Accepted preprint

McGlashan, M., Clarke, I., Gee, M., Grieshofer, T., Kehoe, A. & Lawson, R. (2025). COVID-19 vaccine conspiracy theories, discourses of liberty, and "the new normal" on social media. *Linguistics Vanguard*. https://doi.org/10.1515/lingvan-2024-0121

COVID-19 vaccine conspiracy theories, discourses of liberty, and 'the new normal' on social media

Abstract

Public distrust in government, pharmaceutical companies, healthcare professions, and medical science and technology has been consistently linked with vaccine rejection. Policymakers, therefore, want to better understand links between distrust of institutions and vaccine refusal.

This paper reports on a case study of posts (tweets) to the social media platform Twitter (now X.com) collected as part of the AHRC-funded TRAC:COVID (**TR**ust **A**nd **C**ommunication: a **C**oronavirus **O**nline **Vi**sual **D**ashboard) project. The TRAC:COVID dashboard combines methods from corpus linguistics with various visualisation techniques to enable users to explore ~84million posts containing reference to COVID-19 published between 1st January 2020 and 30th April 2021 (encompassing the dates of UK coronavirus lockdowns). The dashboard and all sampling considerations (including an overview of the detailed search query used) is available at https://www.traccovid.com.

Specifically, the paper analyses a subsample of posts that make reference to vaccines and at contain at least one hashtag relating to various categories of dis/misinformation. By employing Keyword Co-occurrence Analysis – a method for examining statistically significant keywords using Multiple Correspondence Analysis – we find that these posts draw on various 'discourses of liberty' to protest against perceived infringements on 'health freedoms' through the imposition of new norms of behaviour (e.g. mask wearing).

Keywords

COVID-19, conspiracy theories, fake news, Keyword Co-occurrence Analysis, vaccine hesitancy

1 Vaccine conspiracy theories and dis/misinformation

Vaccine confidence is underpinned by public trust that vaccines work, are safe, and are produced by a trustworthy source. Vaccine confidence thus encompasses "trust in the vaccine (the product), trust in the vaccinator or other health professional (the provider), and trust in those who make the decisions about vaccine provision (the policy-maker)" (Larson et al. 2015). Vaccine hesitancy may emerge as a result of a loss of confidence in vaccines with respect to "(i) the effectiveness and safety of vaccines; (ii) the system that delivers them, including the reliability and competence of the health services and health professionals and (iii) the motivations of policy-makers who decide on the needed vaccines" (MacDonald, 2015: 4162). As such, if public trust in vaccines is undermined, it may influence vaccine confidence and result in vaccine hesitancy.

Public distrust in government, pharmaceutical companies, healthcare professions, and medical science and technology has been consistently linked with vaccine hesitancy and rejection and thus policymakers are keen to understand it (Attwell et al. 2017). Some modern examples of vaccine hesitancy or refusal (often resulting from contact with, or stated belief in, misinformation or rumours) have seen the resurgence of – or inability to inoculate against – life-threatening illnesses. Polio saw a resurgence in Nigeria following boycotts based on, amongst other factors, rumours suggesting the oral polio vaccine (OPV) was linked to cancer, HIV and sterility (Ghinai et al. 2013). Polio remains endemic in Pakistan and Afghanistan due to "parental refusal of polio vaccination, conspiracy theories, and misinformation which have rendered the polio eradication initiatives futile" (Ittefaq et al., 2021: 480). Finally, numerous countries previously free of endemic measles (Albania, Brazil, Czech Republic, Greece, Mongolia, UK, and Venezuela) have reported a resurgence in transmissions (Bozzola et al. 2020; Durrheim 2020). In the case of the UK, Andrew Wakefield's now-debunked research that suggested links between the MMR (Measles, Mumps, and Rubella) vaccine and autism has resulted in widespread anxieties, increasing vaccine refusal and by extension measles outbreaks (Larson et al. 2015).

As noted by Hardaker, et al. (2024: 164), "[r]esistance to vaccination has existed as long as vaccination itself" and this resistance has taken the form of organised protests (and even riots) at the local and national level. The internet has enabled the rapid distribution and consumption of (dis/mis)information about vaccines and made possible the worldwide coordination of anti-vaccination movements, which can serve to undermine public trust in vaccination and result in widespread vaccine refusal. Social media specifically has proven to be a prominent avenue for the distribution of vaccine misinformation (Suarez-Lledo & Alvarez-Galvez 2021).

Related to vaccine dis/misinformation, vaccine conspiracy theories may also serve to undermine confidence in vaccine products, providers, and policy-makers. Drawing on Grieve & Woodfield (2023: 14), we position conspiracy theories here as a form of fake news as they may draw variously on misinformation (Type I Fake News, i.e. unintentionally inaccurate information, e.g. errors), disinformation (Type III Fake News, i.e. true but intentionally deceptive information, e.g. omissions), or both simultaneously (Type II Fake News, i.e. intentionally false news, i.e. lies). Conspiracy theory beliefs may be premised on little access to (or ignorance of) information that is either veracious or honest but represent a rational attempt to understand complex phenomena and deal with feelings of powerlessness, for example. Studying the effects of anti-vaccine conspiracy theories on intentions to vaccinate, Jolley & Douglas (2014: 6) have found that "anti-vaccine conspiracy theories may have more than a trivial effect on vaccination intentions" and that "anti-vaccine conspiracy theories appear to introduce undue suspicion about vaccine safety, and increase feelings of powerlessness and disillusionment, whilst decreasing trust in authorities, which in turn introduce reluctance to vaccinate". Anti-vaccination misinformation and conspiracy theories feature a range of tropes, including that vaccines are: toxic (containing foreign DNA, aborted fetal tissue, or formaldehyde), thus harmful (citation of previously harmful vaccines/medical treatments such as Thalidomide is common) and linked to autism (Kata 2012); linked to genetic modification (Lyons, Merola & Reifler

2019) and DNA alteration[1]; and a means of population control through sterilisation[2] or even genocide.[3] Underpinning these tropes in anti-vaccination arguments runs a general distrust of science and government but, particularly, pharmaceutical companies who are accused of bribing researchers "to fake their data, cover up evidence of the harmful side effects of vaccines, and inflate statistics on vaccine efficacy" for financial gain or some other sinister motive (Jolley & Douglas, 2014: 1).

