

## Original Paper

# Psycho-linguistic Differences among Competing Vaccination Communities on Social Media

Jialiang Shi<sup>1</sup>, Piyush Ghasiya<sup>2</sup> and Kazutoshi Sasahara<sup>2\*</sup>

<sup>1</sup>*Graduate School of Informatics, Nagoya University, Nagoya, Japan*

<sup>2</sup>*School of Environment and Society, Tokyo Institute of Technology, Tokyo, Japan*

---

### ABSTRACT

Currently, the significance of social media in disseminating noteworthy information on topics such as health, politics, and the economy is indisputable. During the COVID-19 pandemic, anti-vaxxers have used social media to distribute fake news and anxiety-provoking information about the vaccine. Such social media practice may harm the public. Here, we characterise the psycho-linguistic features of anti-vaxxers on the online social network Twitter. For this, we collected COVID-19 related tweets from February 2020 to June 2021 to analyse vaccination stance, linguistic features, and social network characteristics. Our results demonstrated that, compared to pro-vaxxers, anti-vaxxers tend to have more negative emotions, narrative thinking, and immoral tendencies. Furthermore, we found a tighter network structure in anti-vaxxers even after mass vaccination. This study can advance our understanding of the online anti-vaccination movement, and become critical for social media management and policy action during and after the pandemic.

---

*Keywords:* Anti-vaccine movement, COVID-19, psycho-linguistic features, morality, social media, text mining.

---

\*Corresponding author: Kazutoshi Sasahara, [sasahara.k.aa@m.titech.ac.jp](mailto:sasahara.k.aa@m.titech.ac.jp). We would like thank the members of CREST projects (JPMJCR20D3 and JPMJCR17A4) for fruitful discussions.

---

Received 31 October 2021; Revised 18 March 2022

ISSN 2048-7703; DOI 10.1561/116.00000056

© 2022 J. Shi, P. Ghasiya and K. Sasahara

## 1 Introduction

The opposition to vaccines or anti-vaccinationism is an old phenomenon with roots traced back to the 1850s. Much of the anti-vaccine sentiment of the era was laid out by John Gibbs when he published a booklet against the Vaccination Act of 1853 (British Government)—which required mandatory vaccination for all infants over 4 months old [9]. There are several reasons for this, such as conspiratorial beliefs, disgust, individualism, and hierarchical worldviews for anti-vaccine behaviour or attitudes of people [19].

While early anti-vaccine activists distributed pamphlets and organized rallies, the anti-vaxxers of the 21st century have access to Facebook, Twitter, and other social media platforms to spread their views globally. Due to these communication tools, we see extremely potent (small but far reaching) anti-vaccination movements during the current COVID-19 pandemic. Although as of October 17, 2021, 47.5% of the world’s population has received at least one dose of the COVID-19 vaccine [29], and the immunization program is going strong, researchers believe that the present anti-vaccination movement can undermine the efforts to end the coronavirus pandemic [2]. Therefore, at a time when the COVID-19 pandemic has killed millions of people, it is essential to counter those who actively spread vaccine-related mis/dis-information, fake news, conspiracy, and propaganda on social media platforms, to make it possible to immunize all vulnerable people against the deadly COVID-19 disease.

As social media plays an important role in propagating global anti-vaccination beliefs against the COVID-19 vaccine, several researchers are analysing anti-vaxxers behaviour and attitude on social media [25]. However, few researchers have tackled questions related to the psychological and moral aspects in the language of anti-vaxxers on social media. By understanding the social media posts of anti-vaxxers from psychological, moral, and linguistic perspectives, we can peek into their minds, which can be helpful in developing counter strategies against the anti-vaccination movement, both at individual and platform levels [7].

In this research, we used longitudinal data that captures COVID-19 discussions on Twitter. First, we utilized the Louvain algorithm to distinguish between anti-vaxxers and pro-vaxxers. Then we used the Linguistic Inquiry and Word Count (LIWC) 2015 [41] to quantify their semantic differences, and then used the Moral Foundation Dictionary (MFD) [14] to compare moral tendencies in pro- and anti-vaxxers. As explained later, these are powerful tools to capture anti-vaxxers’ characteristics of spontaneous social media posts implicated in linguistic information. We found that as the COVID-19 pandemic spread and vaccination started, the psycho-linguistic features of anti-vaxxers’ posts also changed.

Here we pose two research questions (RQs):

- **RQ1:** How do anti-vaxxers differ from pro-vaxxers in terms of psycho-linguistic features as well as moral values they convey?
- **RQ2:** How did the mass vaccination (i.e., December 2020) affect pro- and anti-vaxxers' behaviour?

Regarding RQ1, previous research has reported that anti-vaxxers use analytical language less frequently [33] and specific moral foundations are associated with vaccine hesitancy [1, 18]. We intend to confirm if these psycho-linguistic properties are observed in our data. We can leverage such linguistic signals to better detect harmful content from anti-vaxxers. Regarding RQ2, various studies have analysed public moods using social data, but emotional changes have not been measured before and after mass vaccination concerning the evolution of anti-vaxxers. Therefore, we measured mass vaccination effects on emotions in pro- and anti-vaccine groups to get hints on vaccine operations.

Answering these questions and understanding the whole picture of anti-vaxxers helps us develop countermeasures against the online anti-vaccine movement. As detailed later, our results indicate that we need to pay attention to the apparent negative tendency of anti-vaxxers in language expressions and the fact that they are resolute in their beliefs in network structure.

## 2 Related Work

Various social media platforms work as conduits in the circulation and amplification of fake news [24]. Health-related misinformation and fake news is also a burgeoning research topic among social scientists and medical research professionals. In this line of research, a recent study by Suarez-Lledo *et al.* investigated health misinformation on social media and found that “vaccine” is the fastest spreading topic on Twitter [47]. YouTube is another platform on which anti-vaccine narratives are often broadcast. Lahouti *et al.* investigated YouTube videos to understand anti-vaccine sentiments in France where vaccine hesitancy is high and found that anti-vaxxers are very active on YouTube [26].

