


From Facebook to YouTube: The Potential Exposure to COVID-19 Anti-Vaccine Videos on Social Media

Social Media + Society
January-March 2023: 1–24
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20563051221150403
journals.sagepub.com/home/sms


Anatoliy Gruzd¹, Deena Abul-Fottouh²,
Melodie Yunju Song³, and Alyssa Saiphoo¹

Abstract

This article examines the role of Facebook and YouTube in potentially exposing people to COVID-19 vaccine-related misinformation. Specifically, to study the potential level of exposure, the article models a uni-directional information-sharing pathway beginning when a Facebook user encounters a vaccine-related post with a YouTube video, follows this video to YouTube, and then sees a list of related videos automatically recommended by YouTube. The results demonstrate that despite the efforts by Facebook and YouTube, COVID-19 vaccine-related misinformation in the form of anti-vaccine videos propagates on both platforms. Because of these apparent gaps in platform-led initiatives to combat misinformation, public health agencies must be proactive in creating vaccine promotion campaigns that are highly visible on social media to overtake anti-vaccine videos' prominence in the network. By examining related videos that a user potentially encounters, the article also contributes practical insights to identify influential YouTube channels for public health agencies to collaborate with on their public service announcements about the importance of vaccination programs and vaccine safety.

Keywords

misinformation, disinformation, vaccination, COVID-19, Facebook, YouTube, social media, recommender algorithm, social network analysis, infodemic

Introduction

SARS-CoV-2 was first identified in Wuhan, China, in December 2019. The World Health Organization (WHO, 2022) has since reported over 600 million cumulative globally confirmed cases and over 6.4 million confirmed deaths as of 4 September 2022. In a global effort to reduce COVID-19-related hospitalizations and deaths, many vaccines were developed and have been shown to be highly effective in preventing and reducing the severity of the disease (Centers for Disease Control and Prevention [CDC], 2022c). However, COVID-19 vaccine acceptance rates vary greatly regionally. A systematic review of peer-reviewed surveys shows that the acceptance rate of COVID-19 vaccines in some countries in the Middle East, Europe, Africa, and Russia is under 60% (Sallam, 2021). What is particularly problematic is that even in countries with an abundance of freely available COVID-19 vaccines, there is a persistently high prevalence of vaccine hesitancy. For instance, while most of the US population who is eligible (i.e., 5 years or older) has been vaccinated with at least one dose (83.8% as of 7 September 2022),

vaccine hesitancy is estimated to be as high as 26.7% in certain areas of the country (CDC, 2022a, 2022b).

Vaccine hesitancy is described as a “delay in acceptance or refusal of vaccination despite availability of vaccination services” (McDonald & SAGE Working Group on Vaccine Hesitancy, 2015). A leading cause of decline in vaccine coverage and outbreaks of vaccine preventable diseases, vaccine hesitancy burdens the healthcare system and health of the population (Hotez et al., 2020; Omer et al., 2012). Because of this, vaccine hesitancy has been named one of the top 10 global health threats (WHO, 2019). Vaccine hesitancy is a spectrum of beliefs, rather than a single one, stemming from doubtfulness toward vaccines. On one end of the spectrum are those who are

¹Toronto Metropolitan University, Canada

²McMaster University, Canada

³University of Toronto, Canada

Corresponding Author:

Anatoliy Gruzd, Social Media Lab, Ted Rogers School of Management, Toronto Metropolitan University, Toronto, Ontario M5B 2K3, Canada.
Email: gruzd@torontomu.ca



generally cautious about vaccination, mostly due to inconclusive or conflicting research evidence that is presented to them on the efficacy and safety of vaccines. On the other end are those individuals who believe in various conspiracy theories and may even participate in anti-vaccination protests. While the phenomenon of vaccine hesitancy is not unique to COVID-19 vaccines, its presence and impact are especially concerning during the ongoing COVID-19 pandemic.

Social media platforms, particularly those with large user bases, are most culpable of spreading vaccine-related misinformation that may contribute to vaccine hesitancy. In 2019, about 31 million people followed anti-vaccine groups on Facebook, generating as much as US\$989 million in revenue for Meta (Center for Countering Digital Hate, 2020). Vaccine-related misinformation on social media reduces public confidence in vaccines, leading to vaccine hesitancy (Carrieri et al., 2019; Pan American Health Organization, 2021). For example, in a randomized control trial, vaccine intent in participants in the United Kingdom and the United States declined by 6.2% and 6.4%, respectively, when exposed to social media posts containing COVID-19 vaccine-related misinformation (Loomba et al., 2021). In fact, those who relied the most on social media for information during the pandemic were more hesitant to get vaccinated (Lazer et al., 2021). In another study, parents who were exposed to vaccine-related misinformation on Facebook were 1.6 times more likely to perceive vaccines as unsafe (Tustin et al., 2018). Furthermore, between 5% and 30% of vaccine refusals in countries like the United States are estimated to be caused by vaccine-related misinformation. These refusals have caused approximately \$50–\$300 million worth of total estimated harm every day since May 2021 (Bruns et al., 2021).

COVID-19 vaccine-related misinformation has been shared widely since the start of the pandemic. In fact, claims about COVID-19 vaccines started circulating even before clinical trials of COVID-19 vaccines had begun, and have been on the rise since (Gruzd & Mai, 2021). Vaccine-related misinformation can take many forms. For instance, YouTube's Vaccine Misinformation policy recognizes and acts on the following types of vaccine-related misinformation (Google, 2021):

- *Vaccine safety*: content alleging that vaccines cause chronic side effects, outside of rare side effects that are recognized by health authorities;
- *Efficacy of vaccines*: content claiming that vaccines do not reduce transmission or contraction of disease;
- *Ingredients in vaccines*: content misrepresenting the substances contained in vaccines.

Similarly, Meta's Community Standards include established guidelines on the removal of false claims that may discourage vaccination, including the following (Meta, 2022):

- Claims that contribute to vaccine rejection (e.g., COVID-19 vaccines do not exist, are not approved by the FDA);
- Claims about the safety of COVID-19 vaccines (e.g., COVID-19 vaccines can kill you, lead to birth defects);
- Claims about the efficacy of COVID-19 vaccines (e.g., they increase the likelihood of getting sick);
- Claims about the development of the vaccine (e.g., they contain toxic ingredients, they are untested).

The definitions of vaccine-related misinformation used by both platforms include claims consistent with an anti-vaccine stance. This approach to defining vaccine misinformation is in line with the academic literature on this topic. For example, Amith and Tao (2018) developed the Vaccine Misinformation Ontology (VAXMO) where they placed the Anti-Vaccination Information concept as a subclass of the main Misinformation class. In turn, the Anti-Vaccination Information concept included the following subclasses: Vaccine Inefficacy, Alternative Medicine, Civil Liberties, Conspiracy Theories, Falsehoods, and Ideological. Similarly, in the work by Loomba et al. (2021), the researchers defined COVID-19 vaccine-related misinformation as “information questioning the importance or safety of a vaccine.” As such, for the rest of the article we use the term vaccine-related misinformation to refer to anti-vaccine claims like those listed above.

In the current work, we examine the role of social media platforms in exposing people to COVID-19 vaccine-related misinformation through videos on Facebook and YouTube—the two largest social media platforms in terms of user base and monthly total watch time (DataReportal, 2022). To study the potential level of exposure, we model a uni-directional information-sharing pathway shown in Figure 1: from when (1) a Facebook user encounters a vaccine-related post with a YouTube video, (2) follows this video to YouTube, and then (3) sees a list of related videos automatically recommended by YouTube. We are interested in examining this information-sharing pathway (from Facebook to YouTube's recommendations) because, on one hand, as discussed in section “Facebook,” Facebook is frequently used to share YouTube videos; on the other, automated recommendations are a significant driver of watch time on YouTube (Solsman, 2018; Zhou et al., 2010, 2016).

By examining what vaccine-related content is shared on Facebook and what additional vaccine-related videos are potentially recommended after viewing this content on YouTube, this article addresses an important gap in the literature on vaccine hesitancy and social media, as only 11% of the papers published in this area reviewed multiple social media platforms, with the majority of them (64%) examining “how do people talk about vaccines” as opposed to assessing the level of exposure to such content (Neff et al., 2021).

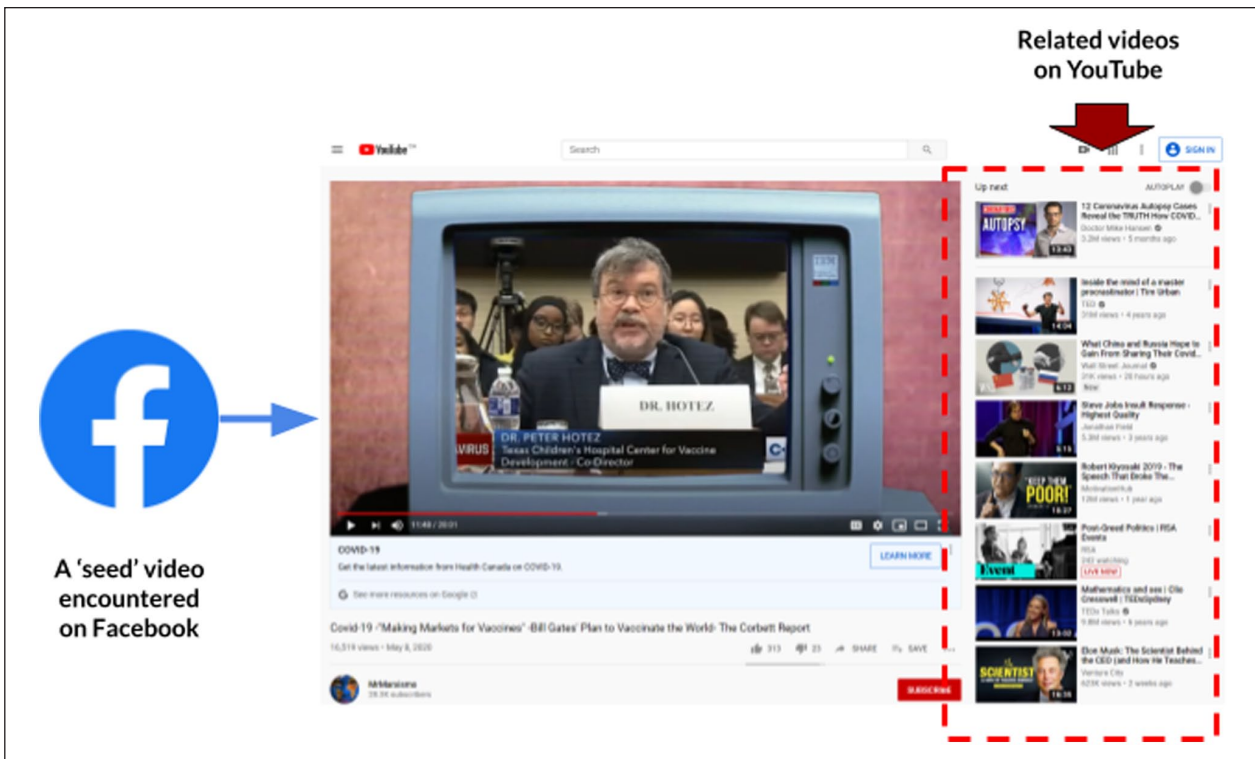


Figure 1. Social media user's information journey: from encountering a video link on Facebook to viewing related videos on YouTube.

Building on previous work that investigated the spread of vaccine-related misinformation on YouTube through video recommendations (Abul-Fottouh et al., 2020; Song & Gruzd, 2017; Tang et al., 2021), we contribute to the scholarship by starting with the examination of Facebook posts to discover vaccine-related “seed” videos that social media users might be exposed to, and then using these “seed” videos to find related videos as recommended by YouTube's Application Programming Interface (API) for developers. We chose Facebook as a starting point because this platform has been implicated as one of the main sources of YouTube videos containing vaccine-related misinformation (Knuutila et al., 2020).

This cross-sectional study differs from previous work in this area as data were collected during the first wave of the COVID-19 pandemic in June 2020. During this time, COVID-19 vaccines in development were still undergoing human trials. As a result, the efficacy of these vaccines was still unknown, which gave rise to misinformation and conspiracy theories on this topic. Coupled with the fact that during this period social media users were exposed to an excess of conflicting vaccine-related information and misinformation, the phenomenon known as “infodemic,” this makes it especially challenging to differentiate between facts and lies (Gruzd et al., 2021; Tangcharoensathien et al., 2020).

While the data for this study were collected in 2020, the findings are still relevant today. This is because even though

many social media platforms have vowed to fight COVID-19 misinformation, there are still many gaps in their misinformation policies, creating opportunities for vaccine-related misinformation to proliferate, as this study will demonstrate. These trends are largely driven by anti-vaccine groups, who find creative ways to bypass social media platform's automated labeling and manual fact-checking. Moreover, while we study misinformation related to COVID-19 vaccines here, our findings are relevant to vaccine misinformation in general.

Facebook and YouTube as Vectors of Vaccine-Related Misinformation

Social media platforms have emerged as major vectors of vaccine-related misinformation (Burki, 2020; Lou & Ahmed, 2019). In this section, we will discuss the role of two most popular social media platforms, Facebook and YouTube (Statista, 2021), in the spread of vaccine-related misinformation.

Facebook

While Facebook is the largest social media platform in the world, it is also one of the biggest sources of vaccine-related misinformation online (Silverman, 2016; Travers, 2020). Meta, the company behind Facebook, Instagram, and WhatsApp, has taken a number of steps to address COVID-19 misinformation

on its many platforms (Burki, 2020). One of the first steps the company took was to redirect users who encountered a COVID-19-related post to evidence-based information from the WHO and other health authorities (Jin, 2020). In October 2020, Facebook banned anti-vaccine advertisements (Brandom, 2020). A few months later in December 2020, the company announced they would increase efforts to remove COVID-19 vaccine-related misinformation from the platform and promote public health messaging for COVID-19 vaccines (Jin, 2020). Despite these efforts, it has been reported that between 41% (Avaaz, 2020) and 88% (Szeto et al., 2021) of COVID-19 misinformation, including vaccine-related misinformation, remained on the platform without a warning label.