The COVID-19 pandemic presented a range of social, medical, and political challenges stemming from the rapid generation and spread of both novel and recontextualised anti-vaccination tropes and conspiracy theories, such as the idea that COVID-19 was created as a biological weapon in China (Pennycook et al. 2020) or that COVID-19 is a false pandemic engineered to administer vaccines containing microchips to track vaccine recipients.[4] Given that conspiracy theories are inherently social phenomena and that their meanings are situated in specific social contexts (Bergmann et al., 2020), the aim of this paper is to explore vaccine-related conspiracy theories in the context of the COVID-19 pandemic. One way to gain a better understanding of social phenomena is to explore language used about them. Given the increasingly important role of social media for our social lives, especially during the lockdowns of COVID-19, this paper investigates tweets about vaccination posted during the COVID-19 pandemic that include hashtags that make reference to various forms of dis/misinformation (including conspiracy theories). It should be noted that while these tweets contribute to narratives specific to various conspiracy theories, they do not necessarily constitute misinformation or disinformation.

2 Data: TRAC:COVID

The data are drawn from TRAC:COVID (**Tr**ust **A**nd **C**ommunication: a **C**oronavirus **O**nline **VI**sual **D**ashboard; Kehoe et al., 2021; see www.traccovid.com). The primary output of the TRAC:COVID project is a freely accessible online dashboard that combines tools from corpus linguistics with various visualisation techniques and enables users to query a corpus of 84,138,394 tweets (i.e. posts on the social media platform Twitter but now known as X) containing reference to COVID-19. Only tweets published in the English language and in Great Britain between 1st January 2020 and 30th April 2021 were collected. TRAC:COVID is intended for use by a wide audience, including members of the general public, researchers, policy makers, and journalists to better understand how social media was used during the pandemic to talk about COVID-19. TRAC:COVID, therefore, facilitates exploration of language use in Britain during the COVID-19 pandemic.

Although TRAC:COVID can be used to explore myriad linguistic and social phenomena as realised in tweets about COVID-19 posted during the pandemic, this paper concentrates specifically on tweets discussing vaccines and conspiracy theories as a way to understand vaccine-related conspiracy theories. This paper builds upon the work of McGlashan et al. (2021), which inductively identified 276 dis/misinformation-related hashtags in tweets containing reference to vaccines from the underlying data of TRAC:COVID. As such, where McGlashan et al. (2021) identified the scale and variety of these hashtags, this paper investigates in more detail the contents of those tweets that contain both reference to vaccines and dis/misinformation (via these hashtags).

[1] https://www.reuters.com/article/uk-factcheck-viral-post-idUSKBN28S2V1
[2] https://jitsuvax.info/conspiracist-ideation/population-control/
[3] https://www.genocidewatch.com/single-post/genocide-watch-rejects-conspiracy-theories-about-covid-19-vaccines
[4] https://www.bbc.co.uk/news/54893437

Specifically, we concentrated on tweets and retweets that include the terms *vaccin* OR *vax*. The plots below, visualisations from www.traccovid.com for the searches “vaccine” and “vax”, show that use of both terms began to increase in frequency of use per day after 9th November 2020. Increased frequency would suggest an increased interest.

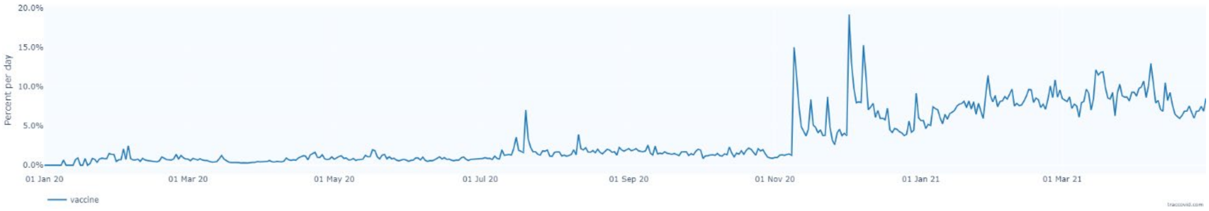

Figure 1: TRAC:COVID timeseries: percentage of tweets per day including the term “vaccine”

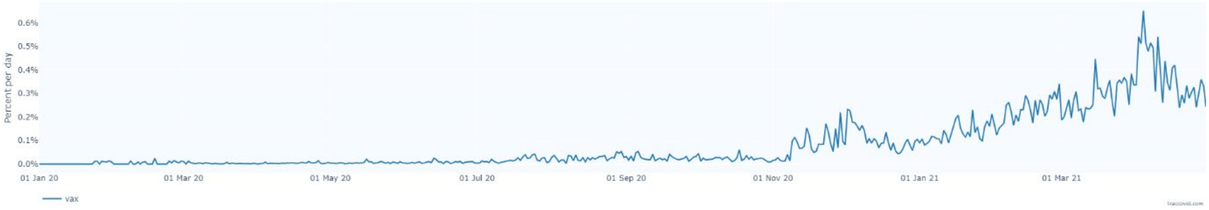

Figure 2: TRAC:COVID timeseries: percentage of tweets per day including the term “vax”

In terms of the size of our corpus (Table 1), there were 2,658,495 tweets and retweets containing the terms *vaccin* OR *vax* but duplicate tweets accounted for 22.63% of the corpus. Removing duplicates left 2,056,747 remaining and, following removal of retweets (which further add to duplication in the corpus), the corpus comprised 1,499,885 unique tweets (or 56.42% of the initial corpus of 2,658,495 tweets and retweets). Unique refers here to tweets that are not total duplicates of any other tweet in the corpus.