Different demographic characteristics and living conditions also play an important role in anti-vaccine views/beliefs. Lyu *et al.* showed these differences in their analysis of Twitter data [28]. Cultural differences also play a critical role in developing pro- or anti-vaccine views. A study by Luo *et al.* observed that the difference between the topic of anti-vaxxers in China and the United States is caused by cultural distinctions between each country [27]. In India,

concerns about anti-vaccination are more likely to stem from health concerns and fear of allergic reactions [43].

In addition to differences in content and geography, curiosity to understand what kind of arguments anti-vaxxers give on Facebook has also influenced research. Wawrzuta *et al.* analysed Polish media fan pages and found that the COVID-19 anti-vaccine movement has new arguments, such as the vaccine not being properly tested. However, the classic argument—not trusting the government—also remains popular [50]. Nuzhath *et al.* analysed Twitter data to understand the prominent topics of discussion among anti-vaxxers in Bangladesh and found that misinformation, vaccine safety and effectiveness, conspiracy theories, and mistrust in government are some of the main topics [39].

Jamison *et al.* also found that both pro- and anti-vaxxers are spreading less reliable information or claims on social media and suggested that while all research focus is on bad actors (anti-vaxxers) to understand the anti-vaccine movement, good actors (pro-vaxxers) also play a role in the spread of the “infodemic” [21].

LIWC is a useful tool in investigating psycho-linguistic features. Mitra *et al.* utilized LIWC to understand anti-vaccination attitudes in social media and found that anti-vaxxers tend to be influenced by conspiracy theories [35]. Linguistic differences often accompany network differences. Johnson *et al.* utilized network analysis to understand the evolution of pro- and anti-vaccine communities [22]. One significant finding of their research is related to undecided individuals. Their finding challenges the current thinking that undecided individuals are a passive background population in the battle of “hearts and minds”. Germani and Biller-Andorno show that, compared to pro-vaxxers, anti-vaxxers on Twitter have a high number of influencers and these influencers lead the anti-vaxxers discussion [11]. They also showed that before the suspension of his Twitter account, Donald Trump was the main driver of anti-vaccine misinformation on Twitter. Lastly, Menon and Carley characterized COVID-19 misinformation communities on Twitter [33]. Their analysis suggested that a large majority of misinformed users may be anti-vaxxers. Further, their socio-linguistic analysis also showed that informed users (who spread true information) use more narrative thinking than misinformed users (who spread misinformation).

Social and behavioural scientists have worked to understand the moral basis of people’s judgment for a long time. Voluminous research and a theoretical framework led the foundation of the Moral Foundations Theory (MFT) [12, 13]. MFT works on the assumption that there are 5 major moral foundations: (1) “Care/Harm,” which focuses on not harming others and protecting the vulnerable; (2) “Fairness/Cheating,” which assumes equivalent exchange without cheating to be good; (3) “Loyalty/Betrayal,” which concerns a collective entity instead of individuals; (4) “Authority/Subversion,” which postulates

respect for authority, resulting in maintaining the hierarchy; and (5) “Sanctity/Degradation,” which involves a feeling of disgust caused by the impure. Moral Foundations Theory (MFT) was also used to understand why morality varies across cultures yet still shows similarities and recurrent themes [12].

Within vaccination hesitancy, past research also shows that core morality will influence people’s attitudes toward vaccination [10]. For example, “Liberty” is likely implicated in the decision to not vaccinate a child [4]; endorsement of the foundations of Purity and Liberty are associated with vaccine hesitancy [1, 18]. Moral Foundations Dictionary (MFD) is often used to quantify and understand the extent to which moral foundations are expressed in a text [14]. We used the original version of MFD for this study. There are also some candidates for moral foundations like “Liberty/oppression” related to MFT [17]. However, since the original MFD does not include this dimension, we do not discuss it in this study.

Although several aspects of anti-vaccine communities have been reported by a series of studies (such as those previously mentioned), the psycho-linguistic features of anti-vaccine posts that may increase vaccine hesitancy, especially in the context of the COVID-19 pandemic, remain unclear and understudied. This knowledge is critical to reduce vaccine mis/dis-information and achieve herd immunity towards a post-pandemic era. Therefore, we investigate psycho-linguistic properties of anti-vaxxers in terms of the above-mentioned research questions.

### 3 Data and Methods

#### 3.1 Data Collection

To obtain the longitudinal data of social media posts we used the Twitter Search API to collect COVID-19 related tweets, replies, and retweets (RTs) with keywords such as “covid,” and “covid-19” from February 20, 2020, to June 30, 2021. In this research, we only focused on English language content.

We then filtered vaccine-related tweets from the English dataset for our analyses, using the keywords “vaccination,” “vaccines,” “vaccine,” “vaccinated,” “vaccination,” “vaccineoutside,” “vaccinate,” “vaccinologist,” “vacciner,” and “coronavirusvaccine.” In total we collected 11,395,103 retweets, 11,395,103 tweets, 465,037 replies and 3,781,447 unique users. Our data and code are available online (<https://osf.io/FSM23/>).

#### 3.2 Characterisation of Anti- and Pro-vax Groups

Previous studies have found that understanding retweet networks and community detection is a useful method to reveal communication patterns among

communities [8]. Therefore, we followed the same approach to classify pro- and anti-vaxxers on Twitter. We first constructed a retweet network from the vaccine-related retweets. A retweet network on Twitter can be defined as a directed weighted graph, where nodes and edges represent users and retweet transmissions, respectively. In our research, the direction from one node to another represents a user retweeting another user’s post. The weight represents the number of times the user retweeted another user’s post.

We used the Louvain algorithm, a standard algorithm for community detection [6], to find clusters including pro- and anti-vaxxers. We then applied the  $k$ -core decomposition ( $k = 1$ ) and retained nodes whose indegree (i.e., the number of retweets by different users) was greater than 20 in order to focus on significantly influential users. The Forceatlas2 layout [20] was used to visualize the structure of the resulting clusters. For these processes we used the Gephi software [3]. By manually confirming the top 10 popular users in each cluster (i.e., looking at users’ profiles, tweets, and retweeted contents), we deciphered which clusters represent pro- and anti-vaxxers, and others.

