandemic (Figure 1A), 17% that the media was covering up information (Figure 1B), 19.6% did not consider COVID-19 to be a natural phenomenon, (Figure 1C) and 9.7% felt that the government imposed restrictions were not justified (Figure 1E). Over the course of one year these responses showed small but statistically significant changes. Most notably, more people assumed the government had ulterior motives behind their response to the pandemic (29.9%) and that the COVID-19 pandemic had unnatural origins (27.9%).

At the same time, fewer people (13.4%) considered that more needed to be done regarding restrictions on social behaviour in January 2022 compared to December 2020 (29.3%).

A similar pattern of results was evident for the compliance questions, with the majority endorsing compliance with safety measures but substantial minorities reporting non compliance (Figure 1F-I). For example, in December 2020, 37.1% of people surveyed stated that they did not avoid leaving the house due to COVID-19 (Figure 1G), 15.5% of respondents did not want to be tracked with the NHS Track and Trace app, and 9.4% did not trust what the government would do with that data (Figure 1I). Only 6.6% of people stated that they would not get vaccinated when a vaccine became available (Figure 1D). Although 0.6% of respondents reported not following social distancing guidelines at all, 37.8% reported not doing so all of the time (Figure 1H). Similarly, although 0.4% of responders reported not wearing a mask (Figure 1F) 13.5% only did so because they had to, a percentage that increased to 18.6% in June. Reflecting changes in guidance, the most striking change from December 2020 to January 2022 was that people no longer avoided leaving their house due to COVID-19, from 18.9% to 49.9% in June 2021 and 49.2% in January 2022 (Figure 1G).

In December 2020 the most common reported source of news about the pandemic was via TV (26.5%), and the Internet (21.9%), with social media accounting for 7.3% of responders' primary source of information surrounding the pandemic (Figure 1J). In May 2020, the proportions changed such that marginally more respondents took their information about COVID-19 from governmental communications and less from other sources.

1.2 Predicting lack of compliance on the basis of trust

Next, we investigated whether individual respondent's compliance with COVID-19 safety measures could be predicted from trust in government and media advice about COVID-19-related issues. The answers given to the five compliance questions (wearing a mask, taking the vaccine, following social distancing guidelines, avoiding leaving the house and using the NHS app) were converted into continuous numeric scales and aggregated as a composite compliance score (see Methods). A multiple linear regression model was fitted with the composite compliance score which was adjusted to demographics measures (age, residence, ethnicity, education, occupation) as the dependent variable and yes/no responses to each trust question as predictors. R^2 for the model was 0.13. Table 1 shows the results of this analysis and identifies which aspects of trust predicted lack of compliance (Table 1). Affirming any of the following four statements predicted a lack of compliance with the suggested measures for containing the pandemic: that restrictions are unjustified, that the government has ulterior motives, that the media is hiding things, and skepticism that COVID-19 has natural origins. Considering that restrictions are poorly

justified had the largest regression coefficient. Therefore, if someone considers that restrictions are poorly motivated then they are less likely to follow them, with other aspects of distrust having small additive effects to this.

Lack of trust	Regression Coefficient	p-value	Meaning
Intercept	-0.12	***	Those who trust are more likely to comply
More restrictions are needed	-0.30	***	Affirmative respondents more likely to comply
Restrictions are not justified	1.51	***	Affirmative respondent less likely to comply
The government has ulterior motives	0.10	***	Affirmative respondent less likely to comply
The media is hiding things	0.17	***	Affirmative respondent less likely to comply
COVID-19 has unnatural origins	0.07	**	Affirmative respondent less likely to comply

Table 1. Predictors of non-compliance with suggested measures to reduce transmission of COVID-19. Effect size (Cohen's *d*) is calculated by dividing the beta coefficients calculated in the linear regression by the square root of the *N* times the standard errors. Significance is denoted as *, **, and *** for $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively.

1.3 Predicting trust and non-compliance based on sociodemographic factors, substance use, preexistent neurological or psychiatric conditions, wellbeing, time spent online and information sources

We examined whether sociodemographic variables, levels of substance use, neuropsychiatric status, news sources and time spent online were predictive of whether respondents were distrustful and/or non-compliant with COVID-19 safety measures. Binomial and multinomial logistic regression models were trained to predict the categorical responses for each of the minority answers for questions in Figure 1 (see methods).

Factor	A.	B.	C.	D.	E.	F.	G.	H.	I.
Age	***	***	***	***	***	***	***	***	***
Sex	***	***	***	0.088	**	***	***	***	***
Residence	0.761	***	**	0.178	**	**	0.806	**	***
Ethnicity	0.97	*	**	0.158	***	*	***	*	***
Education	***	***	**	***	***	**	***	**	*
Occupation	0.383	0.19	0.058	0.062	***	***	*	***	***
Neuropsychiatric status	0.236	***	*	**	0.137	***	0.296	***	**
Drug use	*	**	***	***	*	***	***	***	***
News source	***	***	***	***	***	***	***	***	***
Wellbeing	0.864	***	***	***	**	**	***	***	***
Time online	*	***	*	**	0.288	***	0.249	***	***

Table2. Sociodemographic factors predict trust and non-compliance. Questions B,G, H and I were fitted as multinomial logistic regression models, and questions B-F were multinomial logistic regression models. A - Does not wear a mask at all times, B - Lockdown restrictions are not justified, C - There are ulterior motives behind the government's response to COVID-19, D - The media is hiding things from the public, E - Would not take the vaccine, F - Does not socially distance, G - COVID-19 is not a naturally occurring phenomena, H - Does not avoid leaving the house due to COVID-19, I - Does not use the NHS app. Significance is denoted as *, **, and *** for $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively.

The factors added to the model were, in this order, age decade, sex, residence, ethnicity, education, occupation, psychiatric or neurological conditions, drug use frequency prior to the pandemic, primary news sources about the pandemic, wellbeing score determined from the GAD-7 (25) and PHQ (26) anxiety and depression scales, and number of hours spent online per day. The impact of each of the factors on driving either trust-related or compliance-related questions is illustrated in Table 2. The significance of each factor was calculated from the log likelihood ratio of the model containing all factors and the model without that specific factor.

At the factor level, age, education, frequency of drug use and news sources all showed significant prediction of distrust and compliance across all questionnaire items. Notable factors predicting distrust and noncompliance were also sex for all questions apart from the media hiding things, and wellbeing level for all questions apart from following the social distancing guidelines. Neuropsychiatric status was a predictor of all but trust in vaccination, following the social distancing guidelines and distrusting that COVID-19 is a natural phenomenon. Ethnicity also played a role in driving trust about lockdown

restrictions, the government response to COVID-19, vaccination, the origins of COVID-19, leaving the house and using the NHS app.

At predictor level, compliance and distrust varied significantly with sociodemographics. The most salient sociodemographic and lifestyle predictors are illustrated in Figure 2 and the full spectrum of predictors including odds ratios and significance values are appended in the Supplementary Materials. Odds ratios were calculated from the regression coefficients resulting from all complete binomial and multinomial logistic regression models. Younger people (18-30 years old) had twice the odds of older people (51-60 years) to not follow social distancing guidelines. People with low education levels showed high odds of distrusting the government, media and mainstream COVID-19 narrative, as well as to not follow the social distancing guidelines compared to people with university degrees.

Frequent drug users were more likely to not wear masks and not follow the guidelines compared to people who never used drugs, but also to distrust the government and mainstream media. People with low wellbeing were around 1.5 times more likely than people with high wellbeing to distrust the government including what they would do with their data from the NHS app, distrust the media, only following guidelines for masks because they have to, have given up on following them or never followed them to begin with.