	Raw	Unique	Unique (as % of Raw)
Retweets	892,603	556,862	62.39%
Tweets	1,445,191	1,272,433	88.05%
Quote Tweets	230,701	227,452	98.59%
Total	2,658,495	2,056,747	77.37%
Total (minus Retweets)	1,675,892	1,499,885	89.50%

Table 1: corpus size

Tweets in our corpus were then tagged based on their inclusion of hashtags relating to six areas of COVID-19 dis/misinformation as identified by McGlashan, et al. (2021; Table 2). Our final corpus comprises 35,956 unique tweets containing at least one hashtag relating these categories of dis/misinformation (Table 3).

Category	Subcategory	Example hashtags
Anti	Anti-lockdown	#notolockdown, #antilockdown
	Anti-mask	#antimask, #masksoff
	Anti-vaccine passport	#novaccinepassports, #vaccinepassport
	Anti-vax	#antivaxx, #novaccine
Conspiracy theories	Evil	#evil, #markofthebeast
	Financial corruption	#followthemoney, #kerching

	Flat earth	#flatearth, #flatearthers
	Freedom and censorship	#idonotconsent, #keepbritainfree
	General	#propaganda, #conspiracy
	Human rights and crimes against humanity	#humanrights, #depopulation
	Known individuals/conspiracy theorists	#alexjones, #davidicke
	Media	#msm, #fakenews
	New World Order	#nwo, #oneworldgovernment
COVID-19	Cure	#hydroxychloroquineworks, #ivermectin
	Fake/planned	#covidhoax, #plandemic
Pharma	Corruption	#bigpharma, #nhscorruption
	Vaccine	#vaccinemafia, #autism
Politics	Protest	#fightback2020, #resist
	QANON	#qanon, #wwg1wga
	Trump	#maga, #maga2020
Science and technology	Microsoft/Bill Gates	#Billgates, #billgatesbioterrorist
	Technology	#5g, #4ir

Table 2: categories of dis/misinformation hashtag in COVID-19 vaccination tweets

	Raw				Unique			
Category	tweets	quote tweets	Total	%	tweets	quote tweets	Total	%
Anti	19,953	6,072	26,025	27.46%	10,827	4,439	15,266	35.73%
Conspiracy	19,384	4,745	24,129	25.46%	8,979	2,772	11,751	27.50%
COVID-19	11,997	2,372	14,369	15.16%	3,864	1,229	5,093	11.92%
Pharma	6,608	2,060	8,668	9.15%	4,425	1,283	5,708	13.36%
Politics	2,557	1,146	3,703	3.91%	1,517	626	2,143	5.02%
Science and technology	13,816	4,055	17,871	18.86%	1,886	882	2,768	6.48%
Total	74,315	20,450	94,765	100.00%	31,498	11,231	42,729	100.00%
Total (accounting for tweets that belong to 2 or more hashtag categories	42,781	11,515	54,296	57.30%	27,274	8,682	35,956	84.15%

Table 3: number of tweets containing dis/misinformation-related hashtags

3 Methods: Keyword Co-occurrence Analysis (KCA)

The principal method employed here is KCA, which, as described in Clarke (2023) and Sha and Clarke (2025), is a method for examining keywords – lexical items (types) that are found to occur with an unusual *token* (in)frequency when comparisons are made between a target corpus and a reference corpus – by using Multiple Correspondence Analysis (MCA). MCA is a geometric data analytic method which identifies relationships between three or more categorical variables. In the context of KCA, these categorical variables are a binary measure of whether a keyword is *present* or *absent* in the texts of a corpus and MCA is used to identify keywords that co-occur often in the texts of the corpus.

Sha & Clarke (2025) outline the following four step approach for carrying out KCA, with the first three steps representing methodological procedures to enable later analysis (Step 4):

Step 1. Compute keywords using a traditional keyword analysis
Step 2. Analyse each text in the corpus for the occurrence of these keywords and record in a categorical data matrix (Table 4)
Step 3. Subject the data matrix to MCA to reveal dimensions comprising the most common patterns of co-occurring keywords, and finally
Step 4. Interpret these dimensions of keyword co-occurrence, guided by the principles of linguistic co-occurrence (Biber 1988) and the indicative nature of keywords in discourse (Baker 2023).

As such, Step 1 involved producing a keyword list for our target corpus by comparing tweet text (minus #hashtags and @mentions) against the same Twitter reference corpus used to generate keywords on the TRAC:COVID platform. Keywords were identified using Log-Likelihood as a test for significance and because infrequent and overly frequent features can adversely affect MCA (Step 3), we only considered keywords that occurred within at least 1% (≥ 359.56) and fewer than 95% (≤ 34,158.2) of all 35,956 unique tweets in our corpus. This approach produced a list of 175 keywords.