### 3.3 Network Measures

The retweet network was also quantified using the following standard measures to illuminate its structural features, in addition to the number of nodes (users) and links (unique retweet relations):

- **Network density:** the ratio of actual connections and potential connections.
- **Clustering coefficient:** measures the degree to which nodes in a network tend to form triangles or “clusters.”
- **Average distance:** the average minimum number of connections to be crossed from any arbitrary node to any other.

For this measurement, we converted our retweet network to undirected graphs because these measures are defined for undirected graphs.

### 3.4 Psycho-linguistic Analysis with LIWC

LIWC is a standard tool in social psychology for computerized text analysis [41]. Given a text, LIWC can quantify emotions, thinking style, and social concerns by counting the dictionary words registered in LIWC. The psychological expressions in a text can also reveal critical aspects. Table 1 shows all the LIWC categories and subcategories used in this study. LIWC was used in many studies, such as sentiment analysis and social relationships, to evaluate the impact of psychological expression [48].

Table 1: LIWC2015 categories and subcategories with example words.

Category	Abbrev	Examples	# of words in the category
<b>Summary language variables</b>			
Analytical thinking	Analytic	–	–
<b>Total function words</b>	funct	it, to, no, very	491
Total pronouns	pronoun	I, them, itself	153
Articles	article	a, an, the	3
Prepositions	prep	to, with, above	74
Auxiliary verbs	auxverb	am, will, have	141
Common Adverbs	adverb	very, really	140
Conjunctions	conj	and, but, whereas	43
Negations	negate	no, not, never	62
<b>Psychological processes</b>			
Affective processes	affect	happy, cried	1393
Positive emotion	posemo	love, nice, sweet	620
Negative emotion	negemo	hurt, ugly, nasty	744
<b>Personal concerns</b>			
Work	work	job, majors, xerox	444
Leisure	leisure	cook, chat, movie	296
Home	home	kitchen, landlord	100
Money	money	audit, cash, owe	226
Religion	relig	altar, church	174
Death	death	bury, coffin, kill	74

We used LIWC 2015 to investigate attitudinal and linguistic differences between pro- and anti-vaxxers. LIWC is especially useful for measuring the sentiment and thinking styles in a tweet. According to [23], analytical thinking and narrative thinking are often in opposition, which may characterize pro- and anti-vaxxers. Narrative thinking is identifying conceptual categories and organizing them in hierarchical ways [23], which can be linked to the frequent use of pronouns and function words [42] and the less frequent use of analytic categories in LIWC. Thus, we compare LIWC scores for analytic, pronouns, and function words between pro- and anti-vaxxers to determine whether they are analytic thinkers or narrative thinkers.

### 3.5 Morality Analysis with MFD

As mentioned in Section 2, Graham *et al.* developed the moral foundations dictionary (MFD), which consists of 156 words and 168 word stems to quantify the frequency of words referring to virtues and vices associated with

each moral foundation [14]. They attempted to understand moral tendencies among liberals and conservatives and found that liberals consistently showed greater endorsement and use of the Care and Fairness foundations compared to the other three foundations, whereas conservatives endorsed and used the five foundations more equally. The Japanese version of MFD was used to reveal that a trade-off between the Fairness and Authority foundations plays a key role in the online communication of Japanese users on Twitter [31]. We also used the MFD in this study to measure five moral foundations as additional psycho-linguistic features for pro- and anti-vaccine clusters on Twitter.

## 4 Results

### 4.1 Pro/Anti-vaxxers and Other Communities

The above-mentioned processes resulted in the retweet network based on vaccine-related tweets (77,934 nodes and 1,999,164 edges). In this network, we identified six main clusters: (1) Left-wing (32.0%), (2) Pro-vaxxers (22.2%), (3) Right-wing (13.0%), (4) Anti-vaxxers (11.7%), (5) India-related (5.7%), and (6) Canada-related (3.6%). These clusters are shown in Figure 1.

We noticed the apparent differences in the content of the top 10 users selected by indegree (i.e., the number of retweets by different users) in each cluster. Our examinations identified the second largest cluster as a pro-vaccination

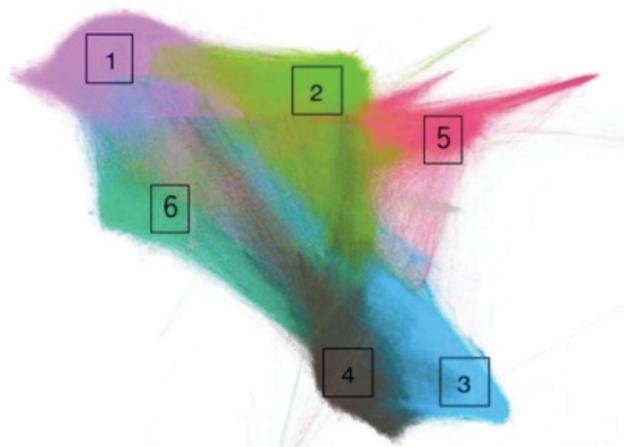


Figure 1: Retweet network and communities constructed from vaccine-related tweets: 1. Left-wing (32.0%), 2. Pro-vaxxers (22.2%), 3. Right-wing (13.0%), 4. Anti-vaxxers (11.7%), 5. India-related (5.7%), 6. Canada-related (3.6%).



Table 2: Sample tweets of pro- and anti-vaxxers.

Pro-vaxxers	Anti-vaxxers
this is a critical step toward developing vaccines and treatments. #coronavirus #ncov	#coronavirus & the georgia guidestones – what you need to know! #billgates #vaccines
citizens 60 and above who have not yet been vaccinated can go to their nearest vaccination site to be vaccinated. here is a list of vaccination sites: # ichoosevaccination	covid-19 vaccine company under federal investigation over allegedly misrepresenting its role in government program

Table 3: Network features of the pro- and anti-vax clusters in the retweet network.