1. **Common beliefs underlying distrust: a topic-modelling analysis of free text responses**

Topic modelling was applied to free-text responses to provide further insights into the most common beliefs people held that motivated distrust in government and media. Specifically, respondents were presented with a free text box if they answered specific questions indicating distrust. They were asked to provide reasons for responding that (1) restrictions were either unjustified or more needed to be done, (2) the government had ulterior motives, (3) the media was hiding things, (4) they would refuse the vaccine, and (5) they thought the origins of COVID-19 were unnatural. LDA models were trained on the resulting free-text data from these questions to capture in a data-driven manner the responders' most common beliefs as commonly occurring latent documents or 'topics'.

After removing infrequent words, non-words, lemmatising and tokenising, the optimal number of topics for each question given the corpus of responses was chosen based on the number of topics from 1 to 30 where Cv coherence was highest. Then, LDA was used to infer this number of latent topics from the corpus (see Methods and Supplementary materials).

2.1 Distribution of beliefs across timepoints

We illustrate the dominant topics derived from the answers to the free text questions in Figure 3, as well as whether there was a change in topic distribution from December 2020 to January 2022. Chi-squared tests illustrated that changes in the prevalence of beliefs was significant for every question assessed. People who indicated that they felt more restrictions were needed, tended to justify this distrust on the

basis of topics such as the need to protect the NHS, infections rising, lockdowns being needed, and that rules had to be stronger and clearer. They also believed that the government needed to be more proactive and the reason for amplifying restrictions was because many people did not follow the guidelines. The highest % change between December 2020 and January 2022 is for the topic related to infections, transmission, hospitalisation and death rates, rising concerns which sky-rocketed from 10.1% in December 2020 to 27.1% in January 2022.

Conversely, people who endorsed restrictions being unjustified stated that this was due to their impact on the youth, the fact that they were not focused specifically on the vulnerable, the contradictory nature of the policies, and most prevalently because of the impact they had on the economy. The changes in the prevalence of belief topics was substantial. The most striking one was related to the government inconsistencies about the restrictions which saw an increase from 13.9% in December 2020 to 27.1% in January 2022.

The sub-population who distrusted the government prevalently believed that they had ulterior motives because they made inconsistent decisions with regards to COVID-19 policies, thereby prioritising the gain of popularity with the public or businesses rather than population health. This topic saw an increase from 18.8% to 25.7% from December 2020 to June 2021, followed by a decrease to 18.9% in January 2022. The remaining topics featured the government's lack of transparency, prioritising economy over health, using COVID-19 to distract from other issues such as Brexit, and using the pandemic for personal gain. The most prevalent belief in January 2022 was that the government has been dishonest with the public and ignored advice of experts (24.5%).

People who distrusted the media believed that they shared only partial information or sides of the stories, focused on the negative and shocking stories, did not encourage discourse critical of the government, and were unreliable. A substantial proportion among the distrustful (20.5% in December 2020, 20.6% in June 2021, 19.9% in January 2022) believed that the media not only did not present an impartial view but also were influenced by politics, the global economy, and investors, making this the most popular topic.

Only a small proportion of people in our sample indicated that they would refuse the COVID-19 vaccination. However, those who did so either quoted personal health circumstances that made them exempt, were unsure about vaccines in general, did not see the benefits compared to the risks, or worried vaccines were being rolled out too quickly and might have long term side effects that had not yet been assessed. The latter topic showed an increase from 36.7% in December 2020 to 53.5% in January 2022.

Lastly, people had a range of views concerning the origins of COVID-19. Respondents believed it was either a bioweapon developed in China which was accidentally released, a virus leaked from a lab, a zoonosis resulted from our interference with the biosphere or from food markets, that it resulted from improper treatment of animals, or were unsure about its origins at all. The most prevalent belief in January 2022 was that the virus is the result of improper treatment of animals and interference with their habitats (35.5%), and the least prevalent, that it was created by humans and deliberately/accidentally released from a lab (8.6%).

2.2 Specific beliefs covary with sociodemographic factors

Topic probabilities for the entire cohort surveyed varied substantially with sociodemographic factors (Figure 4). Chi-squared tests confirmed that the distribution of best-fitting topics varied significantly with sociodemographic factors. Further chi-squared tests and Pearson's correlations (for incremental factors) were conducted to determine if each individual topic varied in prevalence across sociodemographic categories (See Supplementary Materials).

Younger people (18 to 30 years old) believed that more had to be done to protect the frontline workers, the lockdown should have been more prompt and restrictions should have been stricter; they were also the group most likely to distrust the government and media in all major reasoning themes. In comparison, older people (above 60 years old) had the highest probability for distrusting the COVID-19 narrative. Specifically they believed that COVID-19 is either a product of humans mistreating animals or man-made, where opinions diverge into either intentionally as a bioweapon or accidentally released.

People who identify as 'Other' (rather than male or female) were more likely to believe more needs to be done, but also to be more distrustful of the government, media, and believe that COVID-19 is a zoonosis resulting from human interference with the biosphere or be unsure about the virus origins.

Minority ethnic groups were more likely to have divergent beliefs in all domains studied, apart from believing that restrictions are unjustified.

Respondents educated at PhD level were distrustful of the government because they believe that the economy has been prioritised over public health. In comparison, people with lower education levels had higher probability of believing that more needed to be done because the government's response should have been more proactive and many people do not follow the guidelines. They also had higher probabilities to generally distrust the media. The effect of low education levels, however, was most apparent in regards to distrusting the mainstream COVID-19 narrative.

Having a psychiatric or neurological condition was linked to having a higher probability to distrust the authorities in all cases, but lower probability to believe that COVID-19 originated from the wet markets, was a man-made bioweapon or that restrictions were illogical. Overall, people suffering with low wellbeing, frequent drug users, and those who spent most of their day online had higher probabilities for all topics in all cases.

Discussion

Distrust predicts non-compliance

Our results confirm that distrust in government and media, and conspiracy theories more broadly, have motivated a substantial minority within the British public to be non-compliant with key aspects of government rules and advice during the COVID-19 pandemic (6, 13, 27, 28). Indeed, in some respects the

level of distrust appears to have increased as the pandemic progressed. For example, while in December 2020 and June 2021 around a quarter of respondents indicated that the government had ulterior motives behind its response to the virus, this grew to almost a third by January 2022. Distrust in the mainstream COVID-19 narrative increased in similarly large proportions. There was a corresponding change in compliance across this timeframe. Social distancing significantly decreased and more people only wore masks because they had to. Some components in these changes in compliance, such as the dramatic reduction in respondents no longer avoiding leaving the house by January 2022, will have reflected changes in context and government advice. However, our results indicate that increased distrust played a substantial role as well, for example, endorsing the statement that restrictions were unjustified provided the strongest predictor for non-compliant behaviour.

Distrust and the beliefs that underpin distrust are diverse

Free text analyses enabled us to pinpoint in a data-driven, unbiased, and unconstrained manner the nature of beliefs that motivated specific aspects of respondents' distrust in government and media. The respondents highlighted diverse topics that spanned from perceptions of incompetence and inappropriate prioritisation on the part of government to beliefs in nefarious conspiracies. Most relevant to compliance, the substantial proportion of people who endorsed the view that the government had ulterior motives differed in the reasons that they gave. Some highlighted that expert advice was ignored or that the economy was being prioritized above public health. Others held more conspiratorial beliefs regarding corruption, with decision making perceived to be motivated by attempts to bolster popularity or for personal gain. A not insignificant proportion of respondents believed that the pandemic was being used to distract people from the UK's exit from the European Union. Concerning distrust in mainstream media, respondents considered that the media was unreliable, inaccurate, sensationalist, motivated by a lack of impartiality, was pursuing an agenda, or deliberately manipulating public discourse. Topics surrounding the origins of COVID-19 ranged from improper hygiene in wet markets and human-induced destruction of wild habitats to more conspiratorial beliefs about the accidental or deliberate release of an engineered bioweapon. Common reasons given to justify vaccine refusal were narrower, converging on a common theme of concerns about side effects, exemptions due to illnesses or pre-existing medical conditions, and reticence in the light of a fast rollout.