Our first research question will assess the potential prevalence of vaccine-related misinformation on YouTube that is shared on Facebook. Building on the previous work which found that most vaccine-related misinformation on Facebook is shared by anti-vaccination Facebook groups (Johnson et al., 2016, 2019, 2020), we will identify public Facebook entities (i.e., groups and pages) that shared the most popular vaccine-related YouTube videos during the studied period (June 2020) and then will determine their overall vaccine stance (pro-, anti-, or neutral). Thus, our first research question is:

Research Question 1 (RQ1): What is the dominant vaccine stance of the most popular Facebook groups and pages that shared vaccine-related YouTube videos during the studied period?

The reason for examining the groups and pages that produced the most popular content, as opposed to those that were the most active, is because the content with the most engagement correlates with content visibility on the platform. In other words, content in less active groups and pages (i.e., those with a few anti-vaccination messages) with high engagement will be seen by many. In contrast, highly active groups or pages (i.e., those with hundreds of anti-vaccination messages) with low engagement would be less likely to contribute to vaccine hesitancy because no one or very few will see and engage with their content. To address RQ1, we manually reviewed and coded the most popular Facebook groups and pages in our dataset as “pro-vaccine,” “anti-vaccine,” “neutral,” or “not relevant.” In this investigation, videos with the most engagement (measured by the volume of likes, shares, and reactions) are defined as “popular.” As mentioned above, we label pages as “anti-vaccine” if they shared vaccine misinformation that may lead to vaccine hesitancy as defined by both Facebook and YouTube.

From Facebook to YouTube: The Spread of Vaccine-Related Misinformation Across Platforms

What makes it especially challenging when addressing vaccine-related misinformation is that false and misleading

claims are not constrained within a single platform. Something that is posted on one platform can be easily reshared across many others. Indeed, Knuutila et al. (2020) found that the most popular YouTube videos containing COVID-19 misinformation were often cross-posted on multiple social media platforms, and that Facebook frequently directed traffic to YouTube. A case in point is the conspiracy video “*Plandemic: The Hidden Agenda Behind COVID-19*” that promoted a wide range of debunked COVID-19 conspiracy theories, including claims that COVID-19 vaccines are ineffective and will “kill millions” (Plandemic Series, 2021). After its release on Facebook, the documentary spread quickly across multiple social media platforms. Before it was eventually removed by Facebook and many other social media platforms for violating misinformation policies, YouTube clips of the conspiracy video were shared 3.15 million times on Facebook, receiving 9.94 million comments on Twitter and 8.82 million reactions on Reddit (Frenkel et al., 2020). This demonstrates that despite the efforts of some social media platforms to purge COVID-19 misinformation in general and vaccine misinformation specifically, the ease and speed in which YouTube videos circulate across platforms may expose millions of users to harmful vaccine-related misinformation even if they are removed later. With this background in mind and building on RQ1, we ask:

Research Question 2 (RQ2): What is the dominant vaccine stance of YouTube videos shared on Facebook during the pandemic?

Anti-vaccine videos are often based on misinformation and are directly indicative of vaccine hesitancy (Donzelli et al., 2018). Thus, knowing if most vaccine-related YouTube videos that are shared on Facebook promote anti-vaccine stances would likely suggest the presence and prevalence of vaccine-related misinformation on Facebook. To address RQ2, we manually reviewed and coded all Facebook “seed” videos as “pro-vaccine,” “anti-vaccine,” or “neutral,” based on the content of the entire video. We label videos as anti-vaccine if they shared claims defined as misinformation by both Facebook and YouTube.

YouTube and Its Recommender Algorithm

As noted earlier, YouTube is another popular platform hosting vaccine-related misinformation. Shortly after YouTube launched in 2005, anti-vaccination organizations such as the National Vaccine Information Center (NVIC) and the Canary Party, who were previously spreading their messages through purchasable DVDs, quickly took to YouTube’s free uploading-streaming service. These organizations used YouTube to widely share anti-vaccine conference presentations, testimonials, and heavily edited court hearings that falsely claimed that measles, mumps, and rubella (MMR) vaccines cause autism (e.g., Fisher, 2009). As early as 2006, a third of the

vaccine-related content on the platform was classified as anti-vaccine (Keelan et al., 2007).

Like Facebook, YouTube also committed to addressing COVID-19 misinformation (Burki, 2020; Wetsman, 2020). For example, YouTube's official channel collaborated with the Vaccine Confidence Project to launch a campaign promoting evidence-based COVID-19 information (Graham, 2021; Robertson, 2021). They also released a policy to ban all anti-vaccination content starting in Fall 2021 (Reuters, 2021). However, despite these efforts, there are still concerns that COVID-19 misinformation (including anti-vaccine content) will remain on the platform. Based on the analysis of data collected using a browser extension called RegretsReporter, the Mozilla Foundation found that 20% of videos encountered by YouTube users contained some form of misinformation, and an additional 12% of videos were linked to COVID-19-related misinformation specifically (McCrosky & Geurkink, 2021). In another study, YouTube performed worse than Facebook, Instagram, and Twitter in content moderation, having removed only 34% of videos flagged as COVID-19 misinformation—and did this only after the investigative report went public (Szeto et al., 2021).

As mentioned, video recommendations on YouTube are a significant driver of watch time (Solsman, 2018), and thus a key factor when examining how COVID-19 misinformation reaches users on the platform. While we know that YouTube recommendations consider personal (e.g., past viewing behavior, subscription topic), external (e.g., seasonal interest, trending interest), and performance-based metrics (e.g., video quality, topic appeal), the specifics are unknown (Abul-Fottouh et al., 2020). In fact, YouTube's recommendation algorithm is often likened to a "black box" (Stokel-Walker, 2019).

Closely related to the focus of this study, previous work found that YouTube tends to recommend videos that share the same vaccine stance. For example, after watching a pro-vaccine video, users were more likely to be recommended more pro-vaccine videos; the same was true for anti-vaccine videos (Abul-Fottouh et al., 2020; Song & Gruzd, 2017; Tang et al., 2021). This creates a so-called "echo chamber," a phenomenon where individuals are constantly exposed only to messages that support their personal views (Cinelli et al., 2021). In the case of YouTube recommendations, viewers of anti-vaccine videos are less likely to be exposed to content countering their perspective (i.e., pro-vaccine videos). Instead, they were more likely to be exposed to additional anti-vaccine content on YouTube, which in turn may further strengthen their anti-vaccine beliefs (Allgaier, 2018; Moon & Lee, 2020). Other research has also shown the presence of "echo chambers" in YouTube recommendations when it comes to videos with conspiracy-related content, including content about alternative science and political conspiracies (Faddoul et al., 2020)—both categories in which anti-vaccine content is often present.

One of the main concerns with vaccine stance-driven "echo chambers" on YouTube is that a vaccine-hesitant user may start receiving anti-vaccine video recommendations after encountering and watching just one anti-vaccine video. The persistent exposure to a single point of view, especially if it is around anti-vaccination content based on misinformation, may lead to the "majority illusion" paradox (Lerman et al., 2016; Zhang & Centola, 2019): where the minority opinion is thought to be the majority, due to repeated and frequent exposure to minority opinion in a person's perspective network. This brings us to the final two questions. Considering YouTube's recent efforts to address vaccine-related misinformation, we ask:

Research Question 3 (RQ3): Regardless of the vaccine stance of a video being watched, is YouTube more likely to recommend pro-vaccine videos than anti-vaccine videos?

In recognition of the potential implications of vaccine stance divide due to "echo chamber" effects on YouTube, we ask:

Research Question 4 (RQ4): Based on the vaccine stance of a video being watched, is YouTube more likely to recommend videos with the same vaccine stance? Here, we are especially interested in knowing whether anti-vaccine YouTube videos are more likely to lead users (by means of automated related videos) to other anti-vaccine videos than to pro- or neutral-vaccine stance videos.

To answer RQs 3 and 4, we conducted social network analysis (SNA). We discuss the details of this analysis in the "Method" section.

Method

In this section, we will outline the data collection and analysis used in this study. To summarize, we first collected a dataset of Facebook posts that included at least one vaccine-related keyword and a link to a YouTube video ("seed" video). In this dataset, we identified the most popular groups and pages from which these posts originated, coded the overall stance of the group/page as pro-vaccine, anti-vaccine, or neutral (RQ1). Then, we coded each seed video as pro-vaccine, anti-vaccine, or neutral (RQ2). Second, using YouTube's API for developers, we collected a dataset of videos identified as related to each seed video. Locating related videos using this API offers a systematic way to study how YouTube might recommend videos to the average user, without accounting for the user's location, viewing history or other personalization settings. We then coded related videos as pro-vaccine, anti-vaccine, or neutral (RQ3). Third, and finally, using SNA techniques, we created and analyzed a network of seed videos and related videos (RQ3 and RQ4).

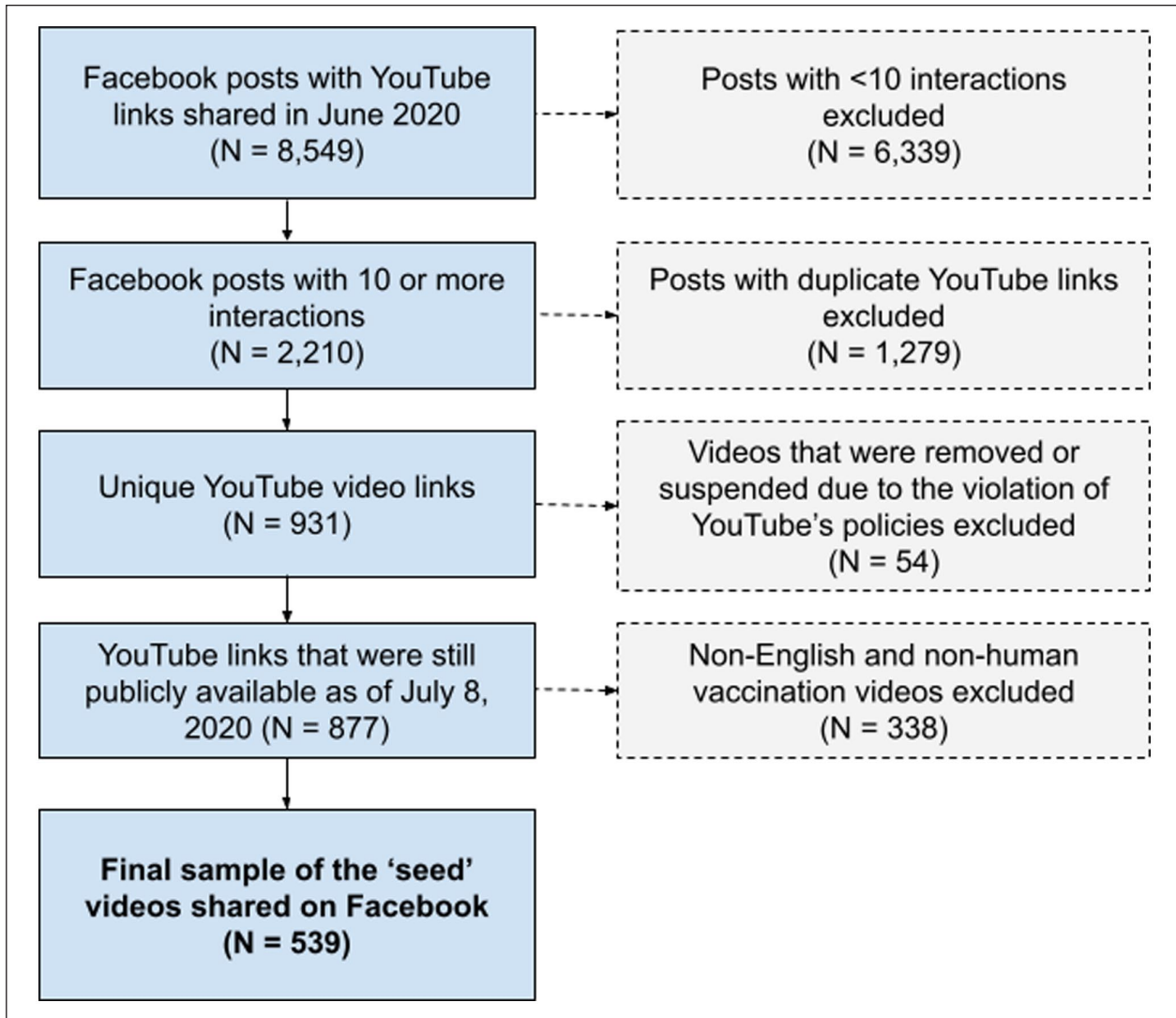


Figure 2. Data collection of seed vaccine-related videos shared on Facebook ($N=539$).

Data Collection: “Seed” Dataset of Popular YouTube Vaccine-Related Videos Shared on Facebook

The initial dataset of YouTube videos shared on Facebook (further referred to as “seed” videos) was developed using data retrieved from CrowdTangle, Meta’s platform that at the time of our research tracked publicly available posts shared on (1) Facebook public pages with more than 50K likes, (2) Facebook public groups with more than 95K members or US-based groups with more than 2K members, and (3) all verified profiles (Fraser, 2021). On 3 July 2020, we collected Facebook posts shared in June 2020 that (1) included a link to a YouTube video and (2) contained at least one vaccine-related keyword in the post or video description (e.g., *vaccine*, *vaccines*, *vaccination*, *vaxx*, *vaxxed*, or *immunization*). Relevant keywords were developed based on the previous

work in this area (Abul-Fottouh et al., 2020; Song & Gruzd, 2017; Tang et al., 2021). We deemed that including keywords specific to COVID-19 was redundant, since upon a manual review of sample posts most vaccine-related posts during the data collection period of June 2020 were about COVID-19 vaccines (as opposed to vaccines against other diseases).