Measure	Overall	Pro-vaxxers	Anti-vaxxers
Nodes (# users)	77,934	11,522	6,909
Links (# unique RTs)	1,999,164	207,343	171,589
Network density	3.29E-04	1.56E-03	3.60E-03
Clustering coefficient	0.039	0.045	0.070
Average distance	6.696	1.581	1.393

group and the fourth largest cluster as an anti-vaccination group. Table 2 shows two example tweets for pro-vaxxers and anti-vaxxers, respectively. Here we can see that anti-vaxxers intimated the Bill Gates conspiracy theory (top) and distrust of the government (bottom). In addition, we observed that some groups have a clear political orientation, such as the first and third largest groups. In the COVID-19 pandemic, vaccine operation is an important political issue; thus, polemical clusters might emerge together with pro- and anti-vaccine clusters. Two other groups turn out to be country-related (i.e., India and Canada), which are related to vaccine strategies in these countries. Because our interest is in psycho-linguistic features in anti-vaccine groups to gain insights into countermeasures, we hereafter restricted our analysis to pro- and anti-vaccine groups (clusters 2 and 4).

The network measures for the entire retweet network and the pro- and anti-vax clusters are summarised in Table 3, which reveals that the number of pro-vaxxers is larger than that of anti-vaxxers. However, anti-vaxxers are more densely connected according to network density and clustering coefficient values. In addition, the distributions of the indegree for pro- and anti-vaccine

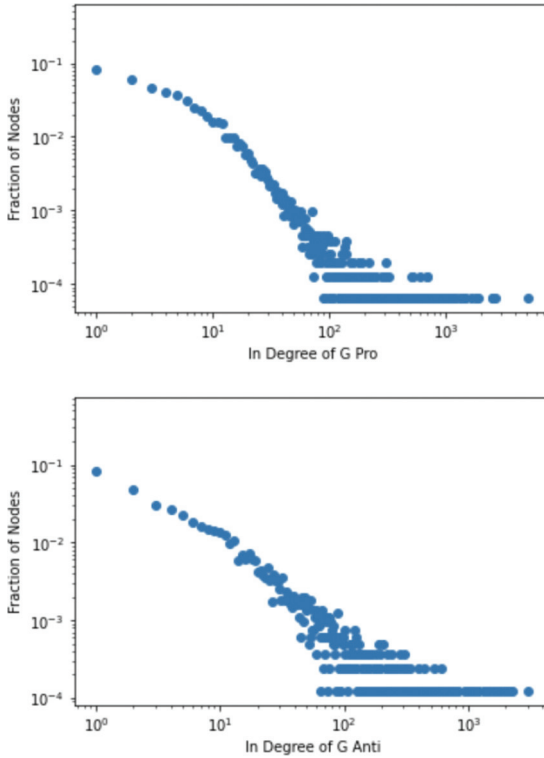


Figure 2: Distribution of the indegree for anti-vaxxers (top) and pro-vaxxers (bottom).

groups have heavy-tailed distributions (Figure 2). This implies the existence of influential accounts in both pro- and anti-vaccine groups, which is often the case in a spreading phenomenon.

Furthermore, we compared the network characteristics of these two communities that changed after mass vaccination (i.e., December 2020). As shown in Table 4, overall, the difference between two period is minute; only changes in network density significantly decreased. When it comes to two groups, the sizes (nodes and links) of pro-vaxxers increased but network density decreased in the after period. Conversely, those of anti-vaxxers decreased, but network density increased in the after period. We can infer from these results that although the pro-vaccine community grew once mass vaccination began, they became sparser than before. However, although the anti-vaccine community is comparatively small and lose members once the vaccination started, those became more tightly knit.

We found that among the top 10 users in these clusters, six accounts belonging to anti-vaxxers have been banned by Twitter, while none of the

Table 4: Network features before and after mass vaccination.

Measure	Overall		Pro-vaxxers		Anti-vaxxers	
	Before	After	Before	After	Before	After
Nodes (# users)	34,283	36,441	4,877	7,031	3,763	3,280
Links (# unique RTs)	736,888	733,073	60,258	100,897	75,950	63,492
Network density	6.27E-04	5.52E-04	2.53E-03	2.04E-03	5.37E-03	5.90E-03

pro-vaxxer accounts have been banned. This observation indicates that our classification of pro- and anti-vaxxers is correct and reliable.

#### 4.2 Psycho-linguistic Features of Pro- and Anti-vaxxers

After identifying pro- and anti-vaccine communities, we compared the two in terms of psycho-linguistic features, including emotion. Specifically, we selected categories such as analytical thinking, affective processes, and personal concerns from LIWC. The LIWC categories and subcategories we used are shown in Table 1. In the following, we used the independent *t*-test to compare the average score between anti- and pro-vaxxers. For all tests, the confidence level is 95%. The confidence intervals were computed from the bootstrapped samples (We randomly sampled tweets ( $n = 45,809$ ) and replies ( $n = 254,660$ ) and repeated it multiple times ( $n = 10,000$ ) for statistical evaluation.) Additionally, to find out the differences between tweets and replies, we separately listed the scores of the two communities.

Affective processes include several emotional words in LIWC. Examples of positive emotion words are “love,” “nice,” and “sweet,” while words such as “hurt,” “ugly,” and “nasty” are seen as negative. Figure 3 shows that, overall, anti-vaxxers expressed more affect (positive and negative combined) than pro-vaxxers in both tweets and replies. If we consider positive and negative affect separately, we find that in tweets and replies, anti-vaxxers have higher negative and lower positive emotions, while the opposite is true for pro-vaxxers.