The distribution of beliefs that underpin distrust changes over time

The fact that a diversity of beliefs underlie distrust presents an obvious challenge for any attempt at population-level persuasion towards improved compliance. Exacerbating this, we observed that distrust and beliefs significantly evolved over a short period of time. From December 2020 to January 2022 we saw not only a significant shift in the proportions of distrust amongst our respondents but also a change in the prevalence of their beliefs. The proportion of people nurturing feelings of distrust towards the government grew from ~ 24% to ~ 30% and the proportion of people who thought that COVID-19 has unnatural origins grew from ~ 19–28%. The most prevalent topics underlying distrust in the government and the mainstream COVID-19 narrative which accompanied this change were that the government

ignored expert advice (from ~ 21% to ~ 25%), and that COVID-19 was caused by humans mistreating animals/nature (from ~ 18% to ~ 35%). Conversely, the proportion of people who would refuse the COVID-19 vaccine dropped from ~ 7% to ~ 2% while over half of the residual unvaccinated population believed that there was a lack of data on long term side effects in January 2022 compared to only a third in December 2022. Finally, the percentage of those surveyed who believed restrictions were unjustified increased from ~ 10% to ~ 13% and the most prevalent belief amongst those was held by ~ 28% in January 2022 of the respondents who indicated that the physical, mental, and economical toll on the population was unjustified compared to ~ 18% in December 2020.

Non-compliance, distrust and beliefs covary with sociodemographic variables

The prevalence of non-compliance, distrust, and the beliefs that underpin distrust also covaried significantly with population demographic variables, and were subject to change over time based on these factors. An overarching pattern was evident, namely that distrust was higher in minority, alternative and otherwise disenfranchised segments of society.

Perhaps most notably, all ethnic minorities in our sample had higher probability to both distrust the government and refuse vaccines, which is consistent with previous studies (29). This is problematic from a public health perspective, especially since these groups also have higher probabilities of suffering negative mental (30) and physical health effects from COVID-19 (31, 32). Previous research has highlighted that marginalised groups are more likely to disbelieve that the government seeks to promote their wellbeing (33). Interestingly, whilst ethnic minorities showed the highest probabilities for being distrustful about the government, the media, vaccines, and COVID-19 narratives, they did not have a significantly higher probability of considering restrictions unjustified. This may indicate that higher non-compliance in ethnic minorities is more the product of distrust in authoritative bodies than it is lack of belief in risks posed by the virus.

Young adults, students, those who identify as 'other' rather than male or female, and frequent drug users in our study were more likely to distrust government and media handling of the pandemic. They were also less likely to comply with social distancing isolation guidelines, which accords with previous research detailing how young people were severely impacted by restrictions in the short- and long-term as well as less likely to comply with pandemic guidelines (34). Those who did not complete education past primary school level had the highest probability of holding conspiratorial beliefs about the origins of COVID-19. Consistent with other studies, education was the strongest predictor we found for distrusting the government and not following COVID-19 related guidelines (6, 14, 29, 35–37). This highlights a need for clear scientific communication of expert advice which is accessible to those of all levels of formal education. Targeting sources of misinformation too is especially important since conspiratorial thinking has been linked to a decrease in willingness to comply with containment measures (38).

People who were retired, disabled, or who had poor mood self-assessment scores were more likely to consider that stricter restrictions were needed. Relatedly, those who had neurological or psychiatric

conditions were more likely to believe that ulterior motives were behind the government's response to the pandemic. Similarly, poor mood self-assessment scores predicted the probability of believing that restrictions needed to be amplified, the government had ulterior motives, the media was hiding things, and COVID-19 had unnatural origins. These observations accord well with past literature showing that people exhibit divergent beliefs when they are more prone to psychological disorders (39, 40), are anxious (40, 41), have patterns of cognition characterised by feelings of helplessness and hopelessness (15), or hallmark characteristics of depression (42). Although warnings have been put forward highlighting that psychological reactions contribute to how individuals and groups respond to the pandemic (43, 44), insufficient efforts aimed at prevention led to a phenomenal rise in the incidence of anxiety, depression (22, 23) and substance use (45). We also found that illicit substance use was predictive of a lack of trust in the government, believing that the media is hiding things, and of the refusal of the vaccine. Furthermore, people who suffer with mental health difficulties or substance use disorders pose higher risks of COVID-19 infection and possible complications associated with the virus, due to increased levels of difficulty accessing disease management services (43).

Relationship of findings to other free text studies

A number of previous studies have investigated reasons for distrust in the government based on text analysis of social media data. These studies have reported results that align with key aspects of our findings. For example, one study using LDA to model Twitter responses focused on understanding the common themes underlying anxiety in the general population and compared this across nations (19). In accordance with the results presented here, a theme identified to cause anxiety in the public was political polarisation and governmental incompetence, which we identified as a reason for distrusting the government. Another study used LDA to model Twitter data in the UK to derive topics of interest during the pandemic and reported that a topic related to lockdown and one related to the government were identified, but no thematic analysis was performed to understand exactly what the problems were (20). Though social media studies pose advantages – such as enabling the passive collection and modelling of data in a near continuous fashion – a major methodological limitation of these studies is that social media allows one to conceal their real identity. This limits the potential for mapping beliefs as a function of population variables. Furthermore, inferences that a topic is wide-spread within a group of individuals should be taken with a grain of salt. This could be due to spam coming from a single individual or set of bots rather than being representative of the views of the general population. Social media platforms are also increasingly riddled with gate-keeping narrative frameworks, which makes transparency an issue when interpreting the available content (46).

Through the use of citizen science studies we provide a compromise whereby public opinion is efficiently surveyed from the general population at large scale but is also accompanied by detailed socio-demographic data. In accordance with results presented here, the limited number of studies carried out in the UK analysing free text data have found general mistrust of governmental politics, concerns about the use of scientific evidence, and quality of communication and political decision making, such as a lack of transparency, corruption and incompetence within the government, inconsistent policies, tensions

between the government and the media and the public, and ultimately, an inability to plan for an uncertain future were highlighted (13, 14).

Limitations

Our study is distinct in evaluating the prevalence of distrust and beliefs at discrete timepoints throughout the pandemic in relation to diverse population variables. Some limitations should however be considered. Though large, our sample is not proportionally representative of the UK population; we highlight a bias toward older adults and those holding a university degree. Nevertheless, we note that the population is inclusive of a diversity of people such that associations of distrust and beliefs can be modeled across a comprehensive set of population variables with considerable sensitivity even to small associations. Moreover, a substantial number of individuals from different minority groups who are typically less likely to volunteer to be part of a research study – such as illicit substance users and persons suffering with neurological or psychiatric conditions – are represented within our study. We posit that this inclusivity reflects that our cohort was recruited primarily through the BBC, that is, as opposed to governmental communication through official survey platforms that the most authority-sceptical are most likely to distrust. In keeping with all sampling methods, it is the case that we likely underestimate levels of distrust and non-compliance in the general population, as those who are distrustful are likely to be more sceptical of taking part in survey-based studies (13, 27). We also included responses from people who live abroad and may not feel the direct effects of the local policies, and hence have somewhat skewed opinions. However, given the low proportions (at most 6.5% of the total sample at each timepoint) and lack of association of country of residence with topic prevalence, this is unlikely to have driven our results. It also is important to highlight that a minority of individuals completed all of the analysed timepoints. Notably though, the demographic profiles remained similar across timepoints. Furthermore, the results are consistent when analysing only those who completed all three timepoints. Finally, our results illustrate a number of robust topics that are distinct and do reflect common themes in the British people surveyed about the COVID-19 pandemic. We used LDA, one of the oldest and most established, and proven to work well topic modelling techniques. However, a diversity of variants on topic modelling methodology are available. Future work should explore application of alternative topic modelling methods and coherence algorithms. Nonetheless, the topics derived using the approach applied here have clear validity, being interpretable and correlating robustly with population variables in a manner that makes intuitive sense, whilst both replicating and extending findings from past studies.