Figure 2 summarizes our data cleaning and preparation steps. We started with a total of 8,549 videos posted across 4,453 Facebook pages or groups (further referred to as Facebook “entities”). To examine content with higher engagement, 6,339 posts (74%) with fewer than 10 interactions (i.e., reactions, comments, and shares) were excluded. We then extracted YouTube links in the remaining 2,210 Facebook posts. After excluding posts from different accounts sharing the same YouTube video, we were left with a dataset of 931 unique YouTube links. Even though data were collected only 3 days after the end of June, 54 YouTube

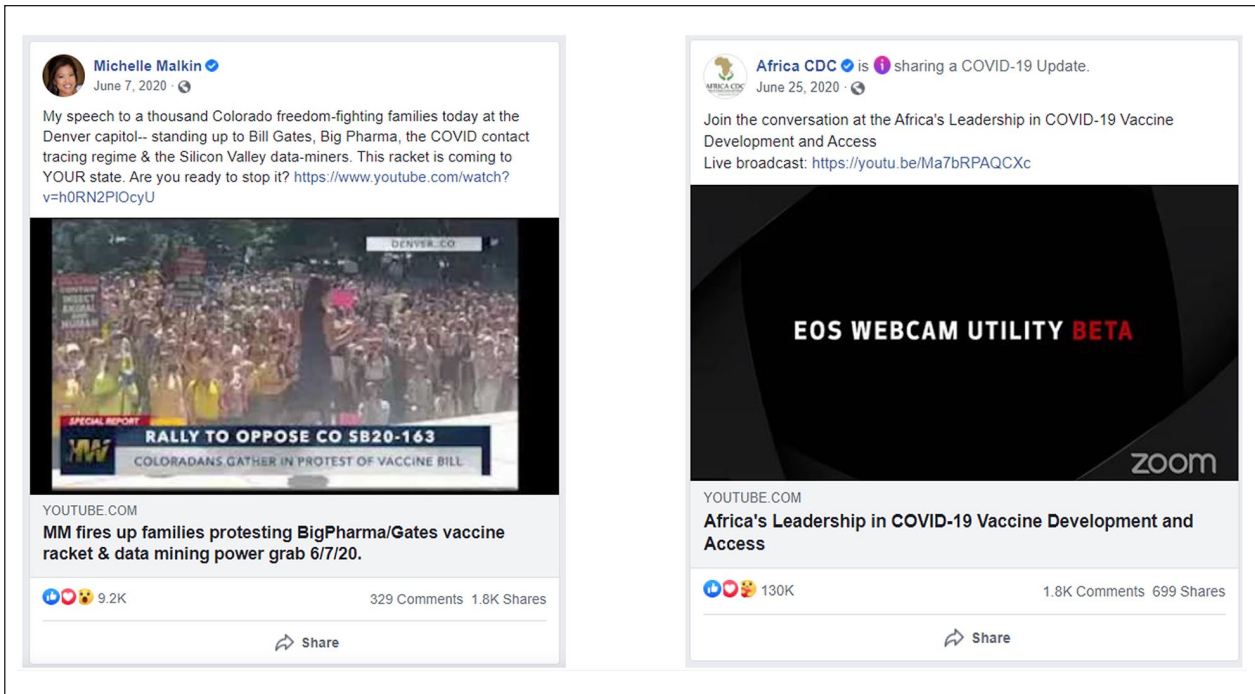


Figure 3. Sample Facebook posts linking to YouTube.

videos had already been removed by either the original poster or by YouTube. As a result, only 877 unique YouTube videos remained in the dataset. During the final data preparation step, we manually reviewed the remaining 877 videos to only include those that were in English and were indeed about human vaccination or immunization. The final “seed” dataset consisted of 539 vaccine-related English YouTube videos. Examples of seed videos and their original posts are shown in Figure 3.

Data Collection: YouTube’s Related Videos

To understand the composition of subsequent videos that users would likely be exposed to after watching COVID-19 vaccine-related YouTube videos that were shared on Facebook, we used YouTube’s API Related Videos call (“relatedToVideoId”) via YouTube Data Tools (Bernhard, 2015) to retrieve up to 50 related videos for each “seed” video based on their relevance. While the exact way of how YouTube recommends videos to a given user depends on a number of factors (including user’s location and viewing history), using the Related Videos API search allowed us to retrieve a pool of video candidates which are likely to be used by YouTube when recommending videos to individual users.

In total, we retrieved 19,083 videos and associated meta-data such as comment count, view count, dislike count, like count, published date, channel ID, and channel title. This step was conducted on 8 July 2020. We excluded related videos whose title, description, or transcript (if available) did not

contain vaccine-related keywords (immun*, vax*, vaccin*), resulting in a dataset of 3,058 videos. Because we were interested in examining only vaccine-related videos, we kept a final dataset of 2,260 English language videos that were manually classified as “pro-vaccine,” “anti-vaccine,” or “neutral.”

Vaccine Stance Coding

To answer RQ1, a sample of 56 Facebook entities (out of 4,453) were manually reviewed and coded as “pro-vaccine,” “anti-vaccine,” “neutral,” or “not relevant.” These were the most popular entities in our dataset based on the total number of interactions (including the number of reactions, comments, and shares) received across all their posts. In sum, these 56 entities shared posts that attracted 75% of all recorded interactions in our initial dataset of 8,549 posts shared by 4,453 Facebook entities.

To classify the most popular Facebook entities, three coders manually reviewed 1,161 posts shared by the 56 entities in the collected dataset. Fifty-eight non-English posts shared by 19 entities were automatically translated into English using Google Translate. After the first round of coding was done by a research assistant with a background in public health, the codes were cross-validated by two of the authors to ensure data quality. The codes were assigned at the entity level. Appendix A lists the examined entities and the resulting codes.

To answer RQs 2–4, 3,058 YouTube videos, including videos in the seed dataset and related videos, were watched

and coded as “pro-vaccine,” “anti-vaccine,” or “neutral.” During the coding, non-English language videos were excluded, which resulted in the final dataset of 2,260 seed and related videos. Pro-vaccine videos expressed support for COVID-19 vaccination, while anti-vaccine videos expressed attitudes of refusal or rejection toward vaccines. Neutral stance videos were neither supportive of nor opposed to vaccines (e.g., news media presenting two sides of the “debate”) (see Appendix B for coding instructions).

Due to the high volume and length of videos that required watching, nine coders participated in the coding process: one of the authors with vaccine-related misinformation expertise and eight research assistants in public health. The merged list of 3,058 videos was split into four batches of roughly equal size (765, 765, 766, 762 videos in each batch). Each of the eight research assistants was assigned to code one of the batches (in no particular order). The ninth coder, one of the authors, coded all 3,058 videos.

Two qualitative coding training sessions were held to ensure code consistency. To mitigate discrepancies in case a video was labeled differently, each video was watched by three coders. When there were coding disagreements, the majority rule was used to assign the final code. The intraclass correlation coefficients were above the recommended threshold of 0.7.

SNA

To answer RQs 3 and 4, we conducted SNA. We first used information provided by YouTube API via YouTube Data Tools (Bernhard, 2015) to create the YouTube’s Related Videos network consisting of 2,260 vaccine-related videos and 20,711 connections (see Figure 4). This is a “baseline” network of related videos (not affected by the user’s personal settings, previous watch history, etc.). We use this network to explore a universe of potential recommendations triggered by “seed” videos. In this network, each node represents a single video, and a connection from Video A to Video B means that after watching Video A, YouTube’s API is likely to recommend Video B as a related video.

To create the model for related videos, we used an SNA method called Latent Order Logistic (LOLOG) modeling (Fellows, 2018). We chose LOLOG for various reasons. First, this type of probability framework is designed to account for processes behind network growth over time, such as those commonly observed in online networks (Fellows, 2018). Second, LOLOG models are exponential family models with clearly defined model parameters whose coefficients can be easily interpreted in the same way logistic regression models are interpreted. Third, LOLOG models are similar to the popular Exponential-Family Random Graph Models (ERGMs; Hunter & Handcock, 2006), which rely on Monte Carlo methods and variational inference to describe probability distributions in large networks. However, the advantage of LOLOG models over ERGMs is

that they can model scale-free degree structures observed in networks, while avoiding problems of model degeneracy.

The dependent variable for the LOLOG model is the probability of forming a tie between Video A and Video B (expressed as the log of the odds) when accounting for (1) exogenous node-level factors such as vaccine stance expressed in a video and (2) endogenous network structural factors of the observed network such as in-degree centrality distribution. See Table 3 (Rows 1–14) for LOLOG model terms used in this study.

The model terms “Incoming Ties-Pro-Vaccine” (Row 3) and “Incoming Ties-Neutral” (Row 4) were examined to answer RQ3. Using anti-vaccine videos as the reference category, these terms are used to determine whether YouTube’s API is more likely to recommend anti-vaccine videos than pro-vaccine videos. A positive and significant p value (i.e., $p < .05$) for “Incoming Ties-Pro-Vaccine” would mean that pro-vaccine videos are more likely to be recommended than anti-vaccine videos (the reference category). Likewise, a positive and significant p value for “Incoming Ties-Neutral” would mean that neutral videos are more likely to be recommended than anti-vaccine videos. Two additional terms were included in the model to account for the number of outgoing ties based on vaccine stance: “Outgoing Ties-Pro-Vaccine” (Row 5) and “Outgoing Ties-Neutral” (Row 6). This is done mostly to ensure that the model has a good fit with the observed data and to account for the artifact of the data collection process, when YouTube Data Tools retrieved a prescribed number of related videos (up to 50) for each “seed” video. Furthermore, to account for online networks’ tendency to follow a power-law distribution, which manifests in a highly skewed, long-tail degree distribution pattern (Artico et al., 2020; S. Johnson et al., 2014; Kunegis et al., 2013), we added the following terms to the model: “Preferential Attachment” (Row 2), “Out-star2” (Row 12, term indicates two-star out-degree configurations), and “Out-star3” (Row 13, term indicates three-star out-degree configurations).

Next, we added two differential homophily terms to model community formation patterns in the network and check for presence of so-called “echo chambers.” Specifically, we used the “stance (Match)” term (Row 7) to determine whether videos with the same stance are likely to link to each other (RQ4). A positive and significant value ($p < .05$) for this term would confirm that there is a statistically significant tendency for videos expressing the same vaccine stance to link to one another.

In addition to modeling differential homophily based on vaccine stance, we included the “Video Category (Match)” term (Row 14) to account for the tendency of YouTube to recommend videos of the same topic (Cooper, 2021). The video category is selected by YouTube channel/user during the video upload process. Appendix D shows the counts of videos in the network by their user-assigned category. Five most popular categories in the dataset are: News & Politics (970 videos), Education (367), People & Blogs (278),

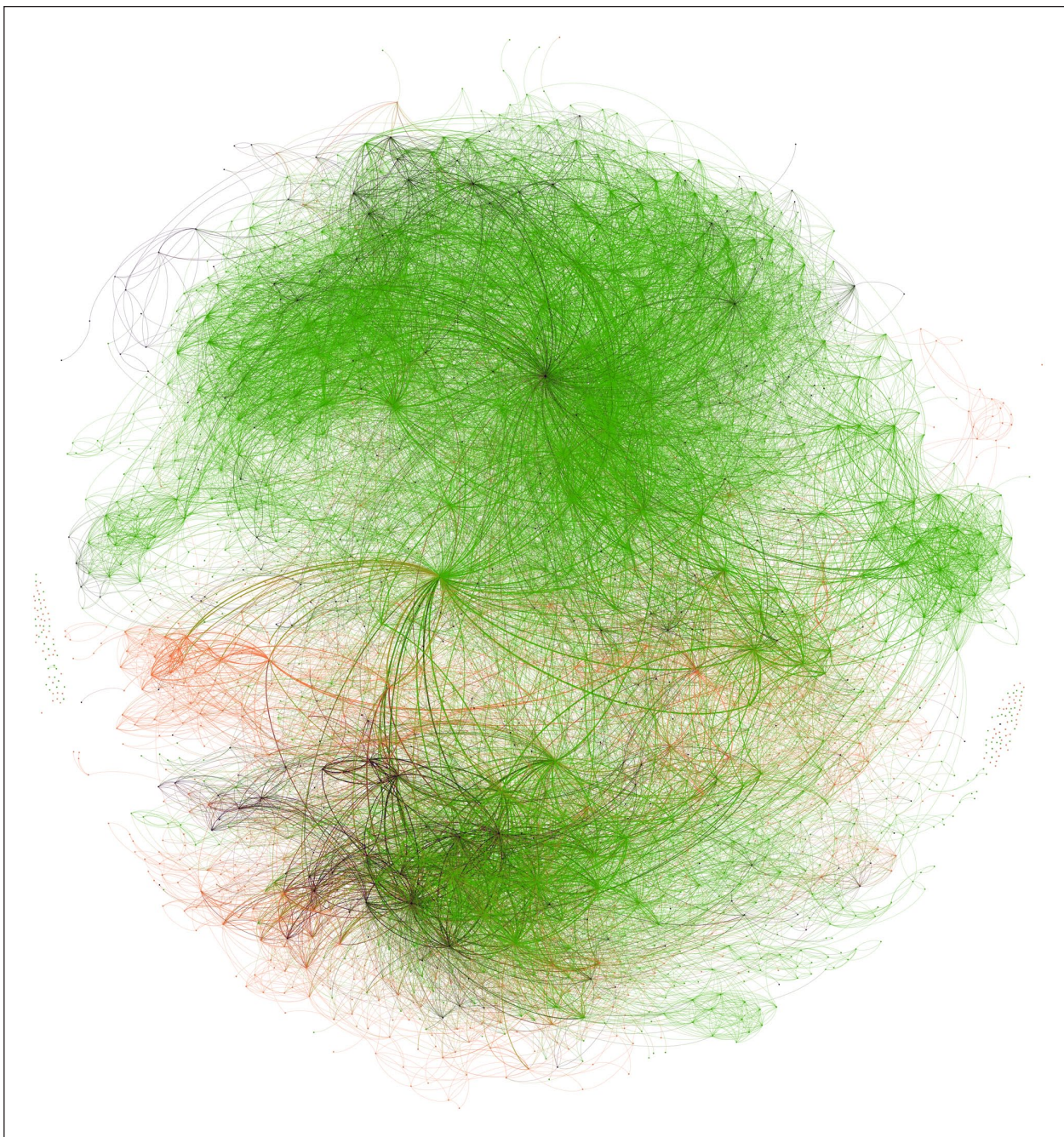


Figure 4. YouTube's Related Videos network of 2,260 vaccine-related videos. Pro-vaccine videos are shown in green, anti-vaccine videos are shown in red, neutral videos are shown in black.