Mass vaccination for COVID-19, which was expected to improve negative emotions both in pro- and anti-vaccination groups, started in December 2021 [44]. However, Figure 4 shows that negative emotions instead increased before the mass vaccination (Period 1: February 2020 to November 2020) and after it (Period 2: December 2020 to June 2021). In addition, for replies and tweets, anti-vaxxers showed higher negative emotions compared to pro-vaxxers during Period 1. This trend increased in Period 2, and similarly anti-vaxxers’

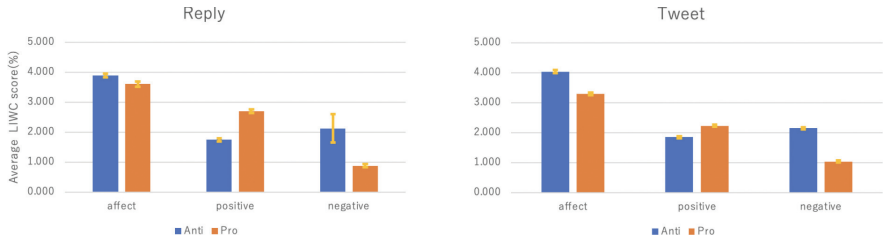


Figure 3: Average emotion score (%) for anti- and pro-vaxxers. Affective processes including positive and negative emotion. Differences between anti- and pro-groups are all significant (independent  $t$ -test,  $p < 0.001$ ).

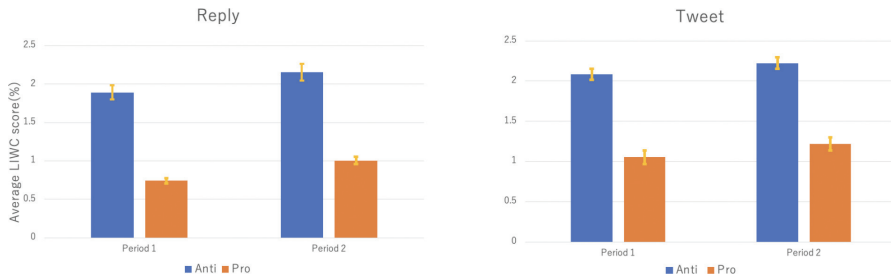


Figure 4: Average negative emotion scores before and after the mass vaccination (Period 1: February 2020 to November 2020; Period 2: December 2020 to June 2021). Differences between anti- and pro-groups are all significant (independent  $t$ -test,  $p < 0.001$ ).

expression of emotion was larger than that of pro-vaxxers. These results suggest that anti-vaxxers propagated their anti-vaccination beliefs more passionately and emotionally even after the start of mass vaccination [16].

There are subcategories of personal concerns in LIWC. In Figure 6, the frequency of words used in personal concerns for anti- and pro-vaxxers are visualized using Wordcloud. We found that the three most highly used subcategories by pro-vaxxers are money (black), religion (orange), and leisure (blue), while anti-vaxxers have shown a higher usage of death (purple) and work (green).

Comprehending the thought process (whether analytical or narrative thinking) of both pro- and anti-vaxxers can also provide critical information that can be utilized to create a mitigation strategy. In both replies and tweets, we found that anti-vaxxers used more function words and pronouns, and had a lower analytic score than pro-vaxxers. Recall that the higher function words, pronouns score and lower analytic scores represent narrative thinking. This suggests that anti-vaxxers use more narrative thinking than pro-vaxxers (Figure 5). Similarly, we found that replies showed higher usage of function

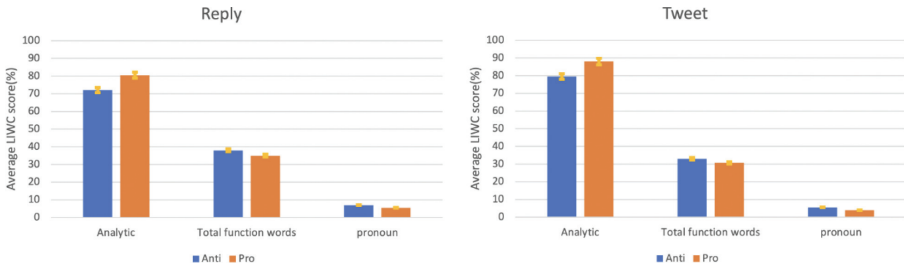


Figure 5: LIWC scores for analytical, total function words, and total pronouns on average for anti- and pro-vaxxers. Differences between anti- and pro-groups are all significant (independent *t*-test,  $p < 0.001$ ).



Figure 6: Word clouds for pro- and anti-vaxxers. Colours correspond to different LIWC subcategories: green for work, blue for leisure, red for home, black for money, orange for religion, and purple for death.

words, pronouns and a lower analytic score than tweets. This result also indicates that the trend of narrative thinking can be stronger in replies, a targeted message). Furthermore, we compared analytic thinking tendency in replies between anti-vaxxers and between anti- and pro-vaxxers. It turns out that anti-vaxxers were in the narrative mode when replying to pro-vaxxers (Figure 7).

### 4.3 Moral Foundations in Pro- and Anti-vaxxers

Morality is another psycho-linguistic feature. It is important to explain the process of making social judgements [45]. As explained, we used MFD to assess morality tendencies among pro- and anti-vaxxers. The comparison of the average scores of both groups is shown in Table 5. We can see that anti-vaxxers show higher scores in the Vice moral foundations (e.g., Harm, Cheating, Betrayal, Subversion, and Degradation), indicating they tend to use moral violating words. While in the Virtue moral foundations, pro-vaxxers show higher scores in Care, Fairness, Loyalty and Sanctity, indicating they expressed

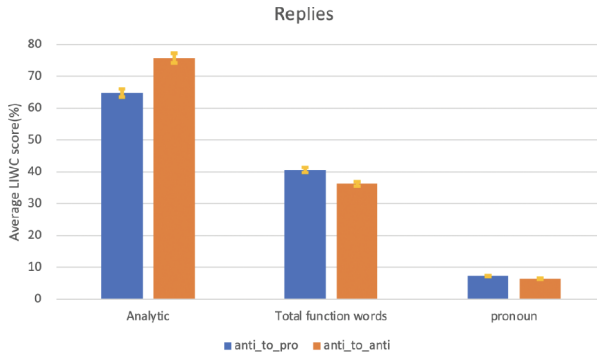


Figure 7: LIWC scores for analytical, total function words, and total pronouns in replies between anti-vaxxers and between anti- and pro-vaxxers. Differences are all significant (independent  $t$ -test,  $p < 0.001$ ).