Future directions

Taken together, the diverse and shifting nature of compliance, distrust, and beliefs that underlie distrust presents a major challenge to any efforts to improve compliance with government guidelines. The COVID-19 pandemic has presented a unique opportunity to examine these relationships between the public and the bodies of authority at a finer grain than has been previously possible. However, the general principles that have been observed have relevance to other future pandemics, crisis situations, and more generally to the interplay of government, media, and population behaviour. It seems reasonable to suggest that aspects of the methodology applied here could be applied for this type of purpose. Specifically, the

combination of large-scale citizen science survey data and machine learning can be used to rapidly identify the currently most prevalent beliefs that underlie distrust and ultimately non-compliance. These may then be addressed in a manner that is tailored to individuals by mapping belief prevalence to the most relevant population variables. We believe that stratifying people according to sociodemographic variables has the potential to be a useful tool in the process of tackling lack of trust and compliance, and suggest to go a step further: instead of targeting with health messaging alone, tailor information campaigns aimed at increasing trust in the government specifically to topics those groups have the highest probability of basing their reasoning on.

Materials And Methods

I. Experimental design

Data collection and participant demographics

In January 2020 the study was promoted by the BBC on the main and on the news homepages. Subsequently, in May 2020, the study was promoted again via the same channels and by AH on an episode of the BBC2 Horizon Documentary named 'the Great British Intelligence Test'. Online links remained in prominent positions over the subsequent month in each promotion launch. The study survey consisted of a set of bespoke cognitive assessment tasks and of questionnaires including items related to socio-demographic and mental health variables. The survey was programmed in HTML5 with JavaScript and deployed via our Cognitron server system, which already supports diverse online studies. Cognitive and mental health results have been reported elsewhere (21–24). At the end of the survey, participants who wished to be recontacted for future studies were asked for their email addresses.

Out of the total of the 765,830 participants recruited in December 2019 and May 2020 only 95,441 signed up with their email addresses and gave permission to be recontacted. The totality of the participants is detailed in Table 3.

In December 2020 the questionnaires were expanded to include additional variables; among these, people were asked a set of fixed-answer questions regarding the coronavirus pandemic. These pandemic-related questions were of three types: i) regarding people's beliefs about the pandemic, ii) regarding people's compliance with the measures suggested to reduce the spread of COVID-19 and iii) regarding people's primary information source on the pandemic. The questions concerning people's beliefs triggered a free-text question in case people picked what was expected to be the 'minoritarian' answer to allow them to explain the motives behind their choice. The pandemic-related questions relevant to the present work are here reported; for the full transcript of all the questions asked in the survey refer to the Supplementary Materials.

Data was analysed from 20,922 individuals in December 2020, from 12,796 individuals in June 2021 and from 14090 individuals in January 2022 who completed all questions that are subject of this study. Their

demographics are shown in the Supplementary Materials.

This study was run in accordance with the Helsinki Declaration of 1975, as revised in 2008. All procedures were approved by the Imperial College Research Ethics Committee (17IC4009). All participants provided informed consent prior to completing the survey.

Table 3

Recruitment timeline and sample sizes. *All 95,441 participants who signed up with their email addresses were recontacted in December 2020. 397 of those deleted their accounts between December 2020 and June 2021 when the next recontact happened. 1303 further deleted their accounts between June 2021 and January 2022. In January 2022 people who signed up to the Great British Intelligence Test between December 2020 and January 2022 were also part of the recontacting scheme. At the recontacting in January 2022 104,516 people received a second email asking for their responses.*

		December 2019	May 2020	December 2020-January 2022
	Recruited	533,248	232,582	30,754
December 2020	Recontacted	95,441		N/A
	Responded	33,227		
	Complete responses	20,922		
June 2021	Recontacted	95,044		
	Responded	33,200		
	Complete responses	12,796		
January 2022	Recontacted	124,496		
	Responded	25,137		
	Complete responses	14,090		

Scoring for text about compliance

We wanted to derive a composite score representative of the overall level of compliance of an individual with the COVID-19 restrictions, to see if we could predict people’s non-compliance on the basis of their beliefs. With this in mind, we scored the possible answers to the compliance questions on a continuous scale from 0 to 2, with 0 being fully compliant, depending on the question. Next, for each individual we evaluated the score for each question based on their answers and we added up all the scores to get a unique composite score. The compliance composite score was, therefore, assigned on a scale from 0 to 10, where 0 means fully compliant and 10 non-compliant at all with the government suggestions regarding conduit during the COVID-19 pandemic. The scoring system for each question is Supplementary Materials.

II. Statistical methods

Predicting lack of compliance on the basis of beliefs and media sources

To investigate whether any of the beliefs as phrased by the fixed-answer belief questions and the primary pandemic news source were predictive of non-compliance with the COVID-19 restrictions, we have fitted a linear regression model taking as features the items of the belief and media questions and predicting as target the compliance composite score.

Predicting beliefs based on questionnaire variables

Next, we wanted to verify whether the responders holding divergent/minoritarian responses in the belief questions were more likely to belong to a certain sub-population, or have certain characteristic traits. Therefore, we performed binomial and multinomial logistic regression on the data using Python `statmodels` library (47) in order to identify questionnaire variables predictive of people more likely to answer a question in a distrustful/noncompliant way. The questionnaire variables investigated include sociodemographic features (age, gender, occupation, education, ethnicity and country of residence), levels of substance use, neuropsychiatric status, wellbeing and number of hours spend online per day and primary source of news about the pandemic. All empty values were dropped, and dummy variables were computed for all categorical data including the predictors and the dependent variables. The predictor categories with the highest number of participants acted as reference categories and were removed from the models. We also checked for highly correlated variables and found that questions such as those asking about mask wearing and about socially distancing had very low numbers of responses in certain categories, so those were collapsed together and the questions entered binomial rather than multinomial regression models. For each model odds ratios were calculated to reflect odds compared to the removed highest scoring dummy. Full logistic regression outputs can be found in Supplementary Materials.

Topic modelling on free-text data

For the belief and counterfactual compliance questions, in case of a particular response (in most cases what was expected to be the minoritarian response), respondents were asked to specify the rationale driving their answers. Topic modelling was used to exploit and extract the major classes of opinions justifying divergent answers from the free-text data in an unsupervised fashion. For each question the following topic modeling pipeline was applied: a) pre-processing of the free-text data; b) selection of the optimal number of topics; c) implementation of LDA as topic modeling algorithm; d) rating of top 20 words and opinions per topic; e) interpretation of top opinions to label/name the topic.

The following subsections cover this in more detail.

i) Free-text data pre-processing

Free-text data was screened manually for non-related responses/spam responses and those entries were removed. Established natural language processing techniques were applied to pre-process the text data using the NLTK package in Python (48). These involved removing stop words, junk words, punctuation, special characters, numbers, words that occurred less than 10 times and lowercasing and lemmatizing the remaining text (bringing each word to its dictionary form). Data was then tokenised and vectorised (bag of words model).

ii) LDA implementation and selection of optimal number of topics

LDA falls under the rubric of topic modelling, a popular approach for both social network analysis and mapping out public opinions from free-text published on social media platforms (49).