Nonprofits & Activism (238), and Science & Technology (209).

Finally, to examine other possible types of homophily and clustering among nodes that are neither based on vaccine stances nor video categories, we added the following terms as proxies for community formation processes in the network (Morris et al., 2008; Sosa et al., 2015): “Reciprocity” (Row 8), “Two-path” (Row 9), “Triangles” (Row 10), and “Geometrically Weighted Distance Shared Partners” (Row 11).

Results

Vaccine-Related Content on Facebook (RQ1)

To answer RQ1, that is, to determine the dominant vaccine stance of the most popular public Facebook groups and pages that shared vaccine-related YouTube videos, we narrowed down our analysis to examine the content posted by 56 public Facebook entities (listed in Appendix A). Table 1 shows the results of the manual review and coding of these Facebook

Table 1. Top 56 public Facebook entities by vaccine stance.

Vaccine stance	Was the entity available as of 1 December 2021?		Total
	No	Yes	
Anti-vaccine	17 (45.9%)	20 (54.1%)	37
Pro-vaccine	1 ^a (12.5%)	7 (87.5%)	8
Neutral	0 (0%)	4 (100%)	4
Not relevant	1 (14.3%)	6 (85.7%)	7
Total	19 (33.9%)	37 (66.1%)	56

^aOne group that shared pro-vaccination content has been removed from Facebook potentially for sharing spam-type messages. As this group is no longer available, we are not able to confirm the actual reason for their removal.

entities, counting how many exhibited pro-vaccine, anti-vaccine, and neutral vaccine stances. The majority, 37 out of 56, promoted anti-vaccine content. Only eight shared pro-vaccine content. Four entities were neutral regarding vaccine stance, and seven were not related to the topic of human vaccination.

Furthermore, about a year after the data collection, nearly half (17 out of 37) of the anti-vaccine groups and pages were no longer available, suggesting that they have been suspended for circulating COVID-19 and/or vaccine-related misinformation. This included groups like “99% unite Main Group ‘it’s us or them’” and “FABUNAN ANTIVIRAL INJECTION SUPPORTERS GROUP,” and pages such as “Collective Action Against Bill Gates. We Wont Be Vaccinated!!” and “We Are Vaxxed.” While the remaining anti-vaccine groups and pages were still available as of December 2021, many of their posts that we collected were removed by the platform, the group moderators, or the original posters. Specifically, out of 98 posts shared by the 20 available anti-vaccine entities, only 33 posts (34%) are still publicly accessible.

Vaccine Stance of YouTube Videos (RQ2)

We coded all Facebook “seed” videos that produced vaccine-related videos ($N=484$) to answer RQ2: *What is the dominant vaccine stance of YouTube videos shared on Facebook during the pandemic?* The content analysis of YouTube seed videos shared on Facebook supports the results reported in the previous section. We found that anti-vaccine is the dominant stance on Facebook in this dataset, with nearly twice as many anti-vaccine seed videos (57.0%, $n=276$) than pro-vaccine videos (28.7%, $n=139$) (see Table 2). This finding in conjunction with the analysis of the most popular Facebook entities in RQ1 demonstrates that anti-vaccine content, shared on Facebook groups and pages, prevails over pro-vaccine content, before the platform takes action to remove this content.

Table 2. Vaccine stance coding of YouTube videos shared on Facebook.

Vaccine stance coding categories	YouTube videos shared on Facebook (“seed” videos)	
	Number of videos (n)	Percentage (%)
Pro-vaccine	139	28.7
Anti-vaccine	276	57.0
Neutral	22	4.5
Excluded		
Not relevant	19	3.9
Not in English	8	1.7
Not available	20	4.1
Total	484 ^a	100.0

^aOut of 539 seed videos, 484 were included in the final dataset. The excluded 55 seed videos did not generate vaccine-related video recommendations based on YouTube API, likely because the videos were permanently or temporarily suspended/unavailable during the data collection phase.

Effect of Video Stance on Video Recommendations (RQ3 and RQ4)

This section examines whether YouTube’s API is more likely to recommend pro-vaccine videos than anti-vaccine videos (RQ3), and whether YouTube is more likely to suggest related videos with the same vaccine stance (RQ4). Figure 4 shows the YouTube’s Related Videos network consisting of 2,260 vaccine-related videos and 20,711 connections. The figure shows clusters of videos sharing the same vaccine stance by color.

Table 3 shows the LOLOG terms used in the model as well as the corresponding network statistics $g(y)$, coefficients θ , standard errors, and p values. The probability, or more specifically the log of the odds, of observing a tie between two videos (the dependent variable in the model) is calculated based on the set of network statistics $g(y)$. The coefficients θ are interpreted in the same way as we would interpret coefficients for a logistic regression.

Regarding RQ3, our results show that even though there were nearly twice as many pro-vaccine videos in the YouTube’s Related Videos network (see Figure 4), pro-vaccine videos were not significantly more likely to be recommended by YouTube’s API than anti-vaccine videos (Table 3, Row 3, $\theta=0.69$, $p=.333$). At the same time, pro-vaccine videos were significantly more likely to produce ties to more videos than anti-vaccine videos (Table 3, Row 5, $\theta=1.00$, $p=.000$); in contrast, watching an anti-vaccine seed video would likely lead to a smaller network of related videos of any vaccine stance. This explains the reason that while anti-vaccine seed videos are more prevalent than pro-vaccine seed videos on Facebook, the Related Videos network on YouTube has conversely more pro-vaccine videos.

Regarding RQ4, we did not observe a homophily effect based on vaccine stance (Table 3, Row 7, $\theta=0.95$, $p=.058$).

Table 3. LOLOG model of YouTube’s Related Videos network.

LOLOG terms	Observed statistics $g(y)$	θ	SE	p
1. Edges	20,711	8.08	1.21	<.001**
2. Preferential Attachment	NA ^a	1.61	0.13	<.001**
3. Incoming Ties-Pro-Vaccine	2,450	0.69	0.72	.333
4. Incoming Ties-Neutral	2,327	-4.95	3.76	.188
5. Outgoing Ties-Pro-Vaccine	2,248	1.00	0.23	.000**
6. Outgoing Ties-Neutral	4,606	-0.11	0.62	.856
7. Stance (Match)	14,419	0.95	0.50	.058
8. Reciprocity	2,518	2.35	0.85	.006*
9. Two-path	377,270	0.54	0.23	.021*
10. Triangles	109,332	-0.74	0.22	.001**
11. Geometrically Weighted Distance Shared Partners (GWDSP 0.5)	153,863	-0.70	0.31	.022*
12. Out-star2	177,882	0.15	0.03	<.001**
13. Out-star3	1,384,269	-0.02	0.01	.024*
14. Video Category (Match)	9,676	-0.55	0.32	.084

LOLOG: Latent Order Logistic.

^aPreferential Attachment accounts for the unobserved order in which dyads were added to the network, hence the presence of NA (not available) under observed statistics (Fellows, 2018).

* p values < .05 are shown with an asterisk.

** p values < .001 are shown with two asterisks.

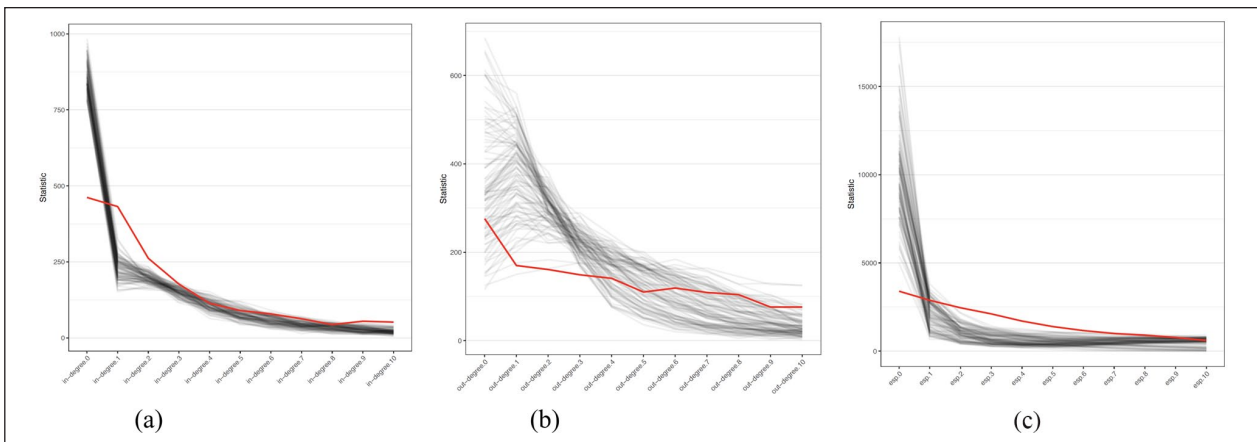


Figure 5. The in-degree (a), out-degree (b), and ESP distributions (c) of 100 simulated networks (black) from the fitted LOLOG compared to the observed network of related videos (red).

or video category (Table 3, Row 14, $\theta = -.055$, $p = .084$). These findings suggest that YouTube users are unlikely to become entrenched in a specific vaccine stance or video category.

All other endogenous factors of the network are significant, suggesting the presence of network closure (i.e., a tendency toward clustering). Furthermore, links to related videos appear to be reciprocal (Table 3, Row 8, $\theta = 2.35$, $p = .006$), meaning that if YouTube’s API recommends Video B based on Video A, it is 10.5 times more likely ($\exp[2.35]$) than by chance alone that Video B will also lead to Video A.

To validate the resulting model, we ran 100 simulations to compare various statistics between the simulated networks

and the observed network (Hunter et al., 2008). Figure 5 shows the goodness of fit for the degree distributions: (1) in-degree, (2) out-degree, and (3) edgewise shared partner (ESP) distribution of simulations of statistics from the fitted model compared to the observed network. These parameters were used to see how well the model fits patterns that are not explicitly represented by the terms in the model. The red line shows the observed network, and the black lines are the simulated ones. As the black and red lines overlap and generally follow each other’s pattern, the model shows a good fit with regard to these metrics.

The following section presents results of our ad hoc analysis to understand what causes some videos in the network to

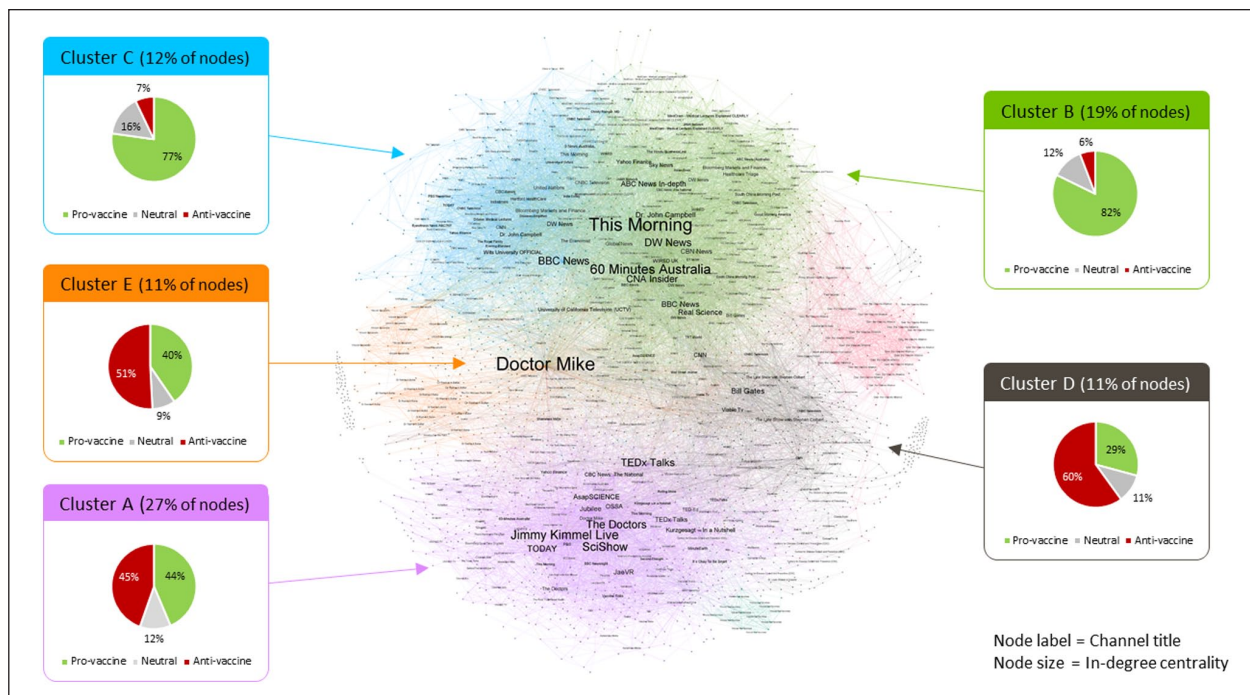


Figure 6. Clustering of videos in YouTube's network of related videos ($N=2,260$).

cluster, especially if it is neither due to their vaccine stance nor video category.

Examination of Factors Behind “Echo Chambers” in YouTube’s Network of Related Videos

To better understand the tendency of some videos to cluster together, we relied on a community detection algorithm called Fast Unfolding¹ (Blondel et al., 2008), as implemented in Gephi v0.9.2 (Bastian et al., 2009). The community detection algorithm allowed us to identify densely connected videos by partitioning the network into clusters in such a way that videos from the same cluster are more likely to relate to each other than to videos from other clusters. Once clusters were detected, we manually examined the most linked to YouTube videos and channels in each cluster to investigate why they were clustered (see Appendix C).

One of the main outputs of the community detection algorithm is the modularity value for the whole network. It is a value that ranges between 0 and 1, indicating how well a network can be partitioned into separate clusters based on Newman’s modularity class detection algorithm (Newman, 2006). Values closer to 1 suggest that a network consists of disconnected or loosely connected clusters. Values closer to 0 suggest the opposite; that is, the network predominantly consists of a single, well-connected cluster. The modularity value for our network was somewhere in the middle (0.549), indicating that while there was a strong overlap between clusters, we still observed well-defined clusters of densely connected nodes.