In line with step 2, using this keyword list, we produced a data matrix that records whether the keywords identified in Step 1 are present (P) and/or absent (A) in each tweet in our corpus and also appended some supplementary information to the matrix. We specified qualitative supplementary variables (quali_sup in Table 4) indicating whether dis/misinformation hashtags occurred in these tweets and which category/categories of COVID-19 dis/misinformation (Table 2) these hashtags belonged to, as well as information about the length of each tweet as a quantitative supplementary variable (quanti_sup in Table 4). Table 4 gives some dummy data to visualise the structure of this matrix.

	Keywords			Quali_sup						Quanti_sup
Tweet	KW1	KW2	KW…	Anti	Conspiracy	Covid-19	Pharma	Politics	Science and technology	Text length
1	A	P	A	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	29
2	P	P	A	FALSE	FALSE	TRUE	TRUE	FALSE	FALSE	48
3	A	A	A	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	32

Table 4: dummy data matrix

The third and final methodological step was to subject this matrix to MCA using the 'FactoMineR' package in R (Lê, Josse & Husson 2008) to identify dimensions comprising the most common patterns of keyword co-occurrence across the corpus. The MCA assigned each category of keyword (present or absent) a score (in percentage) for how much they contribute to a dimension. Categories of keywords that make stronger contributions to a dimension therefore account for a larger share of the total percentage of contributions to a dimension. Our analysis below focusses on only those categories of keywords that made the strongest contributions to a dimension (and therefore are most representative of the patterns of variation within a dimension; Le Roux & Rouanet, 2010: 52) by considering those with a contribution score equal to or greater than the average expected contribution for a dimension. This was achieved by using the following equation, which divides the total possible contribution to a dimension (100%) by the total number of categories of keywords in our feature set. Given that we have 175 keywords and 2 possible categories of keyword, the average contribution score is calculated as follows:

$$\frac{100}{175 \times 2} = 0.2857143$$

Furthermore, keywords are also assigned coordinates, which show how closely associated keywords are to each other based on their distribution in texts; "keywords with strong contributions and positive coordinates co-occur often together in many texts, while keywords with strong contributions and negative coordinates co-occur often together in a different set of texts" (Sha & Clarke 2025). For more detail on MCA for identifying dimensions, see Sha & Clarke in this issue and Clarke et al. (2021).

Concerning Step 4, which focuses on the interpretation of dimensions to explain why keywords co-occur across many tweets, our approach draws broadly from theory and methods in corpus linguistics, discourse analysis, and Corpus-Assisted Discourse Studies (CADS). To wit, we concentrate on keywords so as to identify lexical items that are both distinctive to and generalisable across (Egbert & Biber 2019; McGlashan & Krendel 2023) the texts that make up our target corpus. In other words, keywords enable us to identify language use that is statistically salient within and across – and therefore typical of – the texts that make up our corpus. Furthermore, our consideration of keyword co-occurrence enables us to operationalise the conception of discourse as "a cluster of context-dependent semiotic practices that are situated within specific fields of social action" (Reisigl & Wodak, 2016: 27) and to see how specific dimensions (and potentially discourses) are revealed through the co-occurrence of keywords that also contain reference to vaccines *and* dis/misinformation-related hashtags.

4 Findings

Due to limitations of space, this analysis concentrates on the first dimension of our KCA, which, using Benzécri's (1992) modified rates, explains 23% of the variance in the dataset. The 94 keywords contributing most strongly to Dimension 1 are presented in Table 5. Unlike traditional keyword analysis, wherein a specified number of keywords – or set of keywords – would be isolated from a larger keyword list for independent analysis, KCA is interested in how keywords co-occur as part of larger functional constructs and so keywords need to be interpreted in the context of the texts in which they occur to understand how they function within those constructs.

Keyword	Dim_1_coordinate	Dim_1_contribution	Positive/Negative
normal_P	2.0576229	2.5647752	Positive
back_P	1.6010858	3.0955408	
lives_P	1.4911087	1.4033958	
still_P	1.3865033	3.018071	
every_P	1.3751706	1.3529647	
life_P	1.2727009	1.4361035	
masks_P	1.2528389	1.0054154	
months_P	1.2359167	0.9049022	
believe_P	1.1487665	0.922964	
virus_P	1.1471942	2.192863	
year_P	1.1288024	1.2063926	
got_P	1.112826	0.8263702	
flu_P	1.1104524	1.1180189	
last_P	1.0985308	0.7762277	

go_P	1.0984755	1.4893679	
years_P	1.0967878	0.7834183	
even_P	1.0959157	1.4984956	
want_P	1.0908006	1.9109675	
spread_P	1.0900985	0.6483525	
govt_P	1.0827813	0.6239993	
something_P	1.022433	0.5130448	
think_P	1.0210395	1.5293655	
safe_P	1.0172692	0.8206274	
live_P	1.0065746	0.5121572	
control_P	1.0031533	0.8827912	
let_P	1.0023341	0.5427826	
get_P	0.9957288	3.0495065	
tests_P	0.9954591	0.4889819	
give_P	0.9906465	0.5525086	
without_P	0.9836244	0.6559732	
many_P	0.9816603	1.3788778	
enough_P	0.9801621	0.5768484	
getting_P	0.9627639	0.8329681	
fear_P	0.9500965	0.50941	
yet_P	0.9477657	0.885297	
never_P	0.9422314	1.0067706	
can_P	0.9328018	2.7576975	
vax_P	0.9236201	0.6582368	
need_P	0.9235772	1.9745267	
going_P	0.9031421	1.2052968	
keep_P	0.903124	0.5639036	
thing_P	0.9014364	0.4009739	
make_P	0.8989983	1.0883454	
way_P	0.8942237	0.8244538	
remember_P	0.8731155	0.3700578	
tell_P	0.8715892	0.3880669	
vaccinated_P	0.8613108	1.348214	
us_P	0.855376	1.8786027	
take_P	0.8550116	1.4331693	
people_P	0.8518871	4.0332878	
stop_P	0.8496569	0.9682949	
put_P	0.8300973	0.3317264	
country_P	0.8200349	0.3372217	
ever_P	0.8159401	0.3872803	
social_P	0.8051942	0.4040236	
everyone_P	0.8049456	0.6472518	
risk_P	0.8029194	0.4603676	
population_P	0.7899786	0.4156038	
time_P	0.7824057	0.8693768	
testing_P	0.7806247	0.4017455	