LDA is the most popular algorithm for segregating free text data into topics; it uses a Dirichlet distribution prior to infer latent topics from a set of data (50). Broadly speaking, LDA analyses a *corpus* of *texts* (in our case, the collection of free-text responses to a given question) to identify *latent topics* – putatively, themes or topics which appear across multiple texts within the corpus (we interpret latent topics here as common justifications for rejecting commonly held beliefs). Crucially, in contrast to fixed-answer questionnaires alone, this approach allows for important classes of public opinion to be identified *a posteriori*. LDA works on the assumption that documents input to the model are composed of some common underlying mixture of topics and seeks to both derive those latent topics and quantify each document's relationship to them (49).

To group the opinions of the members of the British public into topics, we ran LDA on our data. Coherence measures identify the optimal number of topics in the dataset. We used the Cv coherence measure previously shown to be the most highly correlated with human judgement (51, 52) than other measures, such as algorithmic perplexity (51, 53).

For finding latent topics in the free-text data we used the LDA multicore implementation in gensim (54) which uses online LDA (55), an implementation based on online, stochastic, multithreaded optimisation with the intent of speeding up the computational inference process.

iii) Topic modelling evaluation

The free-text data was segmented into word pairs for the first step of computing Cv coherence. To calculate Cv coherence the probability that each word pair belongs to a certain topic was computed as well as a measure that quantifies how strongly a wordset supports another word set (52).

iv) Topics labelling and opinions interpretation

The topic mixture, or percentage that each opinion belongs to its assigned topic, was used to determine the 20 opinions that most strongly relate to each topic. These opinions were used as a corpus to perform thematic analysis, where topic labels were generated using thematic analysis as described in (56). For each topic, the opinions were coded, and themes were extracted from the codes. Traditional thematic

analysis often results in several themes generated from the corpus; due to the size of our corpus, we often extracted 1–2 themes. The most dominant theme was used as the label for each topic. Thematic analysis was independently performed by DK, DC, and AL, and the topic labels (themes) were compared to generate a consensus topic label. The saved LDA models were used to infer the topic distribution in the opinions belonging to the newest time point as well.

Declarations

Affiliations

1. Department of Brain Sciences, Imperial College London
2. Neuromodulation Lab, Department of Psychology, Faculty of Health and Medical Sciences, University of Surrey
3. UKRI Centre for Doctoral Training in AI for Healthcare, Department of Computing, Imperial College London
4. UK Dementia Research Institute: Care Research & Technology
5. Engineering and Physical Sciences Research Council CDT Neurotechnology, Imperial College London
6. NMR unit, Queen Square Multiple Sclerosis Centre, UCL, Queen Square Institute of Neurology, Department of Neuroinflammation, Faculty of Brain Sciences, University College London
7. Centre for Neuroimaging Sciences, Institute of Psychiatry, Psychology and Neuroscience, King's College London

Acknowledgments

The authors would like to thank the participants of this study who kindly responded to our emails and therefore allowed us to collect generous amounts of data. We would also like to thank Eyal Soreq for thoughts on the early stage of the free text analysis. Finally, MB would like to thank her mother for repeatedly traveling across the continent to babysit her newborn while she worked on this paper.

Funding

MB is supported by the Medical Research Council Doctoral Training Programme at Imperial College London. DLK is supported by the Vice Chancellor Studentship at the University of Surrey. DCG is supported by the Department of Brain Sciences, Imperial College London. AL is supported by the UK

Research and Innovation Centre for Doctoral Training in AI for Healthcare <http://ai4health.io> (grant number EP/S023283/1) and the UK MS Register. DC is supported by the UK Dementia Research Institute: Care Research & Technology Centre. WT is supported by the EPSRC Centre for Doctoral Training in Neurotechnology. PJH is, in part, supported by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. AEJ is supported by the Queen Square Multiple Sclerosis Centre at University College London. AH is supported by the UK Dementia Research Institute Care Research and Technology Centre and Biomedical Research Centre at Imperial College London.

Author contributions:

Conceptualization: MB, AEJ, AH

Methodology: MB, AH, DLK

Investigation: MB

Visualization: MB, AL, DLK

Supervision: AH

Software: WT, PJH, AH

Writing—original draft: MB, DLK, DCG, DC, AL

Writing—review & editing: MB, DLK, DCG, AL, WT, AEJ, DC, PJH, AH

Competing interests

Professor Hampshire is owner and director of Future Cognition LTD and H2 Cognitive Designs LTD, which support online studies and develop custom cognitive assessment software respectively. Dr. Hellyer is owner and director of H2 Cognitive Designs LTD and reports personal fees from H2 Cognitive Designs LTD outside the submitted work. All other authors declare they have no competing interests.

Data and materials availability

The assessment software used in this study is made available for academics via the Cognitron web server on request to AH. Requests for data should be directed to AH: a.hampshire@imperial.ac.uk

References

1. M. Janssen, H. van der Voort, Agile and adaptive governance in crisis response: Lessons from the COVID-19 pandemic. *International Journal of Information Management*. **55**, 102180 (2020).
2. K. Sharma, S. Seo, C. Meng, S. Rambhatla, Y. Liu, COVID-19 on Social Media: Analyzing Misinformation in Twitter Conversations. *arXiv preprint arXiv:2003.12309* (2020).
3. A. Bridgman, E. Merkle, P. J. Loewen, T. Owen, D. Ruths, L. Teichmann, O. Zhilin, The causes and consequences of COVID-19 misperceptions: Understanding the role of news and social media. *Harvard Kennedy School Misinformation Review*. **1** (2020).
4. R. Imhoff, P. Lamberty, A Bioweapon or a Hoax? The Link Between Distinct Conspiracy Beliefs About the Coronavirus Disease (COVID-19) Outbreak and Pandemic Behavior. *Social Psychological and Personality Science*. **11**, 1110–1118 (2020).
5. B. Duffy, Life under lockdown: coronavirus in the UK. Kings College London (2020).
6. D. Freeman, F. Waite, L. Rosebrock, A. Petit, C. Causier, A. East, L. Jenner, A.-L. Teale, L. Carr, S. Mulhall, E. Bold, S. Lambe, Coronavirus conspiracy beliefs, mistrust, and compliance with government guidelines in England. *Psychological Medicine*. **52**, 251–263 (2022).
7. G. Gozgor, Global Evidence on the Determinants of Public Trust in Governments during the COVID-19. *Applied Research Quality Life* (2021).
8. Q. Han, B. Zheng, M. Cristea, M. Agostini, J. J. Bélanger, B. Gützkow, J. Kreienkamp, P. Collaboration, N. P. Leander, Trust in government regarding COVID-19 and its associations with preventive health behaviour and prosocial behaviour during the pandemic: a cross-sectional and longitudinal study. *Psychological Medicine*, 1–11 (2021).
9. J. Murphy, F. Vallières, R. P. Bentall, M. Shevlin, O. McBride, T. K. Hartman, R. McKay, K. Bennett, L. Mason, J. Gibson-Miller, L. Levita, A. P. Martinez, T. V. A. Stocks, T. Karatzias, P. Hyland, Psychological characteristics associated with COVID-19 vaccine hesitancy and resistance in Ireland and the United Kingdom. *Nat Commun*. **12**, 29 (2021).
10. M. Zajenkowski, P. K. Jonason, M. Leniarska, Z. Kozakiewicz, Who complies with the restrictions to reduce the spread of COVID-19?: Personality and perceptions of the COVID-19 situation. *Personality*