Each cluster is represented in a different color in Figure 6. We will focus on the five largest clusters that contain in total approximately 80% of the nodes ($n=1,804$) in the network. (Each top 5 cluster includes 5% or more of the nodes in the network.) We will further refer to them as Cluster A, B, C, D, and E. The remaining, smaller clusters only contained 2 to 22 videos each.

Based on the manual coding of vaccine stance for all videos in this network, we found that while Clusters B and C were predominantly pro-vaccine, and Cluster D was predominantly anti-vaccine, Clusters A and E contained about the same number of pro- and anti-vaccine videos. The prevalence of ties across videos with different stances in Clusters A and E (which contain 38% of nodes in the network) is likely why we did not observe a statistically significant homophily effect based on vaccine stance in the LOG model. To understand why there are ties between pro- and anti-vaccine videos in these two clusters, we examined the top 10 most linked to pro- and anti-vaccine videos (based on the videos’ in-degree centrality). Rather than examining each video independently, we examined the channels that shared the videos in question and their description. The results of this ad hoc evaluation are summarized in Appendix C.

Cluster A accounts for 27% of the videos collected ($n=605$). This cluster was mixed in terms of vaccine stance, with almost an equal amount of pro-vaccine (44% of nodes in the cluster) and anti-vaccine videos (45% of nodes in the cluster; the remaining 12% were neutral). Some of the most linked to pro-vaccine videos in this cluster were from popular educational channels that discuss dissenting opinions

related to highly contentious issues in medicine and scientific inquiry, the anti-vaccine movement and vaccine hesitancy included. Examples of such channels in the cluster include *SciShow*, *ASAPScience*, and *TEDxTalks*. Other highly linked to pro-vaccine videos in this cluster came from channels that post clips from cable-TV talk shows, such as *Jimmy Kimmel Live* and *The Doctors*. The majority of the frequently linked to anti-vaccine videos in this cluster were from highly sensational channels (e.g., *Thom Hartmann Program*, *CinemaLibre*, *The Real Truth About Health*). One channel, *Vaccine Risks*, features a series of lectures discussing the “risks” of vaccines—the risks that are discussed are highly exaggerated and littered with misinformation but are presented as facts. The series is narrated by Andrew Wakefield—a former physician in the United Kingdom whose medical license was revoked for falsely inferring MMR vaccines cause autism (Omer, 2020). Some of Andrew Wakefield’s anti-vaccine interviews from other channels (e.g., *Thom Hartmann Program*, *CinemaLibre*) were also highly linked to in this cluster.

Cluster E accounted for 11% of the videos in our network ($n=240$). Like Cluster A, this cluster had a mixture of pro and anti-vaccine videos (40% and 51%, respectively, with the remaining 9% of videos being neutral). Highly linked to videos in this cluster—both pro- and anti-vaccine—were related to the COVID-19 conspiracy theory film: “Plandemic: The hidden agenda behind COVID-19.” For example, the most linked to pro-vaccine video was from *Doctor Mike*, a celebrity doctor on YouTube who explains and debunks vaccine-related misinformation discussed in the documentary. On the other hand, the most linked to anti-vaccine videos also talked about “Plandemic.” Conspiracy theorist Rashid Buttar’s channel is one such example: he was interviewed in the *Plandemic* series and promoted the film heavily.

The above shows that even though videos in Clusters A and E are mixed in terms of their vaccine stance, they are homogeneous in other respects. For instance, Cluster A is formed based on videos (of all stances) shared by popular YouTube channels with a primary focus that is not on vaccines (such as channels for TV shows like *Jimmy Kimmel* and *The Doctors*, pop culture channels like *Rolling Stone*, and channels run by online influencers like *Shameless Maya*). Cluster E features videos, of all vaccine stances, that are about or related to the *Plandemic* documentary.

Discussion and Conclusion

Vaccine-related misinformation on social media can reduce social progress and public confidence in vaccines (Center for Countering Digital Hate, 2021), delaying global herd immunity during a pandemic. When weighing trade-offs between curtailing misinformation and restricting freedom of speech, one must consider the burden of disease caused by misinformation and subsequent infections caused by vaccine refusal, such as the cost of out-patient and in-patient care, as well as

the long-term treatment costs for patients with persistent symptoms post-discharge.

Our findings that the majority of the most viral entities on Facebook (66%, 37 out of 56) promoted anti-vaccine videos (RQ1) and that over 50% of YouTube videos (57%, 276 out of 484) shared on Facebook were anti-vaccine (RQ2) are significant in two ways. First, it highlights that Facebook falls short of their declared goals to keep the platform free from COVID-19 and vaccine-related misinformation. This is echoed by other research which revealed that the platform’s regulation policy only moderately impacted posts and endorsements of anti-vaccine content on Facebook (see, for example, Gu et al., 2022). Though Facebook announced their decision to ban anti-vaccine content in February 2021 and have since claimed to remove 3,000 accounts and 20 billion pieces of anti-vaccine content worldwide (Bickert, 2021), some of the most prominent anti-vaccine groups (e.g., Children’s Health Defense, Natural News) and personalities (e.g., Joseph Mercola and Robert Kennedy Jr.) continue to have a presence on the platform.

Second, the high prevalence of pro-vaccine videos in the YouTube Related Videos network demonstrates that YouTube may be effectively removing vaccine-related misinformation from their platform. This is consistent with a released statement of their commitment to address COVID-19 and general vaccine-related misinformation (YouTube, 2021). This finding is in line with another study which found a reduction in vaccine-related misinformation on the platform (Hussein et al., 2020). However, while there are more pro-vaccine videos in the network, pro-vaccine videos are not any more likely to be linked to than anti-vaccine videos (RQ3). In addition, even though the LOLOG model did not confirm a statistically significant homophily effect based on vaccine stance (RQ4), some anti-vaccine videos formed a potential “echo chamber” in YouTube’s Related Videos network. Specifically, we found a cluster of densely connected videos (labeled as D in Figure 6) with predominantly anti-vaccine stance (60%). This cluster primarily consisted of videos related to Bill Gates. Many anti-vaccine videos in this cluster were heavily edited clips of interviews with Bill Gates that were taken out of context to support various conspiracy theories, such as claims of his conspiring a global genocide by microchipping and killing people with COVID-19 vaccines.

The finding that pro-vaccine videos may sometimes relate to anti-vaccine videos, and vice versa, suggests that YouTube users may be exposed to vaccine viewpoints that are opposite to their own. This has both positive and negative implications. On one hand, some studies point to examples of vaccine skeptics declaring positive vaccine intentions after being exposed to information about the risk of communicable diseases (Thaker & Subramanian, 2021); on the other hand, some studies show that people who were pro-vaccine would reduce their intention to get vaccinated after being exposed to COVID-19 vaccine-related misinformation (Loomba et al., 2021). Though it is out of the scope of our

investigation to study changes in vaccine belief, robust experimental studies in social psychology are underway to examine the effectiveness of “inoculating” and “prebunking.” By examining the order of information exposure (e.g., pro-vaccine exposure first, anti-vaccine exposure to follow) and the medium (e.g., imagery, video, and text) in which vaccine information is presented, we may begin to unravel the complex causal pathways that influence or entrench one’s vaccine beliefs (Lewandowsky & van der Linden, 2021).

Taken together, the results demonstrate that despite the efforts by Facebook and YouTube, COVID-19 vaccine-related misinformation in the form of anti-vaccine content finds a way to propagate, and in some cases such content may even be amplified by YouTube through their automated content recommendations. Future research might need to look at techniques used by YouTube channels to bypass the platform’s misinformation policies.

Because of the apparent gaps in the platform-led initiatives to combat misinformation, public health agencies must be proactive in making sure that their public service announcements about the importance of vaccination programs and vaccine safety are highly visible on social media to reach the right audience. By examining how YouTube’s API finds related videos, our study contributes important insights that can be used to identify potential partners for public health agencies to collaborate with. For example, by reviewing the most linked to videos in section “Examination of factors behind ‘echo chambers’ in YouTube’s network of related videos,” we observed that popular YouTube channels have something in common. They often have many subscribers, post engaging and current content, and use search engine-friendly keywords in titles and descriptions. While public health agencies might not have enough resources to develop a strong following base on social media and create viral content, they may partner with marketing firms and influencers to deliver public service announcements to a wider audience. Another avenue is to partner with more traditional news media organizations, such as DW News, BBC News, Channel NewsAsia (CNA) Insider, CNN, and MSNBC, as their YouTube channels produced some of the most successful pro-vaccine contents on YouTube (see Clusters B and C in Figure 6). In addition, if the goal is to reach vaccine-hesitant audiences, public health agencies may try to work with popular educational channels (such as *SciShow*, *ASAPScience*, and *TEDxTalks*) and talk shows (such as *Jimmy Kimmel Live* and *The Doctors*) that have shown to be highly linked to by YouTube and frequently cross-linked with anti-vaccine videos in Cluster A.

In this article, we followed the potential path of COVID-19 vaccine misinformation across Facebook and YouTube. We found that while social media platforms have committed to purging harmful material, anti-vaccine content remains active and relevant, thus representing a challenge to public health efforts in fighting the pandemic. On the brighter side, we found that while there was more anti-vaccine content in

the original videos collected from Facebook, this content is not likely to take the average viewer into a rabbit hole of anti-vaccine content on YouTube.

Study Limitations and Future Directions

We would like to conclude this article with a discussion of the limitations of the current analysis which also inform directions for future work. First, one inclusion criterion for the seed video dataset, as described in the “Method” section, was the inclusion of at least one vaccine-related keyword in English. Thus, our findings are not generalizable to non-English videos. Recent work lends evidence to a convergence across languages insofar as the presence and salience of anti-vaccine videos promulgated via video channel linkages and video recommendations on YouTube and Facebook groups (Donzelli et al., 2018; Tokojima Machado et al., 2020). Though these studies investigated cross-country misinformation transfer across platforms (Bridgman et al., 2021), less focus is placed on inter-language information transfer and its causal pathways. This is a gap in contemporary misinformation research to be filled by future research.

Second, although this dataset does not include all YouTube vaccine-related videos cross-posted to Facebook, our finding that pro-vaccine videos reside in predominantly large clusters but with anti-vaccine videos accounting for much of the smaller clusters suggests that platform self-regulation has much room for improvement. Experimental studies in physics have proposed an “R-nought” criterion that may prevent information contagions from perforating and spreading system-wide (Xu et al., 2022). Future studies may address this limitation by reproducing our data collection methodology and reporting the temporal changes in pro- and anti-vaccine stance cluster characteristics over time.

Third, there are factors other than vaccine stance (studied here) that may influence YouTube’s recommendation algorithm, such as the channel’s number of subscribers, video length, and date of upload. We recommend that future research use other methods, like the ternary interaction item recommendation model (TIIREC; Yu et al., 2017), that can account for these factors. Relatedly, factors linked to individual users may also influence the algorithm, such as watch history and location. We were unable to account for these factors in our analysis as we used YouTube’s API for developers to collect related videos and reconstruct a generic network of vaccine-related videos. To address this limitation and to validate our results, an interesting and necessary future direction would be to collect recommendations based on personal accounts on YouTube, similar to the approach used by Lall et al. (2020) who used Amazon’s Mechanical Turk Platform to collect watch data from individual users.

Fourth, our results and conclusions apply to the network of videos that we collected, during the time period

of interest. This network is a sample of all vaccine-related videos retrieved via YouTube's Related Videos API. This is a common method for data collection from social media in general and YouTube in particular (e.g., Abul-Fottouh et al., 2020; Kaiser et al., 2021; Röchert et al., 2020; Tang et al., 2021). While we are confident in the comprehensiveness of our dataset, we acknowledge that we have not included all vaccine-related YouTube videos cross-posted on Facebook. Future studies may increase the generalizability of the results by reproducing our data collection methodology and validating our findings using additional datasets.

Finally, we investigate potential exposure to the videos in our study, not actual exposure (i.e., we did not ask users if they had seen certain videos), nor impact (e.g., belief in the content they were exposed to), nor behavior (i.e., whether exposure to these videos affected their decision to be vaccinated). Though questions related to these metrics were not the focus of our study, they would certainly make interesting future directions for research.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.


Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Alliance de recherche numérique du Canada PI: Gruzd Government of Canada > Canadian Institutes of Health Research PIs: Veletsianos, Hodson, Gruzd Government of Canada > Natural Sciences and Engineering Research Council of Canada PI: Gruzd.

ORCID iDs

Anatoliy Gruzd  <https://orcid.org/0000-0003-2366-5163>

Melodie YunJu Song  <https://orcid.org/0000-0002-3492-6835>

Alyssa Saiphoo  <https://orcid.org/0000-0002-7942-9285>

Note

1. To run the Fast Unfolding algorithm, we kept the default value for the resolution parameter as 1 (Lambiotte et al., 2014) and indicated that we want to incorporate edge weights into calculation, which in our case represent how many times Video A recommended Video B.