long_P	0.7547259	0.3343972	
around_P	0.7447443	0.3122598	
just_P	0.733271	1.5545485	
jab_P	0.7288697	0.4710129	
used_P	0.7264108	0.3725788	
immunity_P	0.7216591	0.30713	
part_P	0.7183585	0.2981163	
one_P	0.7102265	1.0705087	
work_P	0.6995905	0.5039623	
like_P	0.6980501	1.0504075	
travel_P	0.6837583	0.3413638	
really_P	0.682389	0.4309131	
test_P	0.6790896	0.4125721	
end_P	0.6748093	0.4049518	
know_P	0.6741901	0.7379092	
much_P	0.6733227	0.3534554	
anyone_P	0.6711502	0.3885248	
money_P	0.6649717	0.2980276	
now_P	0.6645651	1.4558328	
lockdown_P	0.6284113	0.3569884	
freedom_P	0.6272749	0.3372822	
right_P	0.6235767	0.4471958	
say_P	0.6066394	0.4931153	
passport_P	0.5809249	0.4138366	
see_P	0.5384966	0.438967	
deaths_P	0.5165496	0.3018659	
amp_P	0.510877	1.590491	
world_P	0.4401124	0.3217128	
vaccine_P	0.2122467	0.6345317	
people_A	-0.1113347	0.5271177	Negative
read_P	-0.6038742	0.3886609	
gates_P	-0.6576466	0.6257948	
bill_P	-0.7260394	0.6591004	
via_P	-0.7947785	0.9519972	

Table 5: Dimension 1 keywords (ranked by coordinate value)

Dimension 1 distinguishes between texts on the negative side of the dimension, which include *specific* forms of COVID-19 dis/misinformation and conspiracy theories, and texts on the positive side of the dimension, which include more *general* concerns about freedom, control, and 'normality' stemming from ongoing (and future) interventions to curtail the spread of COVID-19.

Texts associated with negative Dimension 1 make (and intend to signal/boost/spread) *specific* forms of COVID-19 disinformation, misinformation and conspiracy theory (including reference to specific individuals, external links), as exemplified in the text most strongly associated with negative Dimension 1 (Examples 1Example 1.1). Many of the texts associated with negative Dimension 1 promote conspiracies about Bill Gates' role in the COVID-19 pandemic. The conspiracies suggest that Bill Gates (along with other elite individuals) orchestrated the COVID-19 pandemic to simultaneously profit from the manufacture of vaccines, and harm and control the global population through adulteration of these vaccines to include 'micro chips' and DNA-altering substances. Such tweets often share supporting content, including URLs to videos, images, links to additional content, as a way to add evidence to their claims.

Example	Coord	Contrib	Text
1.1	-0.292	-0.019	**Bill Gates** (**Dr**. Evil) to Address Forty Heads of State at Climate Summit! **Watch** The Full **Video** Here. [URL REDACTED] #COVID19 #CovidVaccine **#MRNA**

Examples 1: text most strongly associated to negative Dimension 1 (keywords emphasised)[5]

By contrast, texts associated with positive Dimension 1 draw from an overarching 'discourse of liberty', which 'shelters' a range of 'subordinate' discourses (Sunderland, 2004: 69) wherein liberty is construed in relation to perceived norms of freedom and restriction in the specific context of the COVID-19 pandemic.

These discourses articulate a variety of perspectives on the tensions between negative liberty (i.e. the desire to be free from outside interference; freedom from imposition) and positive liberty (i.e. the ability to self-determine; freedom to act without imposition) evaluated (thus, ideologically construed) in terms of what is (un)desirably "normal".

Some texts (as exemplified in Examples 2) include a 'discourse of delayed gratification' in which impositions on negative liberty (e.g. using covid passports, taking vaccines, taking tests, #jabmeup) are accepted despite expressing immediate desires for positive liberty ('getting back to normal', seeing and hugging family). The achievement of positive liberty ("normal") is therefore construed as possible but only through temporary acceptance of some imposition on negative liberty.

Example	Coord	Contrib	text
2.1	0.69	0.11	I'm not scared, **just want get back** to **normal life** ASAP and **travel** to **see** my family. **#VaccinePassports** will be **used** in **countries around** the **world**, if not **already**, and #COVID19 isn't **going** away any **time** soon. Plus not exactly **new**, I've **already got** Yellow Fever Certificate. #jabmeup
2.2	0.53	0.066	I will happily **take two coronavirus tests** a **week** if it means I **get** to **see** and hug my family. I **also** can't wait to **get** my **vaccine**! I **want** my **life** to **get back** to **normal**, but I **want** it to be **safe**. #COVID **#CovidPassport** #Covid19UK

Examples 2: Discourse of liberty - delayed gratification (keywords emphasised)

More common in positive Dimension 1, however, are texts that bring into question any COVID-19 prevention and containment measures. For example, texts in Examples 3 illustrate what we interpret to represent a 'discourse of barriers to positive liberty' wherein the existence of COVID-19 is not questioned but measures that cause any imposition on freedom to act (e.g. lockdown, social distancing) are construed as ineffective. Any concession of negative liberty is viewed with suspicion. And so, perceptions of what might be ideally "normal" here are those that favour positive liberty.