- and Individual Differences. **166**, 110199 (2020).
11. D. A. N. Goldstein, J. Wiedemann, "Who Do You Trust? The Consequences of Political and Social Trust for Public Responsiveness to COVID-19 Orders" (SSRN Scholarly Paper 3580547, Social Science Research Network, Rochester, NY, 2020).
 12. C. Clark, A. Davila, M. Regis, S. Kraus, Predictors of COVID-19 voluntary compliance behaviors: An international investigation. *Global Transitions*. **2**, 76–82 (2020).
 13. L. Wright, A. Steptoe, D. Fancourt, Predictors of self-reported adherence to COVID-19 guidelines. A longitudinal observational study of 51,600 UK adults. *The Lancet Regional Health - Europe*. **4**, 100061 (2021).
 14. L. Enria, N. Waterlow, N. T. Rogers, H. Brindle, S. Lal, R. M. Eggo, S. Lees, C. h Roberts, Trust and transparency in times of crisis: Results from an online survey during the first wave (April 2020) of the COVID-19 epidemic in the UK. *PLOS ONE*. **16**, e0239247 (2021).
 15. K. M. Douglas, R. M. Sutton, A. Cichocka, The Psychology of Conspiracy Theories. *Curr Dir Psychol Sci*. **26**, 538–542 (2017).
 16. R. M. Sutton, K. M. Douglas, Conspiracy theories and the conspiracy mindset: implications for political ideology. *Current Opinion in Behavioral Sciences*. **34**, 118–122 (2020).
 17. R. M. Sutton, K. M. Douglas, Agreeing to disagree: reports of the popularity of Covid-19 conspiracy theories are greatly exaggerated. *Psychological Medicine*. **52**, 791–793 (2022).
 18. A. Hampshire, P. J. Hellyer, W. Trender, S. R. Chamberlain, Insights into the impact on daily life of the COVID-19 pandemic and effective coping strategies from free-text analysis of people's collective experiences. *Interface Focus*. **11**, 20210051 (2021).
 19. S. Kwon, A. Park, Understanding user responses to the COVID-19 pandemic on Twitter from a terror management theory perspective: Cultural differences among the US, UK and India. *Computers in Human Behavior*. **128**, 107087 (2022).
 20. I. K. Cheng, J. Heyl, N. Lad, G. Facini, Z. Grout, Evaluation of Twitter data for an emerging crisis: an application to the first wave of COVID-19 in the UK. *Sci Rep*. **11**, 19009 (2021).
 21. A. Hampshire, Great British Intelligence Test Protocol, 4 (2020).
 22. S. R. Chamberlain, J. E. Grant, W. Trender, P. Hellyer, A. Hampshire, Post-traumatic stress disorder symptoms in COVID-19 survivors: online population survey. *BJPsych Open*. **7** (2021).
 23. A. Hampshire, P. J. Hellyer, E. Soreq, M. A. Mehta, K. Ioannidis, W. Trender, J. E. Grant, S. R. Chamberlain, Associations between dimensions of behaviour, personality traits, and mental-health during the COVID-19 pandemic in the United Kingdom. *Nat Commun*. **12**, 4111 (2021).
 24. A. Hampshire, W. Trender, S. R. Chamberlain, A. E. Jolly, J. E. Grant, F. Patrick, N. Mazibuko, S. C. Williams, J. M. Barnby, P. Hellyer, M. A. Mehta, Cognitive deficits in people who have recovered from COVID-19. *eClinicalMedicine*. **39** (2021).
 25. R. L. Spitzer, K. Kroenke, J. B. W. Williams, B. Löwe, A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. *Archives of Internal Medicine*. **166**, 1092–1097 (2006).

26. R. L. Spitzer, K. Kroenke, J. B. W. Williams, and the Patient Health Questionnaire Primary Care Study Group, Validation and Utility of a Self-report Version of PRIME-MD The PHQ Primary Care Study. *JAMA*. **282**, 1737–1744 (1999).
27. L. Wright, A. Steptoe, D. Fancourt, Patterns of compliance with COVID-19 preventive behaviours: a latent class analysis of 20 000 UK adults. *J Epidemiol Community Health*. **76**, 247–253 (2022).
28. O. Bargain, U. Aminjonov, Trust and compliance to public health policies in times of COVID-19. *Journal of Public Economics*. **192**, 104316 (2020).
29. S. Parsons, R. D. Wiggins, Trust in government and others during the COVID-19 pandemic, 23.
30. S. Rothman, S. Gunturu, P. Korenis, The mental health impact of the COVID-19 epidemic on immigrants and racial and ethnic minorities. *QJM: An International Journal of Medicine*. **113**, 779–782 (2020).
31. D. R. Morales, S. N. Ali, COVID-19 and disparities affecting ethnic minorities. *The Lancet*. **397**, 1684–1685 (2021).
32. D. B. G. Tai, A. Shah, C. A. Doubeni, I. G. Sia, M. L. Wieland, The Disproportionate Impact of COVID-19 on Racial and Ethnic Minorities in the United States. *Clinical Infectious Diseases*. **72**, 703–706 (2021).
33. L. R. Arriola, A. N. Grossman, Ethnic Marginalization and (Non)Compliance in Public Health Emergencies. *The Journal of Politics*. **83**, 807–820 (2021).
34. J. Jaureguizar, I. Redondo, N. Galende, N. Ozamiz, Factors related to compliance with the COVID-19 health regulations among young people. *World J Psychiatry*. **11**, 1247–1258 (2021).
35. N. Georgiou, P. Delfabbro, R. Balzan, COVID-19-related conspiracy beliefs and their relationship with perceived stress and pre-existing conspiracy beliefs. *Personality and Individual Differences*. **166**, 110201 (2020).
36. M. Duplaga, The Determinants of Conspiracy Beliefs Related to the COVID-19 Pandemic in a Nationally Representative Sample of Internet Users. *International Journal of Environmental Research and Public Health*. **17**, 7818 (2020).
37. D. Allington, B. Duffy, S. Wessely, N. Dhavan, J. Rubin, Health-protective behaviour, social media usage and conspiracy belief during the COVID-19 public health emergency. *Psychol. Med*. **51**, 1763–1769 (2021).
38. J. Roozenbeek, C. R. Schneider, S. Dryhurst, J. Kerr, A. L. J. Freeman, G. Recchia, A. M. van der Bles, S. van der Linden, Susceptibility to misinformation about COVID-19 around the world. *Royal Society Open Science*. **7**, 201199.
39. M. Bruder, P. Haffke, N. Neave, N. Nouripanah, R. Imhoff, Measuring Individual Differences in Generic Beliefs in Conspiracy Theories Across Cultures: Conspiracy Mentality Questionnaire. *Frontiers in Psychology*. **4** (2013).
40. A. Goreis, M. Voracek, A Systematic Review and Meta-Analysis of Psychological Research on Conspiracy Beliefs: Field Characteristics, Measurement Instruments, and Associations With Personality Traits. *Frontiers in Psychology*. **10** (2019).