References

- Abul-Fottouh, D., Song, M. Y., & Gruzd, A. (2020). Examining algorithmic biases in YouTube's recommendations of vaccine videos. *International Journal of Medical Informatics*, 140, Article 104175. <https://doi.org/10.1016/j.ijmedinf.2020.104175>
- Allgaier, J. (2018). Science and medicine on YouTube. In J. Hunsinger, L. Klastrup, & M. M. Allen (Eds.), *Second international handbook of internet research* (pp. 1–21). Springer Netherlands. https://doi.org/10.1007/978-94-024-1202-4_1-1
- Amith, M., & Tao, C. (2018). Representing vaccine misinformation using ontologies. *Journal of Biomedical Semantics*, 9(1), 1–13.
- Artico, I., Smolyarenko, I., Vinciotti, V., & Wit, E. C. (2020). How rare are power-law networks really? *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 476(2241), Article 20190742. <https://doi.org/10.1098/rspa.2019.0742>
- Avaaz. (2020). *How Facebook can flatten the curve of the coronavirus infodemic*.
- Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. *Proceedings of the International AAAI Conference on Web and Social Media*, 3(1), Article 1. <https://ojs.aaai.org/index.php/ICWSM/article/view/13937>
- Bernhard, R. (2015). *YouTube data tools* (Version 1.22) [Computer software]. <https://tools.digitalmethods.net/netvizz/youtube/>
- Bickert, M. (2021, August 18). How we're taking action against vaccine misinformation superspreaders. *Meta*. <https://about.fb.com/news/2021/08/taking-action-against-vaccine-misinformation-superspreaders/>
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>
- Brandom, R. (2020, October 13). Facebook announces ban on anti-vaccination ads. *The Verge*. <https://www.theverge.com/2020/10/13/21514535/facebook-anti-vaxx-ad-ban-moderation-covid-19-vaccine>
- Bridgman, A., Merkley, E., Zhilin, O., Loewen, P. J., Owen, T., & Ruths, D. (2021). Infodemic pathways: Evaluating the role that traditional and social media play in cross-national information transfer. *Frontiers in Political Science*, 3, Article 648646. <https://www.frontiersin.org/article/10.3389/fpos.2021.648646>
- Bruns, R., Hosangadi, D., Trotochaud, M., & Kirk Sell, T. (2021). *COVID-19 vaccine misinformation and disinformation costs an estimated \$50 to \$300 million each day*. Johns Hopkins Center for Health Security. <https://www.centerforhealthsecurity.org/our-work/publications/covid-19-vaccine-misinformation-and-disinformation-costs-an-estimated-50-to-300-million-each-day>
- Burki, T. (2020). The online anti-vaccine movement in the age of COVID-19. *The Lancet Digital Health*, 2(10), e504–e505. [https://doi.org/10.1016/S2589-7500\(20\)30227-2](https://doi.org/10.1016/S2589-7500(20)30227-2)
- Carrieri, V., Madio, L., & Principe, F. (2019). Vaccine hesitancy and (fake) news: Quasi-experimental evidence from Italy. *Health Economics*, 28(11), 1377–1382. <https://doi.org/10.1002/hec.3937>
- Center for Countering Digital Hate. (2020). *Failure to act: How tech giants continue to defy calls to rein in vaccine misinformation*. <https://counterhate.com/wp-content/uploads/2022/05/201201-Failure-to-Act.pdf>
- Center for Countering Digital Hate. (2021). *The Disinformation Dozen: Why platforms must act on twelve leading online anti-vaxxers*. <https://www.counterhate.com/disinformationdozen>
- Centers for Disease Control and Prevention. (2022a). *COVID data tracker*. <https://covid.cdc.gov/covid-data-tracker>
- Centers for Disease Control and Prevention. (2022b). *Estimates of vaccine hesitancy for COVID-19*. <https://data.cdc.gov/stories/s/Vaccine-Hesitancy-for-COVID-19/cnd2-a6zw/>
- Centers for Disease Control and Prevention. (2022c). *COVID-19 Vaccine Effectiveness*. <https://www.cdc.gov/coronavirus/2019-ncov/vaccines/effectiveness/index.html>

- Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociochi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences of the United States of America*, 118(9), e2023301118. <https://doi.org/10.1073/pnas.2023301118>
- Cooper, P. (2021, June 21). How does the YouTube algorithm work in 2021? The complete guide. *Social Media Marketing & Management Dashboard*. <https://blog.hootsuite.com/how-the-youtube-algorithm-works/>
- DataReportal. (2022). *Digital 2022: Global digital overview*. https://datareportal.com/?utm_source=Statista&utm_medium=Data_Citation_Hyperlink&utm_campaign=Data_Partners&utm_content=Statista_Data_Citation
- Donzelli, G., Palomba, G., Federigi, I., Aquino, F., Cioni, L., Verani, M., Carducci, A., & Lopalco, P. (2018). Misinformation on vaccination: A quantitative analysis of YouTube videos. *Human Vaccines & Immunotherapeutics*, 14(7), 1654–1659. <https://doi.org/10.1080/21645515.2018.1454572>
- Faddoul, M., Chaslot, G., & Farid, H. (2020). A Longitudinal Analysis of YouTube’s Promotion of Conspiracy Videos. *ArXiv:2003.03318* [Cs]. <http://arxiv.org/abs/2003.03318>
- Fellows, I. E. (2018). A new generative statistical model for graphs: The Latent Order Logistic (LOLOG) model. *arXiv*. <http://arxiv.org/abs/1804.04583>
- Fisher, B. L. (2009). *NVIC’s October 2009 vaccine conference: It’s about informed choices—NVIC newsletter*. National Vaccine Information Center. <https://www.nvic.org/NVIC-Vaccine-News/May-2009/Thursday,-May-14,-2009-NVIC-s-October-2009-Vaccine.aspx>
- Fraser, L. (2021). *What data is CrowdTangle tracking?* CrowdTangle. <http://help.crowdtangle.com/en/articles/1140930-what-data-is-crowdtangle-tracking>
- Frenkel, S., Decker, B., & Alba, D. (2020, May 20). How the “pandemic” movie and its falsehoods spread widely online. *The New York Times*. <https://www.nytimes.com/2020/05/20/technology/plandemic-movie-youtube-facebook-coronavirus.html>
- Google. (2021). *COVID-19 medical misinformation policy—YouTube help*. <https://support.google.com/youtube/answer/11161123>
- Graham, G. (2021). *COVID-19 vaccines: Get back to what you love* [Blog]. YouTube. <https://blog.youtube/news-and-events/covid-19-vaccines-get-back-what-you-love/>
- Gruzd, A., De Domenico, M., Sacco, P. L., & Briand, S. (2021). Studying the COVID-19 infodemic at scale. *Big Data & Society*, 8(1), 20539517211021116. <https://doi.org/10.1177/20539517211021115>
- Gruzd, A., & Mai, P. (2021). *COVID-19 misinformation portal—A rapid response project from the Social Media Lab at Toronto Metropolitan University*. <https://covid19misinfo.org/about-page/>
- Gu, J., Dor, A., Li, K., Broniatowski, D. A., Hatheway, M., Fritz, L., & Abrams, L. C. (2022). The impact of Facebook’s vaccine misinformation policy on user endorsements of vaccine content: An interrupted time series analysis. *Vaccine*, 40(14), 2209–2214. <https://doi.org/10.1016/j.vaccine.2022.02.062>
- Hotez, P. J., Nuzhath, T., & Colwell, B. (2020). Combating vaccine hesitancy and other 21st century social determinants in the global fight against measles. *Current Opinion in Virology*, 41, 1–7. <https://doi.org/10.1016/j.coviro.2020.01.001>
- Hunter, D. R., & Handcock, M. S. (2006). Inference in curved exponential family models for networks. *Journal of Computational and Graphical Statistics*, 15(3), 565–583. <https://doi.org/10.1198/106186006X133069>
- Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., & Morris, M. (2008). ergm: A package to fit, simulate and diagnose exponential-family models for networks. *Journal of Statistical Software*, 24(3), nihpa54860.
- Hussein, E., Juneja, P., & Mitra, T. (2020). Measuring misinformation in video search platforms: An audit study on YouTube. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW1), 048:1–048:27. <https://doi.org/10.1145/3392854>
- Jin, K.-X. (2020, December 18). Keeping people safe and informed about the coronavirus. *Facebook*. <https://about.fb.com/news/2020/12/coronavirus/>
- Johnson, N. F., Leahy, R., Restrepo, N. J., Velasquez, N., Zheng, M., Manrique, P., Devkota, P., & Wuchty, S. (2019). Hidden resilience and adaptive dynamics of the global online hate ecology. *Nature*, 573(7773), 261–265. <https://doi.org/10.1038/s41586-019-1494-7>
- Johnson, N. F., Velásquez, N., Restrepo, N. J., Leahy, R., Gabriel, N., El Oud, S., Zheng, M., Manrique, P., Wuchty, S., & Lupu, Y. (2020). The online competition between pro- and anti-vaccination views. *Nature*, 582(7811), 230–233. <https://doi.org/10.1038/s41586-020-2281-1>
- Johnson, N. F., Zheng, M., Vorobyeva, Y., Gabriel, A., Qi, H., Velasquez, N., Manrique, P., Johnson, D., Restrepo, E., Song, C., & Wuchty, S. (2016). New online ecology of adversarial aggregates: ISIS and beyond. *Science*, 352(6292), 1459–1463. <https://doi.org/10.1126/science.aaf0675>
- Johnson, S., Faraj, S., & Kudaravalli, S. (2014). Emergence of power laws in online communities: The role of social mechanisms and preferential attachment. *Management Information Systems Quarterly*, 38(3), 795–808.
- Kaiser, J., Rauchfleisch, A., & Córdova, Y. (2021). Comparative approaches to mis/disinformation| fighting Zika with honey: An analysis of YouTube’s video recommendations on Brazilian YouTube. *International Journal of Communication*, 15, 1244–1262. <https://ijoc.org/index.php/ijoc>
- Keelan, J., Pavri-Garcia, V., Tomlinson, G., & Wilson, K. (2007). YouTube as a source of information on immunization: A content analysis. *JAMA*, 298(21), 2482–2484. <https://doi.org/10.1001/jama.298.21.2482>
- Knuutila, A., Herasimenka, A., Au, H., Bright, J., Nielsen, R., & Howard, P. N. (2020). *COVID-related misinformation on YouTube* [Report]. University of Oxford. https://canucklaw.ca/wp-content/uploads/2020/09/covid.related.misinformation.on_youtube.pdf
- Kunegis, J., Blattner, M., & Moser, C. (2013). Preferential attachment in online networks: Measurement and explanations. In *Proceedings of the 5th annual ACM web science conference* (pp. 205–214). <https://doi.org/10.1145/2464464.2464514>
- Lall, S., Agarwal, M., & Sivakumar, R. (2020). A YouTube dataset with user-level usage data: Baseline characteristics and key insights. In *ICC 2020-2020 IEEE international conference on communications (ICC)* (pp. 1–7). Institute of Electrical and Electronics Engineers. <https://doi.org/10.1109/ICC40277.2020.9148782>
- Lambiotte, R., Delvenne, J.-C., & Barahona, M. (2014). Laplacian dynamics and multiscale modular structure in networks. *IEEE Transactions on Network Science and Engineering*, 1(2), 76–90. <https://doi.org/10.1109/TNSE.2015.2391998>

- Lazer, D., Green, J., Ognyanova, K., Baum, M. A., Kin, J., Druckman, J., Perlis, R. H., Santilana, M., Simonson, M., & Uslu, A. (2021). *Social media news consumption and COVID-19 vaccination rates*. The COVID States Project. <https://www.covidstates.org/reports/social-media-news-consumption-and-covid-19-vaccination-rates>
- Lerman, K., Yan, X., & Wu, X.-Z. (2016). The “majority illusion” in social networks. *PLoS ONE*, *11*(2), Article e0147617. <https://doi.org/10.1371/journal.pone.0147617>
- Lewandowsky, S., & van der Linden, S. (2021). Countering misinformation and fake news through inoculation and prebunking. *European Review of Social Psychology*, *32*(2), 348–384. <https://doi.org/10.1080/10463283.2021.1876983>
- Loomba, S., de Figueiredo, A., Piatek, S. J., de Graaf, K., & Larson, H. J. (2021). Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nature Human Behaviour*, *5*(3), 337–348. <https://doi.org/10.1038/s41562-021-01056-1>
- Lou, M., & Ahmed, S. (2019, March 14). The American Medical Association is asking tech companies to stop the spread of vaccine misinformation. *CNN Wire Service*. <http://search.proquest.com/docview/2191086633/citation/71D17C6ABEE4A9CPQ/1>
- McCrosky, J., & Geurkink, B. (2021). *YouTube regrets report*. Mozilla Foundation.
- MacDonald, N. E., & SAGE Working Group on Vaccine Hesitancy. (2015). Vaccine hesitancy: Definition, scope and determinants. *Vaccine*, *33*(34), 4161–4164. <https://doi.org/10.1016/j.vaccine.2015.04.036>
- Meta. (2022). *COVID-19 and vaccine policy updates & protections*. Facebook Help Center. <https://www.facebook.com/help/230764881494641/>
- Moon, H., & Lee, G. H. (2020). Evaluation of Korean-language COVID-19-related medical information on YouTube: Cross-sectional infodemiology study. *Journal of Medical Internet Research*, *22*(8), Article e20775. <https://doi.org/10.2196/20775>
- Morris, M., Handcock, M. S., & Hunter, D. R. (2008). Specification of exponential-family random graph models: Terms and computational aspects. *Journal of Statistical Software*, *24*(4), 1548–7660.
- Neff, T., Kaiser, J., Pasquetto, I., Jemielniak, D., Dimitrakopoulou, D., Grayson, S., Gyenes, N., Ricaurte, P., Ruiz-Soler, J., & Zhang, A. (2021). *Vaccine hesitancy in online spaces: A scoping review of the research literature, 2000-2020*. Harvard Kennedy School Misinformation Review. <https://doi.org/10.37016/mr-2020-82>
- Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(23), 8577–8582. <https://doi.org/10.1073/pnas.0601602103>
- Omer, S. B. (2020). The discredited doctor hailed by the anti-vaccine movement. *Nature*, *586*(7831), 668–669. <https://doi.org/10.1038/d41586-020-02989-9>
- Omer, S. B., Richards, J. L., Ward, M., & Bednarczyk, R. A. (2012). Vaccination policies and rates of exemption from immunization, 2005-2011. *New England Journal of Medicine*, *367*(12), 1170–1171.
- Pan American Health Organization. (2021). *Misinformation fueling vaccine hesitancy, PAHO Director says*. <https://www.paho.org/en/news/21-4-2021-misinformation-fueling-vaccine-hesitancy-paho-director-says>
- Plandemic Series. (2021). <https://plandemicseries.com/>
- Reuters. (2021). *YouTube blocks all anti-vaccine content*. <https://www.reuters.com/technology/youtube-blocks-all-anti-vaccine-content-washington-post-2021-09-29/>
- Robertson, A. (2021, April 26). YouTube launches PSAs encouraging Americans to get vaccinated. *The Verge*. <https://www.theverge.com/2021/4/26/22403987/youtube-covid-19-vaccine-psa-because-everything>
- Röchert, D., Weitzel, M., & Björn Ross, B. (2020). The homogeneity of right-wing populist and radical content in YouTube recommendations. In *International conference on social media and society (SMSociety'20)* (pp. 245–254). Association for Computing Machinery. <https://doi.org/10.1145/3400806.3400835>
- Sallam, M. (2021). COVID-19 vaccine hesitancy worldwide: A concise systematic review of vaccine acceptance rates. *Vaccines*, *9*(2), Article 160. <https://doi.org/10.3390/vaccines9020160>
- Silverman, C. (2016, November 16). This analysis shows how viral fake election news stories outperformed real news on Facebook. *BuzzFeed News*. <https://www.buzzfeednews.com/article/craig-silverman/viral-fake-election-news-outperformed-real-news-on-facebook>
- Solsman, J. E. (2018). Ever get caught in an unexpected hourlong YouTube binge? Thank YouTube AI for that. *CNET*. <https://www.cnet.com/news/youtube-ces-2018-neal-mohan/>
- Song, M. Y.-J., & Gruzd, A. (2017). Examining sentiments and popularity of pro- and anti-vaccination videos on YouTube. In *Proceedings of the 8th international conference on social media & society* (pp. 1–8). <https://doi.org/10.1145/3097286.3097303>
- Sosa, M. E., Gargiulo, M., & Rowles, C. (2015). Can informal communication networks disrupt coordination in new product development projects? *Organization Science*, *26*(4), 1059–1078.
- Statista. (2021). *Most used social media 2021*. <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>
- Stokel-Walker, C. (2019, June 14). Opinion: Algorithms won't fix what's wrong with YouTube. *The New York Times*. <https://www.nytimes.com/2019/06/14/opinion/youtube-algorithm.html>
- Szeto, E., Pedersen, K., & Tomlinson, A. (2021, March 30). Marketplace flagged over 800 social media posts with COVID-19 misinformation. Only a fraction were removed. *CBC News*. <https://www.cbc.ca/news/marketplace/marketplace-social-media-posts-1.5968539>
- Tang, L., Fujimoto, K., Amith, M. (Tuan), Cunningham, R., Costantini, R. A., York, F., Xiong, G., Boom, J. A., & Tao, C. (2021). “Down the rabbit hole” of vaccine misinformation on YouTube: Network exposure study. *Journal of Medical Internet Research*, *23*(1), Article e23262. <https://doi.org/10.2196/23262>
- Tangcharoensathien, V., Calleja, N., Nguyen, T., Purnat, T., D’Agostino, M., Garcia-Saiso, S., Landry, M., Rashidian, A., Hamilton, C., AbdAllah, A., Ghiga, I., Hill, A., Hougendobler, D., Andel, J., van Nunn, M., Brooks, I., Sacco, P. L., Domenico, M. D., Mai, P., . . . Briand, S. (2020). Framework for managing the COVID-19 infodemic: Methods and results of an online, crowdsourced WHO technical consultation. *Journal of Medical Internet Research*, *22*(6), Article e19659. <https://doi.org/10.2196/19659>