Example	Coord	Contrib	text

[5] Example text in **bold** indicates either:

1) A keyword identified during Step 1 of our methods.
   or
2) A covid misinformation hashtag (as identified by McGlashan, et al. 2021) used in the sampling of the texts in our corpus.

3.1	0.541	0.068	Hang on, I thought masks worked? And then we could 'Cry **freedom'** after our #vaccine? Now we **need #CovidPassport** and #MassTesting - does the **vaccine** actually **work**, and if so, why all of this nonsense to **stop us ever really** getting **back** to **normal**? **#NoVaccinePassports**
3.2	0.531	0.066	If lockdowns **work** If **social** distancing works If hand washing works If masks **work** If vaccines work Why are they so reluctant to **let everyone go back** to "**normal**"? It's almost **like** this has **never** been about **Covid** and all about **control**! **#EndTheLockdown** #COVID19
3.3	0.53	0.065	Lockdowns are pointless as once restrictions lifted it comes **back**, **like** with the **flu** we have to learn to **live** with this and **get** on. They **may never** be a covid19 **vaccine**, but the **flu vaccine** has been **around** for over 70 **years yet** it's **still** very **much around #nolockdown** #lockdownUK

Examples 3: Discourse of liberty – barriers to positive liberty

More radical scepticism is seen throughout texts associated with positive Dimension 1 and include outright rejection of prevention and containment measures and what we interpret to be a more radical 'discourse of libertarianism'. Examples range from texts that include some general reference to (and the rejection of) government/state control enabled by COVID-19 prevention and containment measures (Examples 4) to texts that express fears of coercive government/state control (Examples 5). Within such texts, COVID-19 measures are construed as direct acts of control rather than a byproduct of COVID-19 that enable governments/states to pursue more insidious agendas such as eugenics programmes and population control (Example 5.1), and government/state tyranny (Examples 5.2, 5.3, 5.4). Here, articulation of the 'discourse of libertarianism' is achieved through direct references to "control" (Examples 5.1, 5.2) and "permanent control over our liberties" (Example 5.2), "#tyranny" (Example 5.3), and intertextual reference to George Orwell's *Nineteen Eighty-Four* – a dystopian novel about government totalitarianism – which is recontextualised to the COVID-19 context through the hashtag "#COVID1984" (Example 5.4).

Example	Coord	Contrib	text
4.1	0.48	0.053	I so wish this was true. Right **now** - and for the past **year** - the **government** have owned and controlled **us**. We **must get** all of our freedoms **back**. No **new normal**. Just **normal life**. #Covid19 #nomoremasks **#NoVaccinePassportsAnywhere**
4.2	0.47	0.052	Tell **us something** we don't **know**! Time to #GetAGrip Start acting **like** grown adults &**amp**; learning to **live** with this common cold **virus**. "**Vaccines**" won't **stop** it. House arrest &**amp**; shutting everything down was **never** the policy for Influenza nor should it be **now** with #WuFlu **#EndTheLockdown**

Examples 4: Discourse of liberty – libertarianism: rejection of state control

Example	Coord	Contrib	text
5.1	0.57	0.076	The "**pandemic**" has been over for some **time**. The only reason **lockdown** continues is to **control** the **population** to **get us vaccinated** (depopulation/eugenics). **Even** then the **vax** doesn't **stop** infection or transmission of #Covid. **Take back** your **life**. **#scamdemic**

5.2	0.56	0.073	Bravo This foot dragging is about **getting** more **people vaccinated** so that they can justify permanent **control** over our liberties in the form of internal an **#VaccinePassport** - the **last thing** they **want** is **people** ignoring them, ripping off **masks** and **getting back** to **normal**.
5.3	0.52	0.063	The #Covid19 **virus** #masks #lockdown & #vaccine are the tool **used** to **CONTROL US** & **make us** OBEDIENT so they **can** implement **#TheGreatReset** revolution! If U **still think** this **#tyranny** is to save **lives**? then you've **got** ur eyes closed OR U are **part** of their **#Bilderberg** BackBetter plans.
5.4	0.499	0.058	Lockdowns are **going** to continue until **people** are begging for **testing**, **immunity passports** or a **vaccine** under the false pretences of **getting** their **life back** to "**normal**." Truth doesn't matter, all that matters is that **people believe** it to be true - **already** successful **#COVID1984**

Examples 5: Discourse of liberty – libertarianism: coercive government/state control

Though brief, this analysis of a specific subset of COVID-19 tweets containing reference to both vaccines and dis/misinformation-related hashtags identifies an overarching 'discourse of liberty' and several subordinate discourses of 'delayed gratification', 'barriers to positive liberty', and 'libertarianism' through which 'normality' is construed most typically as a life free from government imposition (here, a range of COVID-19 prevention and containment measures). Reading these findings together with a summary of how frequently different categories of dis/misinformation hashtags are used in texts associated with positive Dimension 1 (Figure 3), it appears that these tweets serve to function as forms of visible (and searchable) protest that are informed by a range of conspiracy theories; the hashtag categories *Anti* and *Conspiracy theories* account for 63.5% (36.71% and 26.79% respectively) of the total number of hashtags in positive Dimension 1 texts.