41. M. Grzesiak-Feldman, The Effect of High-Anxiety Situations on Conspiracy Thinking. *Curr Psychol.* **32**, 100–118 (2013).
42. V. Henkel, P. Bussfeld, H.-J. Möller, U. Hegerl, Cognitive-behavioural theories of helplessness/hopelessness: valid models of depression? *European archives of psychiatry and clinical neuroscience.* **252**, 240–249 (2002).
43. W. Cullen, G. Gulati, B. D. Kelly, Mental health in the COVID-19 pandemic. *QJM: An International Journal of Medicine.* **113**, 311–312 (2020).
44. B. Pfefferbaum, C. S. North, Mental Health and the Covid-19 Pandemic. *New England Journal of Medicine.* **383**, 510–512 (2020).
45. R. A. Schmidt, R. Genois, J. Jin, D. Vigo, J. Rehm, B. Rush, The early impact of COVID-19 on the incidence, prevalence, and severity of alcohol use and other drugs: A systematic review. *Drug and Alcohol Dependence.* **228**, 109065 (2021).
46. S. Shahsavari, P. Holur, T. Wang, T. R. Tangherlini, V. Roychowdhury, Conspiracy in the time of corona: automatic detection of emerging COVID-19 conspiracy theories in social media and the news. *J Comput Soc Sc.* **3**, 279–317 (2020).
47. S. Seabold, J. Perktold, in *Statsmodels: Econometric and Statistical Modeling with Python* (Proceedings of the 9th Python in Science Conference, 2010), vol. 57, pp. 92–96.
48. S. Bird, *Natural Language Processing with Python.* O'Reilly Media, Inc, 504 (2009).
49. R. Albalawi, T. H. Yeap, M. Benyoucef, Using Topic Modeling Methods for Short-Text Data: A Comparative Analysis. *Frontiers in Artificial Intelligence.* **3** (2020).
50. D. M. Blei, A. Y. Ng, M. I. Jordan, Latent dirichlet allocation. *Journal of machine Learning research.* **3**, 993–1022 (2003).
51. M. Röder, A. Both, A. Hinneburg, in *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining* (ACM, Shanghai China, 2015), pp. 399–408.
52. S. Syed, M. Spruit, in *2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA)* (2017), pp. 165–174.
53. J. Chang, S. Gerrish, C. Wang, J. Boyd-graber, D. Blei, in *Advances in Neural Information Processing Systems* (Curran Associates, Inc., 2009), vol. 22.
54. R. Rehurek, P. Sojka, Gensim–python framework for vector space modelling. *NLP Centre, Faculty of Informatics, Masaryk University, Brno, Czech Republic.* **3**, 2 (2011).
55. M. D. Hoffman, D. M. Blei, F. Bach, in *Proceedings of the 23rd International Conference on Neural Information Processing Systems - Volume 1* (Curran Associates Inc., Red Hook, NY, USA, 2010), *NIPS'10*, pp. 856–864.
56. V. Braun, V. Clarke, Using thematic analysis in psychology. *Qualitative Research in Psychology.* **3**, 77–101 (2006).

Figures

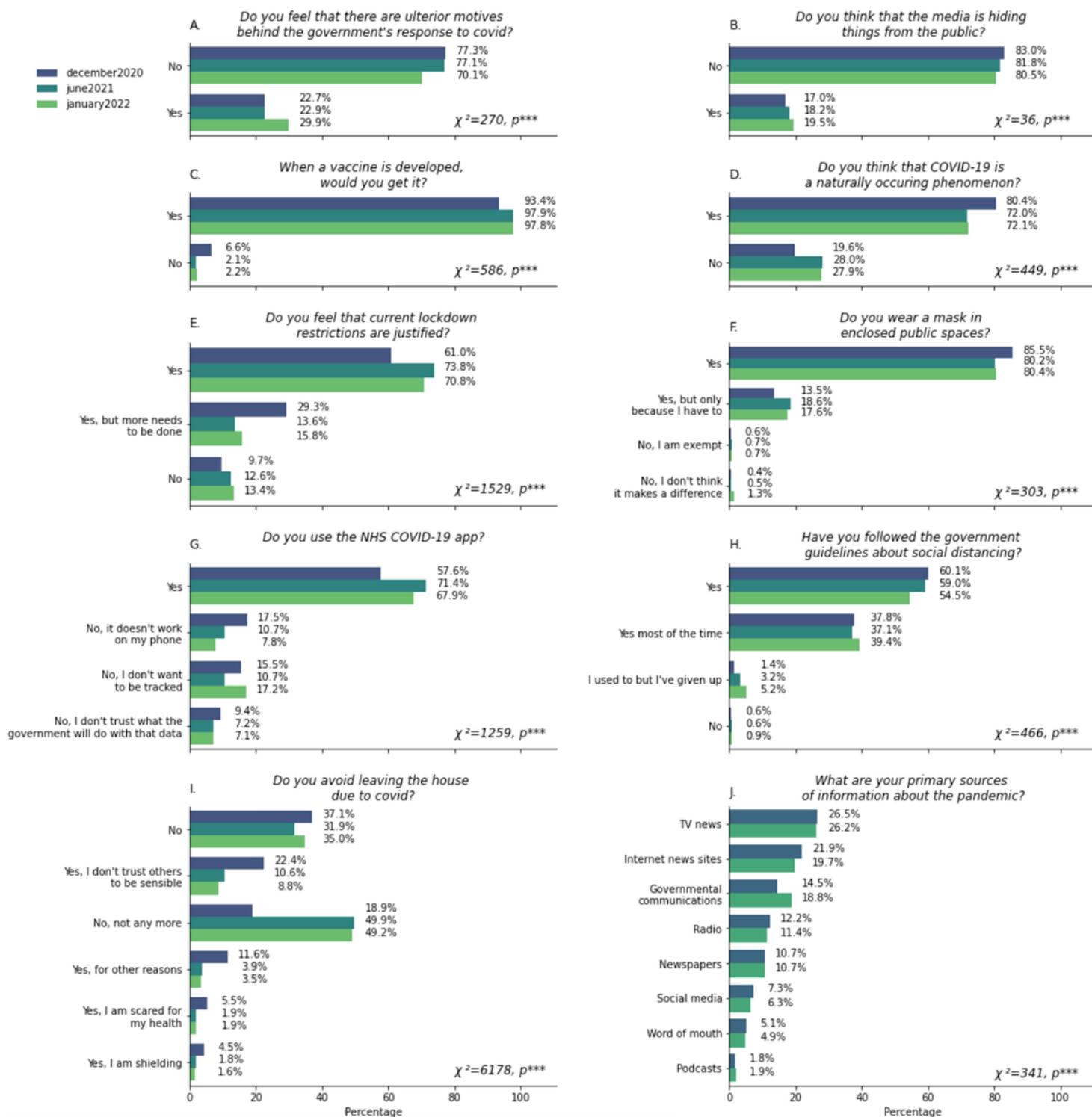


Figure 1

Distribution of trust, compliance and media sources. Panels A-E illustrate the proportions of various responses to trust-class questions. Panels F-I illustrate responses received to questions regarding compliance to measures that attempt to reduce COVID-19 transmission. Figure 1J illustrates the distribution of primary sources about the pandemic that people use. All x-axes show percentages. Each panel contains χ^2 and p values of tests for dissimilarity between the December 2020, June 2021 and

January 2022 data. Answers are displayed in cascade order based on the most prevalent answers given in December 2020. Information regarding sources of information about the pandemic was not collected in January 2022. Significance is denoted as *, **, and *** for $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively.