Table 5: Comparing the average score of pro-vaxxers (pro) and anti-vaxxers (anti).  $p$ -values from independent  $t$ -test are listed.

Measure	Mean Diff	$p$ -value
Care	Pro > Anti	<0.01
Harm	Pro < Anti	<0.0001
Fairness	Pro > Anti	<0.0001
Cheating	Pro < Anti	<0.0001
Loyalty	Pro > Anti	<0.0001
Betrayal	Pro < Anti	<0.0001
Authority	n.s.	n.s.
Subversion	Pro < Anti	<0.0001
Sanctity	Pro > Anti	<0.0001
Degradation	Pro < Anti	<0.0001

moral content. Taken together, anti-vaxxers frequently used immoral language in their posts, thereby distributing anxiety-provoking (fake)news and messages.

## 5 Discussion

We have shown that anti-vaxxers used more affective process and negative words, while pro-vaxxers used more positive words (Figure 3). Another research also showed that the strategy of the anti-vax groups involves strong emotions with highly toxic and negative words during the early COVID-19 pandemic [36, 37]. Thus, our finding is consistent with the previous result. In addition, we showed that anti-vaxxers tend to use vice moral languages in all of the five

moral foundations, whereas pro-vaxxers exhibit the opposite tendency. These findings are the answer to RQ1.

We have seen positive effects globally since the mass vaccination started. How did anti-vax communities change their interactions on Twitter after the mass vaccination? Our results showed that, compared to pro-vaxxers, anti-vaxxers expressed even higher negative emotions after the mass vaccination, which suggests their firmness (Figure 4). These results are the answer to RQ2. Strangely, similar results were observed in pro-vaxxers, although we expected that pro-vaxxers became more positive emotionally. These results cannot be explained by the “backfire effect,” where exposure to opposing views can lead to the increased commitment to preexisting beliefs [40]. Therefore, further investigation is needed to unveil this emotional shift phenomenon.

In addition, we found that anti-vaxxers show more narrative thinking (higher score of total function words and total pronouns) and a lower score in analytic thinking, as shown in Figure 5. Unlike analytic thinking, narrative thinking tends to be more personal, and rumours are easier to spread [46]. However, a previous study showed that informed groups (pro-vaxxers) have more narrative thinking [33], which is contrary to our results but that study was conducted on a small dataset for a short period (less than a month). We used larger and more longitudinal data, and our results show that narrative thinking is more dominant among anti-vaxxers, especially during the COVID-19 pandemic. As mentioned above, anti-vaxxers showed negative moral tendencies (i.e., usage of vice words) based on MFD, consistent with anti-vaxxers using more negative emotions because vice words are correlated with those [30]. Figure 6 shows that anti-vaxxers use more “death” words, a subcategory belonging to personal concern. This result suggests that anti-vaxxers may be more concerned about death caused by vaccination.

Because replying can reach beyond follow-follower relations, reply patterns on Twitter can also provide valuable information, and previous studies have provided such evidence (e.g., [5, 38, 49]). We also investigated reply patterns to understand the psycho-linguistic features of Twitter conversations on vaccination. We found that replies showed more narrative thinking (Figure 5), because replying is an immediate reaction and thus less inclined toward analytic thinking.

These findings provide useful implications for spreading credible content from pro-vaxxers while reducing the exposure of misinformation and anxiety-provoking posts from anti-vaxxers. Because anti-vaxxers’ posts are typically charged with higher negativity and vice morality, algorithms can detect their harmful content by measuring psycho-linguistic features. Then, the SNS system can hide such posts until users agree to avoid unnecessary exposure. At the same time, psycho-linguistic features as well as network structures help us identify those pro-vaxxers who spontaneously transmit trustful vaccine information in near real-time. Thus, we may leverage such information for social fact-checking to oppose anti-vaccine narratives at scale.

## 6 Limitations and Future Work

Several previous studies have shown that Twitter and other social media platforms are not representative of the general population [15, 32, 34]. Even though our findings are statistically significant, we should be aware of the gap between online social networks and reality. Besides this, the focus of our study is only Twitter, but there are several other popular social media platforms such as Facebook, Reddit, and Instagram. Analysis of vaccination communities on these platforms can also bring forward critical insights. In our future study, we will focus on the comparative analysis of vaccination communities on Twitter, Facebook, Instagram, and Reddit. Comparison between different languages and cultures is also an important future direction (e.g., [36]).

## 7 Conclusion

In this study, we quantified psycho-linguistic differences among competing vaccination communities on Twitter during the COVID-19 pandemic. Based on the differences in linguistic usage, we found that anti-vaxxers tend to show more negative, narrative thinking, and immoral tendencies. In terms of network difference, anti-vaxxers showed a tighter network structure and strengthening of their anti-vaccine beliefs.

The mass vaccination of people during the COVID-19 pandemic shows that these vaccines are working. However, even this news does not deter anti-vaxxers from their beliefs; rather, it strengthens these negative emotions. This vicious circle needs to break if we want to immunize all high-risk people in the world and achieve herd immunity, while preventing the online anti-vaccine movement. Our results provide key insights for developing countermeasures against the online anti-vaccine movement, both at individual and society levels.

## Financial Support

This work was supported by JST, CREST Grant Number JPMJCR20D3, Japan.

## Biographies

**Jialiang Shi** received the B.E. degrees from the Chang'an University, Xi'an, China in 2019. He is currently enrolled as a Master student at the Graduate



School of Informatics, Nagoya University. His research interests focus on fake news on social media and natural language processing.

**Piyush Ghasiya** received Ph.D. from Kyushu University in 2021. Presently, he is a postdoctoral research fellow at the School of Environment and Society, Tokyo Institute of Technology. His research interests include computational social science and international relations.