- Thaker, J., & Subramanian, A. (2021). Exposure to COVID-19 vaccine hesitancy is as impactful as vaccine misinformation in inducing a decline in vaccination intentions in New Zealand: Results from pre-post between-groups randomized block experiment. *Frontiers in Communication*, 6, Article 721982. <https://doi.org/10.3389/fcomm.2021.721982>
- Tokojima Machado, D. F., de Siqueira, A. F., & Gitahy, L. (2020). Natural stings: Selling distrust about vaccines on Brazilian YouTube. *Frontiers in Communication*, 5, Article 577941. <https://www.frontiersin.org/article/10.3389/fcomm.2020.577941>
- Travers, M. (2020, March 21). Facebook spreads fake news faster than any other social website, according to new research. *Forbes*. <https://www.forbes.com/sites/traversmark/2020/03/21/facebook-spreads-fake-news-faster-than-any-other-social-website-according-to-new-research/>
- Tustin, J. L., Crowcroft, N. S., Gesink, D., Johnson, I., & Keelan, J. (2018). Internet exposure associated with Canadian parents' perception of risk on childhood immunization: Cross-sectional study. *JMIR Public Health and Surveillance*, 4(1), e7. <https://doi.org/10.2196/publichealth.8921>
- Wetsman, N. (2020, October 14). YouTube will remove videos with COVID-19 vaccine misinformation. *The Verge*. <https://www.theverge.com/2020/10/14/21515796/youtube-covid-vaccine-misinformation-policy>
- World Health Organization. (2019). Ten threats to global health in 2019. <https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019>
- World Health Organization. (2022). *WHO coronavirus (COVID-19) dashboard*. <https://covid19.who.int>
- Xu, C., Hui, P. M., Jha, O. K., Xia, C., & Johnson, N. F. (2022). Preventing the spread of online harms: Physics of contagion across multi-platform social media and metaverses. *arXiv*. <http://arxiv.org/abs/2201.04249>
- YouTube. (2021). *COVID-19 medical misinformation policy—YouTube help*. <https://support.google.com/youtube/answer/9891785?hl=en>
- Yu, L., Huang, J., Zhou, G., Liu, C., & Zhang, Z.-K. (2017). TIIREC: A tensor approach for tag-driven item recommendation with sparse user generated content. *Information Sciences*, 411, 122–135. <https://doi.org/10.1016/j.ins.2017.05.025>
- Zhang, J., & Centola, D. (2019). Social networks and health: New developments in diffusion, online and offline. *Annual Review of Sociology*, 45(1), 91–109. <https://doi.org/10.1146/annurev-soc-073117-041421>
- Zhou, R., Khemmarat, S., & Gao, L. (2010). The impact of YouTube recommendation system on video views. In *Proceedings of the 10th annual conference on internet measurement—IMC '10* (pp. 404–410). <https://doi.org/10.1145/1879141.1879193>
- Zhou, R., Khemmarat, S., Gao, L., Wan, J., & Zhang, J. (2016). How YouTube videos are discovered and its impact on video views. *Multimedia Tools and Applications*, 75(10), 6035–6058. <https://doi.org/10.1007/s11042-015-3206-0>

Author Biographies

Anatoliy Gruzd (PhD, University of Illinois at Urbana-Champaign) is a Canada Research Chair, Professor and Director of Research at the Social Media Lab at Toronto Metropolitan University. Situated at the intersection of social media research, information management, and communication, Dr Gruzd's multi-disciplinary program explores how social media and the growing availability of user data are changing the ways in which people and organizations communicate, collaborate, and disseminate information and how these changes impact the social, economic, and political norms and structures of modern society (email: gruzd@torontomu.ca).




Deena Abul-Fottouh (PhD, McMaster University) is a computational social scientist and a postdoctoral research fellow at the Digital Society Lab at McMaster University. Her research investigates how social media data can guide our understanding of various societal problems including online misinformation, online extremism, anti-social behavior such as hate speech, and digital activism. She is also interested in the use of social media in governance. Dr Abul-Fottouh uses methods of social network analysis, data science, and natural language processing to study the structure and content of social media platforms (email: abulfodm@mcmaster.ca).

Melodie YunJu Song (BSN, MSc, PhD) is a postdoctoral research fellow at the Centre for Vaccine Preventable Diseases, Dalla Lana School of Public Health, University of Toronto. Interested in information as a social determinant of health, Dr Song conducts mixed-methods research to understand public health authorities' perception and use of health communication strategies on social media to increase vaccine confidence. Previously, she was a recipient of the Canadian Health Systems Impact Fellowship and served as a member of the Board of Directors of the Canadian Association for Health Services and Policy Researchers (email: yunju.song@utoronto.ca).

Alyssa Saiphoo (PhD, Toronto Metropolitan University) is a post-doctoral research fellow at the Social Media Lab at Toronto Metropolitan University. Her research interests include social psychology, social comparison theory, and the effects of social media on well-being. Dr Saiphoo uses experiments, qualitative methods, and social network analysis to understand how people use and incorporate social media platforms into their daily lives (email: alyssa.saiphoo@torontomu.ca).

Appendix A

Most viral public Facebook entities (pages and groups) with vaccine-related posts in June 2020.

Rank	Page/Group name	Still available? (as of December 2021)	Vaccine stance	In English?	#Posts	Total interactions ^a (across all posts by each page)	Cumulative count of interactions ^b (running total)	Cumulative % of interactions ^c (running total relative to all interactions)
1	Africa Centers for Disease Control and Prevention	Yes	Pro	Yes	3	141,316	141,316	35
2	Collective Action Against Bill Gates. We Wont Be Vaccinated!!	No	Anti	Yes	567	38,935	180,251	44
3	somoynews.tv	Yes	Neutral	No	1	16,018	196,269	48
4	Michelle Malkin	Yes	Anti	Yes	1	11,712	207,981	51
5	We Are Vaxxed	No	Anti	Yes	22	10,882	218,863	54
6	Lone Star Dog Ranch & Dog Ranch Rescue	Yes	Not relevant	Yes	2	7,468	226,331	56
7	Quỳnh Trần JP	Yes	Not relevant	No	1	6,635	232,966	57
8	The Truth About Cancer	Yes	Anti	Yes	5	4,946	237,912	59
9	99% unite Main Group "it's us or them"	No	Anti	Yes	114	4,430	242,342	60
10	UNTV News and Rescue	Yes	Neutral	Yes	2	4,277	246,619	61
11	John Pavlovitz	Yes	Pro	Yes	11	3,386	250,005	61
12	FABUNAN ANTIVIRAL INJECTION SUPPORTERS GROUP	No	Anti	No	13	3,231	253,236	62
13	ARY News	Yes	Neutral	No	3	2,731	255,967	63
14	SABC News	Yes	Neutral	Yes	6	2,659	258,626	64
15	Fauci, Gates, & Soros to prison worldwide Resistance	Yes	Anti	Yes	44	2,594	261,220	64
16	Aster Bedane	Yes	Anti	No	1	2,248	263,468	65
17	OFFICIAL Q / QANON / r / ra / rawgr / Q+ / Q++++	No	Anti	Yes	103	1,993	265,461	65
18	Rehana Fathima Pyarijaan Sulaiman	Yes	Not relevant	No	1	1,850	267,311	66
19	 BBC News  नेपाली 	No	Pro	No	4	1,771	269,082	66
20	Didier Raoult professeur Marseille	Yes	Anti	No	6	1,654	270,736	67
21	Wits—University of the Witwatersrand	Yes	Pro	Yes	4	1,622	272,358	67
22	The Wild Doc	Yes	Anti	Yes	4	1,617	273,975	67
23	Vaxxed Global Movement	No	Anti	Yes	30	1,597	275,572	68
24	Glorious And Free	Yes	Anti	Yes	12	1,386	276,958	68
25	Larry Elder	Yes	Not relevant	Yes	1	1,311	278,269	68
26	Energy Therapy	No	Anti	Yes	9	1,304	279,573	69
27	Stop Mandatory Vaccination	No	Anti	Yes	31	1,278	280,851	69
28	Charlie Wards Group	No	Anti	Yes	8	1,269	282,120	69
29	Thibaan Channel	Yes	Not relevant	No	1	1,238	283,358	70
30	ScotNepal.Com	Yes	Pro	No	1	1,211	284,569	70
31	Chemtrails Global Skywatch	No	Anti	Yes	20	1,206	285,775	70
32	Yellow Vests Canada	No	Anti	Yes	4	1,130	286,905	71
33	La Pèlerine des Étoiles	Yes	Anti	No	1	1,084	287,989	71
34	Jairam Sarkar Report Card	Yes	Pro	No	1	1,057	289,046	71
35	United States for Medical Freedom	No	Anti	Yes	21	1,029	290,075	71

(Continued)

Appendix A. (Continued)

Rank	Page/Group name	Still available? (as of December 2021)	Vaccine stance	In English?	#Posts	Total interactions ^a (across all posts by each page)	Cumulative count of interactions ^b (running total)	Cumulative % of interactions ^c (running total relative to all interactions)
36	Henri Joyeux	Yes	Anti	No	1	1,021	291,096	72
37	COALITION MONDIALE EN SOUTIEN AU DOCTEUR DIDIER RAOULT	No	Anti	No	16	955	292,051	72
38	Paul Thomas, M.D.	Yes	Anti	Yes	1	929	292,980	72
39	Support Glenn Chong	Yes	Anti	Yes	2	923	293,903	72
40	Afrikaners	Yes	Anti	No	1	911	294,814	73
41	Dr. John Bergman	Yes	Anti	Yes	4	899	295,713	73
42	Down the Rabbit Hole	No	Anti	Yes	9	847	296,560	73
43	The Trump Republicans	Yes	Anti	Yes	1	845	297,405	73
44	Gyanendra Shahi—हाम्रो नेपाल, हामी नेपाली समर्थक	Yes	Pro	No	3	833	298,238	73
45	Citizens Unite UK #wakeup	No	Anti	Yes	43	829	299,067	74
46	AltHealthWORKS	Yes	Anti	Yes	1	781	299,848	74
47	Maasim Ta Vines	No	Not relevant	No	1	765	300,613	74
48	Weston A. Price Foundation	No	Anti	Yes	1	758	301,371	74
49	Rabi Lamichhane & Apil Tripathi Fans Club Nepal	Yes	Pro	No	1	758	302,129	74
50	JABS: Justice, Awareness & Basic Support	Yes	Anti	Yes	6	712	302,841	74
51	New York Alliance for Vaccine Rights	Yes	Anti	Yes	3	684	303,525	75
52	The Canadian Revolution	No	Anti	Yes	4	672	304,197	75
53	Oregon Republican League	Yes	Anti	Yes	1	666	304,863	75
54	Bayan Ko Ph	Yes	Anti	Yes	1	636	305,499	75
55	मुसिक MUSiC	Yes	Not relevant	No	1	627	306,126	75
56	CULT 45 DEPLORABLE AMERICANS FOR TRUMP	Yes	Anti	Yes	2	616	306,742	75

^aThe “Total Interactions” column adds up the number of interactions (including reactions, shares, and comments) across all posts shared by each entity in the dataset. The number of posts is indicated in the “#Posts” column.

^bThe “Cumulative Count of Interactions” column is the running total of interactions across all 56 Facebook entities in this analysis. This column sums the total number of interactions from all entities from the beginning of the ranked list to the current row.