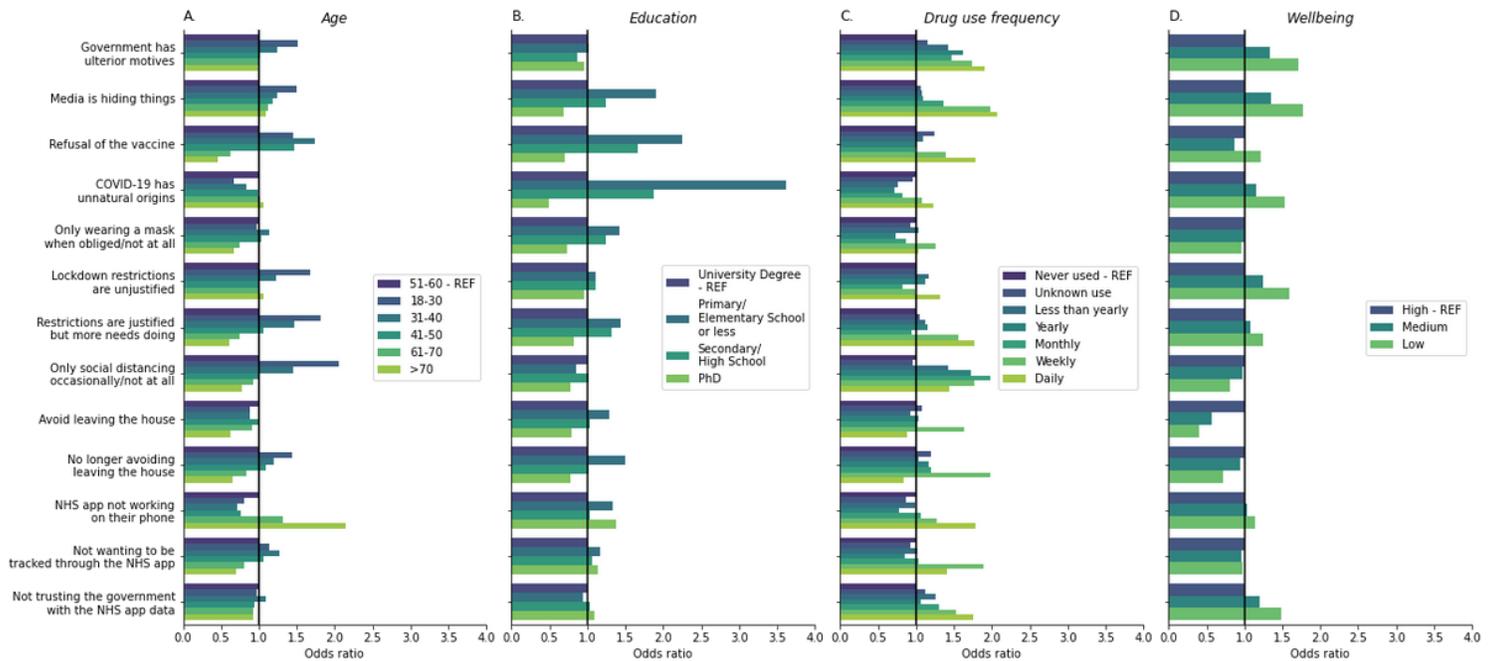


Figure 2

Odds of being distrustful and noncompliant. Panels A-D illustrate the odds ratios of being distrustful for different age groups, ethnicities, drug use levels prior to the COVID-19 pandemic and wellbeing levels recorded in December 2020. On the y-axis answers related to the government response to COVID-19, the media hiding things, vaccination, the origins of COVID-19, wearing a mask and following the social distancing guidelines are resulting from binomial logistic regression models. In the figure legend, categories which act as reference for the odds ratios are labeled “-REF”. Figure for all socio-demographic variables is presented in Supplementary Materials.



Figure 3

Topics and representative opinions per question at different stages of the COVID-19 pandemic. *The different topics that constitute the beliefs of those surveyed are presented alongside their percentage distribution. Detailed topic label descriptions for each of the free text questions can be found in Supplementary Materials.*

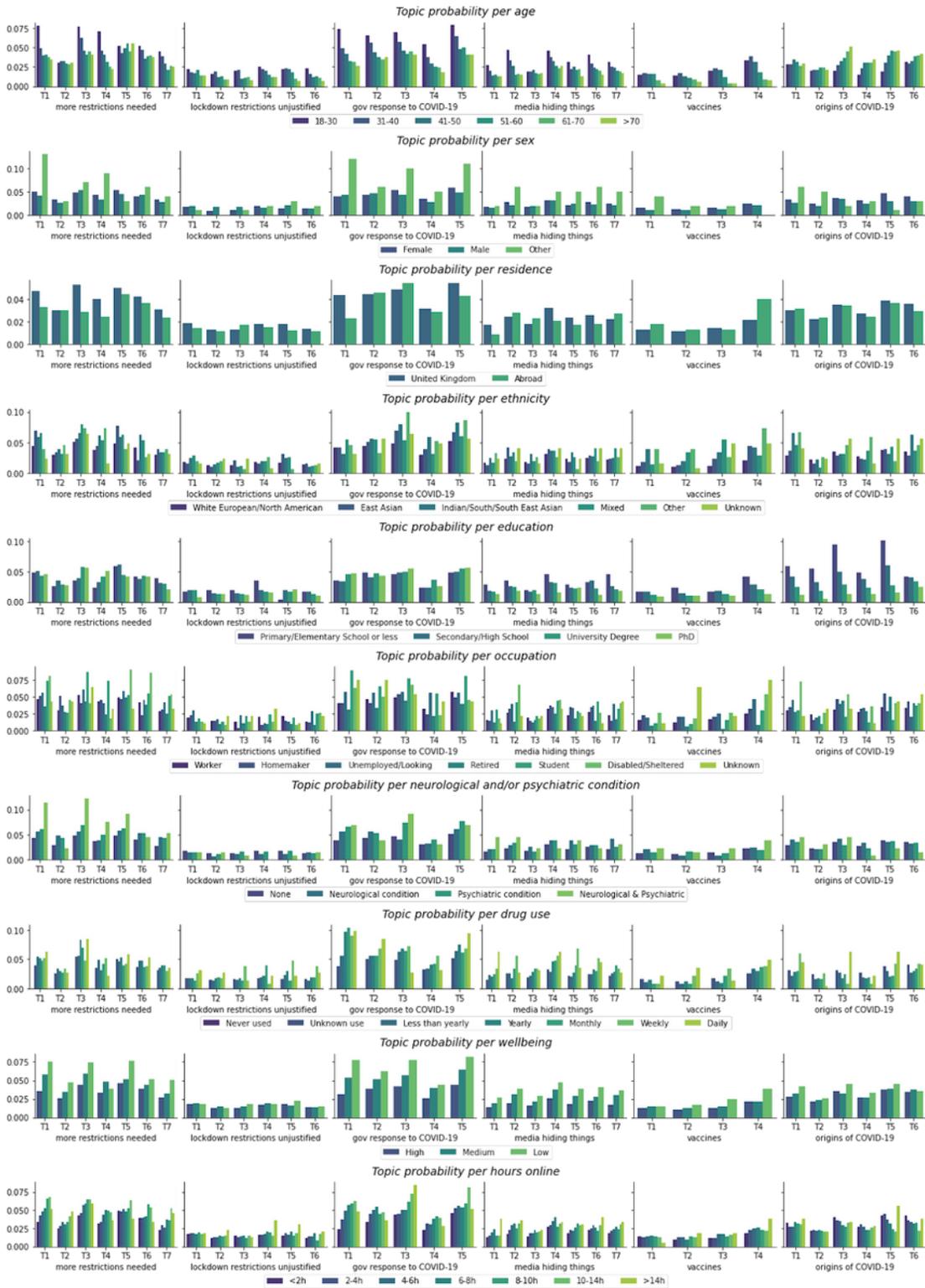


Figure 4

Topic probability varies with sociodemographic variables in December 2020. The probability of a topic (*y* axis) in different levels of sociodemographic factors is presented. Topics are labelled T= Topic. Topic labels in figure 2. Detailed topic label descriptions are in the Supplementary Materials.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [DataS1.csv](#)
- [R1.txt](#)
- [SupplementaryMaterialsMBFinal.docx](#)
- [DataS2.csv](#)
- [TableS6.pdf](#)
- [TableS4.csv](#)
- [TableS5.csv](#)
- [DataS4.csv](#)
- [DataS3.csv](#)