**Kazutoshi Sasahara** received his Ph.D. from The University of Tokyo in 2005. He worked as a researcher at several institutions including RIKEN, The University of Tokyo, and UCLA. From 2012 to 2020, he was an Assistant Professor in Graduate School of Informatics, Nagoya University. Since 2020, he is an Associate Professor in the School of Environment and Society, Tokyo Institute of Technology. His research interests are computational social science and social innovation.

## References

- [1] A. B. Amin, R. A. Bednarczyk, C. E. Ray, K. J. Melchiori, J. Graham, J. R. Huntsinger, and S. B. Omer, “Association of Moral Values with Vaccine Hesitancy,” *Nature Human Behaviour*, 1(12), 2017, 873–80.
- [2] P. Ball, “Anti-vaccine Movement Could Undermine Efforts to End Coronavirus Pandemic, Researchers Warn,” *Nature*, 581(7808), 2020, 251–2.
- [3] M. Bastian, S. Heymann, and M. Jacomy, “Gephi: An Open Source Software for Exploring and Manipulating Networks,” in *Third International AAAI Conference on Weblogs and Social Media*, 2009.
- [4] F. H. Beard, J. Leask, and P. B. McIntyre, “No Jab, No Pay and Vaccine Refusal in Australia: The Jury is Out,” *Medical Journal of Australia*, 206(9), 2017, 381–3.
- [5] C. A. Bliss, I. M. Kloumann, K. D. Harris, C. M. Danforth, and P. S. Dodds, “Twitter Reciprocal Reply Networks Exhibit Assortativity with Respect to Happiness,” *Journal of Computational Science*, 3(5), 2012, 388–97.
- [6] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, “Fast Unfolding of Communities in Large Networks,” *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), 2008, P10008.
- [7] T. Burki, “The Online Anti-vaccine Movement in the Age of COVID-19,” *The Lancet Digital Health*, 2(10), 2020, e504–5.
- [8] D. Cherepnalkoski and I. Mozetič, “Retweet Networks of the European Parliament: Evaluation of the Community Structure,” *Applied Network Science*, 1(1), 2016, 1–20.

- [9] A. Day, “Jonathan M. Berman, Anti-Vaxxers: How to Challenge a Misinformed Movement,” 2021.
- [10] M. V. Day, S. T. Fiske, E. L. Downing, and T. E. Trail, “Shifting Liberal and Conservative Attitudes Using Moral Foundations Theory,” *Personality and Social Psychology Bulletin*, 40(12), 2014, 1559–73.
- [11] F. Germani and N. Biller-Andorno, “The Anti-vaccination Infodemic on Social Media: A Behavioral Analysis,” *PloS One*, 16(3), 2021, e0247642.
- [12] J. Graham, J. Haidt, S. Koleva, M. Motyl, R. Iyer, S. P. Wojcik, and P. H. Ditto, “Moral Foundations Theory: The Pragmatic Validity of Moral Pluralism,” in *Advances in Experimental Social Psychology*, Vol. 47, Elsevier, 2013, 55–130.
- [13] J. Graham, J. Haidt, M. Motyl, P. Meindl, C. Iskiwitch, and M. Mooijman, “Moral Foundations Theory,” *Atlas of Moral Psychology*, 2018, 211–22.
- [14] J. Graham, J. Haidt, and B. A. Nosek, “Liberals and Conservatives Rely on Different Sets of Moral Foundations,” *Journal of Personality and Social Psychology*, 96(5), 2009, 1029.
- [15] S. Greenwood, A. Perrin, and M. Duggan, “Social Media Update 2016,” *Pew Research Center*, 11(2), 2016, 1–18.
- [16] C. Guo, J. Cao, X. Zhang, K. Shu, and M. Yu, “Exploiting Emotions for Fake News Detection on Social Media,” *arXiv preprint arXiv:1903.01728*, 2019.
- [17] J. Haidt, *The Righteous Mind: Why Good People Are Divided by Politics and Religion*, Pantheon, 2013.
- [18] F. Heine and E. Wolters, “Using Moral Foundations in Government Communication to Reduce Vaccine Hesitancy,” *PloS One*, 16(11), 2021, e0259435.
- [19] M. J. Hornsey, E. A. Harris, and K. S. Fielding, “The Psychological Roots of Anti-vaccination Attitudes: A 24-nation Investigation,” *Health Psychology*, 37(4), 2018, 307.
- [20] M. Jacomy, T. Venturini, S. Heymann, and M. Bastian, “ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software,” *PloS one*, 9(6), 2014, e98679.
- [21] A. M. Jamison, D. A. Broniatowski, M. Dredze, A. Sangraula, M. C. Smith, and S. C. Quinn, “Not Just Conspiracy Theories: Vaccine Opponents and Proponents Add to the COVID-19 ‘Infodemic’ on Twitter,” *Harvard Kennedy School Misinformation Review*, 1(3), 2020.
- [22] N. F. Johnson, N. Velásquez, N. J. Restrepo, R. Leahy, N. Gabriel, S. El Oud, M. Zheng, P. Manrique, S. Wuchty, and Y. Lupu, “The Online Competition between Pro- and Anti-vaccination Views,” *Nature*, 582(7811), 2020, 230–3.