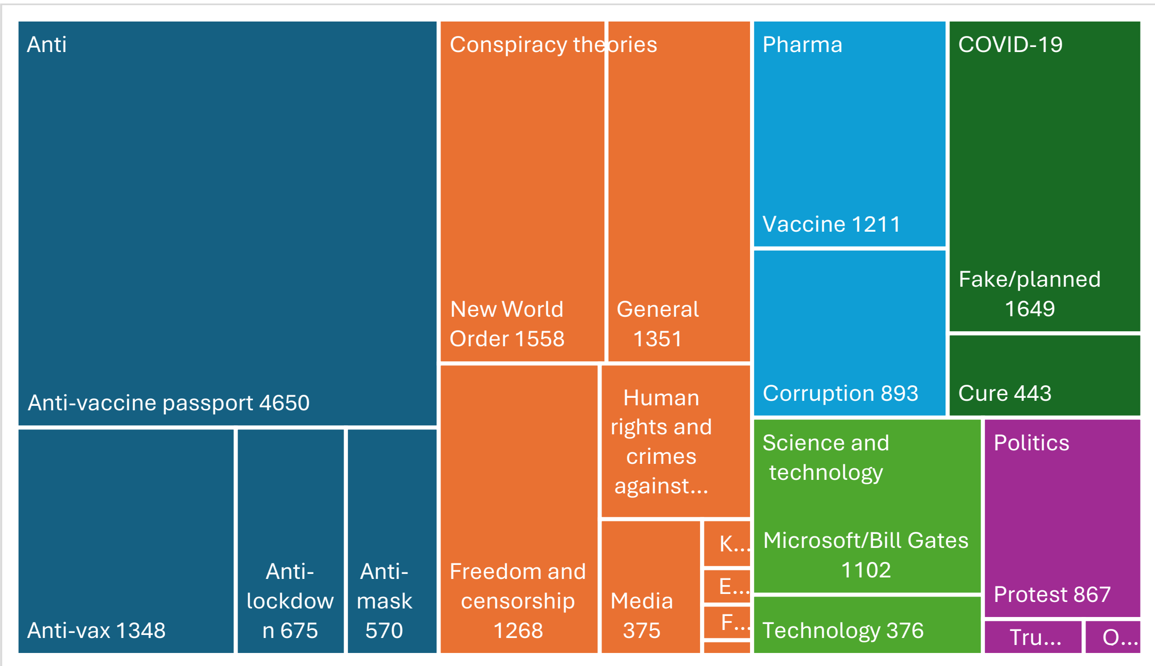

Figure 3: treemap visualising hierarchically (left-to-right) the frequency of occurrence of each of the COVID-19 dis/misinformation hashtag categories (described in Table 2) as they occur in texts associated with positive Dimension 1

Overall, Dimension 1 indicates that the texts in the corpus most often vary according to those which present Bill Gates as the villain protagonist in the COVID-19 pandemic and those which call for freedom. This dimension thus represents a continuum of blame: on the one hand, blame is put squarely on Bill Gates as *the* root cause for violating 'health freedoms', whereas on the other hand, multiple conspiracy theories are co-present to argue for 'health freedoms' by questioning the effect of restrictions and their legitimacy. In other words, where the target for blame on the positive side of the dimension can be (and is in our data) more nebulous, the target for blame on the negative side is clear and absolute.

5 Contributions

This study concentrated on a very small subset of the total number of tweets that make up the TRAC:COVID corpus (35,956/84,138,394 = 0.04%), which include both references to vaccination (*vax* and *vaccine*) and dis/misinformation hashtags. Any findings, therefore, are not generalisable beyond this restricted sample. However, this study found that tweets posted during the COVID-19 pandemic containing both dis/misinformation hashtags and reference to vaccines draw on 'discourses of liberty' in order to protest against the perceived imposition of new norms of behaviour (e.g. mask wearing), which are viewed with suspicion and as overreaching. In doing so, these tweets function to both reject any infringement on 'health freedoms' as well as to identify (through drawing on a range of conspiracy theories) and apportion blame to some cause(s) of these infringements. These findings suggest that changes to perceived norms (especially those that may impede on subjective understandings of personal liberty) may be perceived as threats and, thus, these findings present a challenge for managing change, whether in this specific context or not.

Funding statement

This research was made possible by funding from the AHRC grant number AH/V012630/1 and the Leverhulme Trust ECF-2020-590.

References

Attwell, Katie, Julie Leask, Samantha B. Meyer, Philippa Rokkas & Paul Ward. 2017. Vaccine Rejecting Parents' Engagement With Expert Systems That Inform Vaccination Programs. *Journal of Bioethical Inquiry* 14(1). 65–76. https://doi.org/10.1007/s11673-016-9756-7.

Baker, Paul. 2023. *Using corpora in discourse analysis* (Bloomsbury Discourse). Second edition. London: Bloomsbury Academic.

Biber, Douglas. 1988. *Variation across Speech and Writing*. 1st edn. Cambridge University Press. https://doi.org/10.1017/CBO9780511621024.

Bozzola, Elena, Giulia Spina, Alberto Eugenio Tozzi & Alberto Villani. 2020. Global Measles Epidemic Risk: Current Perspectives on the Growing Need for Implementing Digital Communication Strategies. *Risk Management and Healthcare Policy* Volume 13. 2819–2826. https://doi.org/10.2147/RMHP.S201279.

Clarke, Isobelle. 2023. The Discourses of Climate Change Denialism across Conspiracy and Pseudoscience Websites. In *The Routledge Handbook of Discourse and Disinformation*, 325–341. 1st edn. London: Routledge. https://doi.org/10.4324/9781003224495-24.