^cThe “Cumulative % of Interactions” column calculates the percentage of the “Cumulative Count of Interactions” in the current row (i.e., for entities from the beginning of the ranked list to the entity in the current row) relative to the total number of interactions (406,530) recorded in the full dataset of Facebook posts ($N=8,549$) across 4,453 Facebook entities.

Appendix B**Vaccine Stance Coding Guide for YouTube Videos****Neutral video:**

1. The video is about memorizing vaccine schedules for a test, provided by a cram school or high school in India, or a MCAT test prep center, for instance.
2. A news clip or video that presents BOTH SIDES of the vaccine “debate,” OR a video that presents a fact related to vaccines but DOES NOT TAKE SIDES. For example, [https:// link-to-a-sample-video]—this is a news clip of “Amazon removes anti-vaccine films” without a stance or opinion on whether it is a good move or bad move.

Pro-vaccine video:

1. The video is clearly supportive of vaccines.
2. The video explains how vaccines work and may say the keywords “vaccines are effective” or “vaccines protect us from diseases.”
3. Not all news channels/videos are neutral, in fact most are pro-vaccine: if the news agency *ends* the interview with “health officials say that vaccinations saves lives” and closes the segment with a doctor criticizing anti-vaccine or advocating for vaccines, it is a pro-vaccine video.
4. ALL videos from WHO, UNICEF, CDC, GAVI, Bill and Melinda Gates Foundation, Mayo Clinic, are clearly pro-vaccine.

- Examples: [https:// link-to-a-sample-video] (“How to pack and use a vaccine carrier”) is uploaded by “Immunization Academy” which is a source for enhancing vaccine training and delivery in the field, funded by the Bill and Melinda Gates Foundation.
 - [https:// link-to-a-sample-video]—The title of the video is “are fetal cells used to make vaccines?” This is a pro-vaccine video made by Dr Paul Offit, published by the Children’s Hospital of Philadelphia in the United States. The title speaks directly to people who are anti-vaccine and who share false information that fetal tissues are used to cultivate vaccines to protest vaccines based on moral righteousness. While given the lack of information on the background context of the “fetal tissue” debate, this title may seem “neutral,” but this is a pro-vaccine video.
 - [https:// link-to-a-sample-video]—this is a pro-vaccine video (not neutral); in this video, a reporter discussing the passing of California’s mandatory vaccine law and asks the question “should more States follow California.” The title “Science vs. The Antivaxxers” already gives away that this video is pro-vaccine because the condescending tone of the video.
5. Some videos have a very clear agenda even without watching the video, for instance, images with health-care professionals going to rural areas to inject vaccination (even without voice that describes the action) is a likely indicator of pro-vaccine stance.
- [https:// link-to-a-sample-video]—“A mother’s quest by canoe to vaccines in Sierra Leone—UNICEF”—this is a pro-vaccine video.
6. Some videos start out as educational or informational, but fast forwarding to the end, there is often a strong pro-vaccine statement. If the video is a lesson on “*how vaccines work*”—they are usually pro-vaccine (you’ll find that at the end of the video with a strong statement critiquing people who are anti-vaccine or endorsing vaccination). Note: most of these videos are produced to explain that vaccines are safe. If that is the case, these videos ARE NOT NEUTRAL in stance toward vaccination.

Please pay special attention to concepts like “*herd immunity*,” or anything related to “*vaccination is good*,” “*unvaccinated is bad*”. For example, below are two clipped images from the videos showing how these terms may be used and their pro-vaccine orientation:

[https:// link-to-a-sample-video]

Other examples of pro-vaccine videos include:

- [https:// link-to-a-sample-video]—this video is clearly pro-vaccine (not neutral). The video called “*demystifying medicine*” from McMaster University is a

seminar on addressing vaccine hesitancy from a historical point of view as indicated in the description.

- [https:// link-to-a-sample-video]—this is a pro-vaccine video (not neutral) because it is an introduction of a tool that educates children on the benefits of vaccines.
- [https:// link-to-a-sample-video]—another example of a pro-vaccine video, not a neutral video. The presenter is a doctor who says “*this is an awesome vaccine*” at 0:37.
- [https:// link-to-a-sample-video]—this is a pro-vaccine educational video where the video description outlines “*we also discuss how anti-vaccine stance was rooted in the . . .*”
- [https:// link-to-a-sample-video]—this video is a pro-vaccine video; it encourages taking vaccines before traveling.
- [https:// link-to-a-sample-video]—this channel is “*India for vaccines*,” the video is pro-vaccine. You can see that the doctors will discuss how scary the diseases are, and why you should take vaccines.

Anti-vaccine video:

1. The video contains language that disapproves of vaccines.
 2. The video may feature clips from documentaries (e.g., truth about vaccines, vaccine nation), talk shows, public protests, presentations from Del Big Tree, chiropractors, naturopaths, homeopaths, osteopaths, dietitians. Note: Beware that these videos are often disguised as pro-vaccine or neutral videos.
- [https:// link-to-a-sample-video]—this anti-vaccine video is by a popular anti-vaccine channel “The truth about vaccines.”
 - [https:// link-to-a-sample-video]—this video is anti-vaccine. It is produced by one of the biggest anti-vaccine YouTube channels in North America “iHealthTube.com.”
 - [https:// link-to-a-sample-video]—this is an anti-vaccine video because the description was a badly edited video of an ACIP hearing and how vaccination recommendations are bad (or so the person who posted the video thought). The narrative is anti-vaccine even though the hearing is a vaccine recommendation hearing. The video description may help with coding in such cases.
 - [https:// link-to-a-sample-video]—this is an anti-vaccine video.
 - [https:// link-to-a-sample-video]—this is an anti-vaccine video. Note: Del Big Tree is one of the biggest anti-vaccine proponents who says, “*he’s not against vaccines but.*” Often anti-vaccine videos are tricky, you must watch at least one anti-vaccine video to

learn how people who are anti-vaccine draw parents into false narratives.

- [https:// link-to-a-sample-video]—this is an anti-vaccine video. This news piece is by Next New network interviewed an individual who is anti-vaccine (Dr Mayer Eisenstein) and who tells audiences how to acquire vaccine exemption; in the video he uses phrases such as pharma is dangerous/corrupt; first amendment rights, politicians are evil or misinformed.
 - [https:// link-to-a-sample-video]—Vaccine Nation is an anti-vaccine documentary.
 - [https:// link-to-a-sample-video] “The Real REason Aluminum is in Vaccines!” This is by iHealthTube.com, which was demonetized by YouTube for being an anti-vaccine channel. Most commonly they would hire doctors who would use pseudoscience keywords like “supercharged” or “live vaccines/killed vaccines,” “immunization stimulation.”
3. Some videos are posted by individuals who try to link vaccines to conspiracy theories. Read the description and then quickly watch videos listed below for practice.
- [https:// link-to-a-sample-video]—this is an anti-vaccine video by a conspiracy theorist (also see the video description).
 - [https:// link-to-a-sample-video]—similarly, this is an anti-vaccine video. The title says “Vaccines destroying children’s health,” it shows a little girl’s optic neuritis (a type of adverse event following immunization that can lead to vision loss but mostly temporary).
 - [https:// link-to-a-sample-video]—this is an anti-vaccine video. It is posted by a conspiracy theorist who disagrees with getting vaccinated.
 - [https:// link-to-a-sample-video]. Note: If a video has “create your own vaccine schedule,” this is likely shared by a vaccine-hesitant group in an area with mandatory vaccine regulations for school-aged children.
 - [https:// link-to-a-sample-video]—this is a person talking about how immunization is bad in this long video at 33:39.

Step-by-step identification of hard-to-determine anti-vaccine videos:

1. Look at the channel of the video: If the user has a channel that has first+last name, or an ambiguous

name that implies this is not an official channel, look at the description carefully.

2. Look at the channel and the “dislike/like ratio”: if a video on a reputable channel (e.g., CBC, BBC, CNN) has more dislikes / than likes, it is likely a pro-vaccine video.
 3. Look at the video description: If the description section says anything about conspiracy theories, such as planting chips in humans using vaccines, it is an anti-vaccine video.
 4. Look at the content of the video: If the video contains an interview of Bill Gates talking about vaccines, look carefully at the content and see if there’s anything on “population control,” or “why is Bill Gates choosing Africa?” or “genocide,” or “mind control,” “choice not mandates.” Any keywords as such in the description should be treated with care. Some traits of an anti-vaccine video—of low resolution, cropped, or reverse-image, with commentaries involving “secret,” “unreleased,” “behind doors,” and so on—are strongly suggestive of a conspiracy theory video.
- Note that content with doctors is not necessarily pro-vaccine—they may be osteopaths and naturopaths. How do we know? Look up the doctor’s name on Google and determine whether they have a track record of anti-vaccine statements.
 - Note that videos that seem like they are cropped from a reputable source, such as Bill Gates being interviewed by a reporter, are not always pro-vaccine videos (and most likely are not). Refer to Steps 1 to 3.

In addition to the stance coding, please indicate one of the following:

“**Not Available**”: Is the video still available?

Select this option if the video has been removed by the user or blocked by the platform.

“**Not English**”: Is the video content not in English?

Select this option if the video content is in a foreign language (even if the title is in English).

BUT: Videos with description and subtitles in English, and/or with partial interviews in English should be included.

“**Not Relevant**” Should we exclude this video because it is not about vaccination? Example: a video is not relevant, if it’s a finance-related course that mentions “immunization.”

Appendix C

Top 10 YouTube channels based on the most linked to videos (in-degree) by Stance by Cluster.

Cluster A (27% of all nodes in the network)

Pro-vaccine (44% of nodes in the cluster ^a)	Anti-vaccine (45% of nodes in the cluster)	Neutral (12% of nodes in the cluster)
Jimmy Kimmel Live	Vaccine Risks	TODAY
SciShow	Rolling Stone	Jubilee
The Doctors	Shameless Maya	OSSA
TEDx Talks	Thom Hartmann Program	The Doctors
AsapSCIENCE	Vaccine Risks	Yahoo Finance
JaeVR	CinemaLibre	60 Minutes Australia
TEDx Talks	The Charlotte Observer	This Morning
Kurzgesagt—In a Nutshell	CinemaLibre	Bloomberg QuickTake Originals
CBC News: The National	paulthomasmd	Larry King
TED-Ed	The Real Truth About Health	PBS NewsHour

Cluster B (19% of all nodes in the network)

Pro-vaccine (82%)	Anti-vaccine (6%)	Neutral (12%)
60 Minutes Australia	700 Club Interactive	This Morning
DW News	LifeSiteNews	South China Morning Post
CNA Insider	LifeSiteNews	MedCram—Medical Lectures
		Explained CLEARLY
BBC News	ljahstars THE MINDSET	MedCram—Medical Lectures
		Explained CLEARLY
ABC News In-depth	ROME REPORTS in English	Sky News
Real Science	Louis B	CNBC International TV
Dr. John Campbell	Sabins Studio	BBC
CBN News	paulthomasmd	Sky News Australia
Sky News	LifeSiteNews	MSNBC
Yahoo Finance	Dota 2 Gaming Highlights	Bloomberg Markets and Finance

Cluster C (12% of all nodes in the network)

Pro-vaccine (77%)	Anti-vaccine (7%) (Only 7 anti-vaccine videos had in-degree > 0)	Neutral (16%)
BBC News	China in Focus—NTD	Christy Risinger MD
DW News	NTD	CNBC Television
Wits University OFFICIAL	Ben Swann	7NEWS Australia
Dr. John Campbell	The Last American Vagabond	ABC 7 Chicago
Bloomberg Markets and Finance	thx1138mindlock	Yahoo Finance
This Morning	Dr Mumbi LIVE	The Infographics Show
United Nations	ODANA NETWORK	Al Jazeera English
CNBC Television		MedCram—Medical Lectures
		Explained CLEARLY
CNN		The Telegraph
The Economist		Dr Khan Show

Cluster D (11% of all nodes in the network)

Pro-vaccine (29%)	Anti-vaccine (60%)	Neutral (11%)
Bill Gates	Viable Tv	KTN News Kenya
CNN	Kenya CitizenTV	NDTV
The Late Show with Stephen Colbert	Kenya CitizenTV	Al Jazeera English
The Late Show with Stephen Colbert	Bible Baptist Potch	Channels Television

(Continued)

Appendix C. (Continued)

Cluster D (11% of all nodes in the network)

Pro-vaccine (29%)	Anti-vaccine (60%)	Neutral (11%)
CNBC Television	Kristofer Aspen	Newsy
CNBC Television	NewsClickin	Hindustan Times
CNN	Dal Khalsa UK	Footnotes
CNBC Television	Nathan Riddett	DailyNation
Bloomberg QuickTake Originals	SupergirlFan	NDTV
FRANCE 24 English	Adventist World Radio	CBS 17

Cluster E (11% of all nodes in the network)

Pro-vaccine (40%)	Anti-vaccine (51%)	Neutral (9%)
Doctor Mike	Viable Tv	Reuters
University of California Television (UCTV)	Dr Rashid A Buttar	MSNBC
The University of Arizona	Dr Rashid A Buttar	Vincent Racaniello
Vincent Racaniello	Dr Rashid A Buttar	Vincent Racaniello
Los Angeles Times	Dr Rashid A Buttar	The Social CTV
Stanford	Dr Rashid A Buttar	University of California Television (UCTV)
University of California Television (UCTV)	Viable Tv	Steve Judd Astrology
Vincent Racaniello	Dr Rashid A Buttar	Vincent Racaniello
Vincent Racaniello	Dr Rashid A Buttar	VaccineSafetyConf
Vincent Racaniello	Dr Rashid A Buttar	Sunny

^aThe total is over 100% due to the rounding.

Appendix D

Distribution of video categories in the YouTube's Related Videos network ($N=2,260$).

Video category	Number of videos
News & Politics	970
Education	367
People & Blogs	278
Nonprofits & Activism	238
Science & Technology	209
Entertainment	102
Film & Animation	27
Comedy	25
Howto & Style	17
Music	11
Travel & Events	6
Sports	4
Pets & Animals	3
Gaming	1
Autos & Vehicles	1
(not provided)	1
Grand Total	2,260