- [23] K. Jordan and J. Pennebaker, “How the Candidates are Thinking: Analytic versus Narrative Thinking Styles,” *Retrieved January, 21, 2016, 2016.*
- [24] B. Kalsnes, “Fake News,” in *Oxford Research Encyclopedia of Communication*, 2018.
- [25] J. Keelan, V. Pavri, R. Balakrishnan, and K. Wilson, “An Analysis of the Human Papilloma Virus Vaccine Debate on MySpace Blogs,” *Vaccine*, 28(6), 2010, 1535–40.
- [26] M. Lahouati, A. De Coucy, J. Sarlangue, and C. Cazanave, “Spread of Vaccine Hesitancy in France: What about YouTube™?” *Vaccine*, 38(36), 2020, 5779–82.
- [27] C. Luo, A. Chen, B. Cui, and W. Liao, “Exploring Public Perceptions of the COVID-19 Vaccine Online from a Cultural Perspective: Semantic Network Analysis of Two Social Media Platforms in the United States and China,” *Telematics and Informatics*, 2021, 101712.
- [28] H. Lyu, J. Wang, W. Wu, V. Duong, X. Zhang, T. D. Dye, and J. Luo, “Social Media Study of Public Opinions on Potential COVID-19 Vaccines: Informing Dissent, Disparities, and Dissemination,” *Intelligent Medicine*, 2021.
- [29] E. Mathieu, H. Ritchie, E. Ortiz-Ospina, M. Roser, J. Hasell, C. Appel, C. Giattino, and L. Rodés-Guirao, “A Global Database of COVID-19 Vaccinations,” *Nature Human Behaviour*, 2021, 1–7.
- [30] K. Matsumoto, R. Fuji, and S. Kuroiwa, “An Algorithm for Estimating Human Emotions by Using Semantic Information of Words,” in *Proceeding of the Association for Natural Language Processing 11th Annual Conference*, 2005, 145–8.
- [31] A. Matsuo, B. Du, and K. Sasahara, “Appraisal of the Fairness Moral Foundation Predicts the Language Use Involving Moral Issues on Twitter Among Japanese,” *Frontiers in Psychology*, 12, 2021, 1235.
- [32] J. Mellon and C. Prosser, “Twitter and Facebook are not Representative of the General Population: Political Attitudes and Demographics of British Social Media Users,” *Research & Politics*, 4(3), 2017, 2053168017720008.
- [33] S. A. Memon and K. M. Carley, “Characterizing Covid-19 Misinformation Communities Using a Novel Twitter Dataset,” *arXiv preprint arXiv:2008.00791*, 2020.
- [34] A. Mislove, S. Lehmann, Y.-Y. Ahn, J.-P. Onnela, and J. Rosenquist, “Understanding the demographics of Twitter users,” in *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 5, No. 1, 2011.
- [35] T. Mitra, S. Counts, and J. W. Pennebaker, “Understanding Anti-vaccination Attitudes in Social Media,” in *Tenth International AAAI Conference on Web and Social Media*, 2016.

- [36] K. Miyazaki, T. Uchiba, K. Tanaka, and K. Sasahara, "Aggressive Behaviour of Anti-vaxxers and their Toxic Replies in English and Japanese," *Research Square* (<https://doi.org/10.21203/rs.3.rs-1289044/v1>), 2022.
- [37] K. Miyazaki, T. Uchiba, K. Tanaka, and K. Sasahara, "The Strategy Behind Anti-Vaxxers' Reply Behavior on Social Media," *arXiv preprint arXiv:2105.10319v2*, 2021.
- [38] R. Nishi, T. Takaguchi, K. Oka, T. Maehara, M. Toyoda, K.-i. Kawarabayashi, and N. Masuda, "Reply Trees in Twitter: Data Analysis and Branching Process Models," *Social Network Analysis and Mining*, 6(1), 2016, 26.
- [39] T. Nuzhath, S. Tasnim, R. K. Sanjwal, N. F. Trisha, M. Rahman, S. F. Mahmud, A. Arman, S. Chakraborty, and M. M. Hossain, "COVID-19 Vaccination Hesitancy, Misinformation and Conspiracy Theories on Social Media: A Content Analysis of Twitter Data," 2020.
- [40] B. Nyhan and J. Reifler, "When Corrections Fail: The Persistence of Political Misperceptions," *Political Behavior*, 32(2), 2010, 303–30.
- [41] J. W. Pennebaker, R. L. Boyd, K. Jordan, and K. Blackburn, "The Development and Psychometric Properties of LIWC2015," *tech. rep.*, 2015.
- [42] J. W. Pennebaker, C. K. Chung, J. Frazee, G. M. Lavergne, and D. I. Beaver, "When Small Words Foretell Academic Success: The Case of College Admissions Essays," *PloS One*, 9(12), 2014, e115844.
- [43] S. Praveen, R. Ittamalla, and G. Deepak, "Analyzing the Attitude of Indian Citizens Towards COVID-19 Vaccine—A Text Analytics Study," *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 15(2), 2021, 595–9.
- [44] H. Ritchie, E. Mathieu, L. Rodés-Guirao, C. Appel, C. Giattino, E. Ortiz-Ospina, J. Hasell, B. Macdonald, D. Beltekian, and M. Roser, "Coronavirus Pandemic (COVID-19)," *Our World in Data*, 2020.
- [45] M. Singh, R. Kaur, A. Matsuo, S. Iyengar, and K. Sasahara, "Morality-based Assertion and Homophily on Social Media: A Cultural Comparison between English and Japanese Languages," *arXiv preprint arXiv:2108.10643*, 2021.
- [46] V. Singh, R. Dasgupta, D. Sonagra, K. Raman, and I. Ghosh, "Automated Fake News Detection Using Linguistic Analysis and Machine Learning," in, 2017.
- [47] V. Suarez-Lledo and J. Alvarez-Galvez, "Prevalence of Health Misinformation on Social Media: Systematic Review," *Journal of Medical Internet Research*, 23(1), 2021, e17187.
- [48] Y. R. Tausczik and J. W. Pennebaker, "The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods," *Journal of Language and Social Psychology*, 29(1), 2010, 24–54.

- [49] N. Usher, J. Holcomb, and J. Littman, “Twitter Makes it Worse: Political Journalists, Gendered Echo Chambers, and the Amplification of Gender Bias,” *The International Journal of Press/Politics*, 23(3), 2018, 324–44.
- [50] D. Wawrzuta, M. Jaworski, J. Gotlib, and M. Panczyk, “What Arguments against COVID-19 Vaccines Run on Facebook in Poland: Content Analysis of Comments,” *Vaccines*, 9(5), 2021, 481.