Clarke, Isobelle, Tony McEnery & Gavin Brookes. 2021. Multiple Correspondence Analysis, newspaper discourse and subregister: A case study of discourses of Islam in the British press. *Register Studies* 3(1). 144–171. https://doi.org/10.1075/rs.20024.cla.

Durrheim, David N. 2020. Measles eradication—retreating is not an option. *The Lancet Infectious Diseases* 20(6). e138–e141. https://doi.org/10.1016/S1473-3099(20)30052-9.
Egbert, Jesse & Doug Biber. 2019. Incorporating text dispersion into keyword analyses. *Corpora* 14(1). 77–104. https://doi.org/10.3366/cor.2019.0162.
Ghinai, Isaac, Chris Willott, Ibrahim Dadari & Heidi J. Larson. 2013. Listening to the rumours: What the northern Nigeria polio vaccine boycott can tell us ten years on. *Global Public Health* 8(10). 1138–1150. https://doi.org/10.1080/17441692.2013.859720.
Grieve, Jack & Helena Woodfield. 2023. *The Language of Fake News*. 1st edn. Cambridge University Press. https://doi.org/10.1017/9781009349161.
Hardaker, Claire, Alice Deignan, Elena Semino, Tara Coltman-Patel, William Dance, Zsófia Demjén, Chris Sanderson & Derek Gatherer. 2024. The Victorian anti-vaccination discourse corpus (VicVaDis): construction and exploration. *Digital Scholarship in the Humanities* 39(1). 162–174. https://doi.org/10.1093/llc/fqad075.
Ittefaq, Muhammad, Annalise Baines, Mauryne Abwao, Sayyed Fawad Ali Shah & Tayyab Ramzan. 2021. "Does Pakistan still have polio cases?": Exploring discussions on polio and polio vaccine in online news comments in Pakistan. *Vaccine* 39(3). 480–486. https://doi.org/10.1016/j.vaccine.2020.12.039.
Jolley, Daniel & Karen M. Douglas. 2014. The Effects of Anti-Vaccine Conspiracy Theories on Vaccination Intentions. (Ed.) Ralph Tripp. *PLoS ONE* 9(2). e89177. https://doi.org/10.1371/journal.pone.0089177.
Kata, Anna. 2012. Anti-vaccine activists, Web 2.0, and the postmodern paradigm – An overview of tactics and tropes used online by the anti-vaccination movement. *Vaccine* 30(25). 3778–3789. https://doi.org/10.1016/j.vaccine.2011.11.112.
Kehoe, Andrew, Matt Gee, Robert Lawson, Mark McGlashan & Tatiana Tkacukova. 2021. TRAC:COVID – Trust and Communication: A Coronavirus Online Visual Dashboard. https://traccovid.com.
Larson, Heidi J, William S Schulz, Joseph D Tucker & David M D Smith. 2015. Measuring Vaccine Confidence: Introducing a Global Vaccine Confidence Index. *PLoS Currents*. https://doi.org/10.1371/currents.outbreaks.ce0f6177bc97332602a8e3fe7d7f7cc4.
Le Roux, Brigitte & Henry Rouanet. 2010. *Multiple Correspondence Analysis*. 2455 Teller Road, Thousand Oaks California 91320 United States of America: SAGE Publications, Inc. https://doi.org/10.4135/9781412993906.
Lê, Sébastien, Julie Josse & François Husson. 2008. **FactoMineR** : An *R* Package for Multivariate Analysis. *Journal of Statistical Software* 25(1). https://doi.org/10.18637/jss.v025.i01.
Lyons, Benjamin, Vittorio Merola & Jason Reifler. 2019. Not Just Asking Questions: Effects of Implicit and Explicit Conspiracy Information About Vaccines and Genetic Modification. *Health Communication* 34(14). 1741–1750. https://doi.org/10.1080/10410236.2018.1530526.
MacDonald, Noni E. 2015. Vaccine hesitancy: Definition, scope and determinants. *Vaccine* 33(34). 4161–4164. https://doi.org/10.1016/j.vaccine.2015.04.036.
McGlashan, Mark, Matt Gee, Andrew Kehoe, Robert Lawson & Tatiana Tkacukova. 2021. *TRAC:COVID Case study 2: misinformation, authority, and trust*. Birmingham City University, Birmingham: Working Paper.
McGlashan, Mark & Alexandra Krendel. 2023. Keywords of the manosphere. *International Journal of Corpus Linguistics*. https://doi.org/10.1075/ijcl.22053.mcg.
Pennycook, Gordon, Jonathon McPhetres, Yunhao Zhang, Jackson G. Lu & David G. Rand. 2020. Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. *Psychological Science* 31(7). 770–780. https://doi.org/10.1177/0956797620939054.
Reisigl, M & Ruth Wodak. 2016. The Discourse-Historical Approach (DHA). In Ruth Wodak & Michael Meyer (eds.), *Methods of Critical Discourse Studies*, 23–61. 3rd edn. London ; Thousand Oaks, California: SAGE.

Sha, Yuze & Isobelle Clarke. 2025. Using ATLAS.ti to interpret Keyword Co-occurrence Analysis: A case study on the representation of vaccin* across pseudoscience and conspiracy websites. *Linguistic Vanguard*.
Suarez-Lledo, Victor & Javier Alvarez-Galvez. 2021. Prevalence of Health Misinformation on Social Media: Systematic Review. *Journal of Medical Internet Research* 23(1). e17187. https://doi.org/10.2196/17187.
Sunderland, Jane. 2004. *Gendered discourses*. Houndmills, Basingstoke, Hampshire ; New York: Palgrave Macmillan.