Examining Sentiments and Popularity of Pro- and Anti-Vaccination Videos on YouTube

Melodie Yun-Ju Song  
McMaster University  
1280 Main Street West  
Hamilton, ON, Canada L8S4K1  
+1-905-525-9140 ext.22879  
songy42@mcmaster.ca

Anatoliy Gruzd  
Ryerson University  
350 Victoria Street  
Toronto, ON, Canada, M5B 2K3  
+1-416-979-5000 ext.3509  
gruzd@ryerson.ca

ABSTRACT

Vaccine misinformation on social media poses significant drawbacks to the efforts of vaccine coverage rates. This research studies the interlinkages between pro- and anti-vaccine YouTube videos to help public health professionals explore new ways to reach anti-vaccine and vaccine-hesitant audiences. Using YouTube’s API, we retrieved 9,489 recommended videos from 250 seeds using keywords such as “vaccines” and its derivatives. We then manually identified 1,984 videos directly related to vaccination and then categorized their vaccine sentiment into pro-, anti-, and neutral. Results show that 65.02% of the videos were anti-vaccine, and only 20.87% were pro-vaccine, 14.11% were neutral. Anti-vaccine videos were significantly more prevalent in the “News & Politics” and “People & Blogs” video categories; while pro-vaccine videos were more prevalent in the “Education” and “Science & Technology” categories. Results also showed that anti-vaccine sentiment videos have higher values of closeness centrality (p=0.05), suggesting that watching an anti-vaccine video will likely lead to more anti-vaccine video recommendations. Moreover, videos that had more dislikes than likes (dislike/like ratio) are positively related to pro-vaccine videos (OR=3.912), suggesting that pro-vaccine videos are more ill-received on YouTube than anti-vaccine videos. This study is the first to examine the network of vaccine-related videos on YouTube and their centralities. The results highlight some possible limitations of YouTube-based vaccination awareness campaigns and also emphasize the need to diversify how YouTube makes its recommendations to help viewers break out of the anti-vaccine “bubble.”

CCS Concepts

- Networks → Network types → Overlay and other logical network structures → Social media networks Information systems → Information retrieval → Retrieval tasks and goals → Recommender systems.

Keywords

YouTube; Anti-vaccine; Vaccine; Vaccine hesitancy; Videos; Sentiment analysis; Network analysis; Social Media.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

#SMSociety'17, July 28-30, 2017, Toronto, ON, Canada © 2017 Copyright is held by the owner/author(s). Publication rights licensed to ACM. ACM 978-1-4503-4847-8/17/07... $15.00  
http://dx.doi.org/10.1145/3097286.3097303

1. INTRODUCTION

Vaccination is considered the greatest public health achievement of the 20th century by the Centers for Disease Control [1], it is one of the safest public health interventions that effectively reduces mortality, disability, and poverty [2]. However, negative sentiment towards vaccines is becoming an increasing concern over the past 20 years. A claim widely considered to originate from a now-retracted paper published in an influential medical journal The Lancet, the authors falsely claimed a positive correlation of Mumps-Measles-Rubella (MMR) vaccine and autism, the results of the paper was widely shared within an influential Autism community online that sponsored American talk-show host Jenny McCarthy in the early 2000’s as their spokesperson [3,4]. This led to a growing interest among researchers to start examining media, celebrity culture, and the Internet to understand how and why misinformation spreads [5-7]. The following is a literature review on the socioeconomic and attitudinal variables that helps us understand vaccine-hesitancy from different paradigms.

1.1 Three Paradigms to Interpret Anti-Vaccination Movement

Anti-vaccination movement is first and foremost a health systems delivery issue. Under the evidence-based paradigm which promotes that evidence provides the basis for effective treatment, healthcare providers tend to see anti-vaccine movement as a population-based “knowledge deficit” exaggerated by health information-seeking on the Internet. Literature produced under the evidence-based paradigm emphasizes the responsibility of healthcare providers to present facts to debunk myths, to take time to listen and address concerned parents’ doubts, and ultimately, to earn the trust of patients [8-12]. Systematic reviews of 17 studies suggested limited effectiveness of educational interventions on vaccine coverage [13]. More recently, a randomized clinical trial indicated that pro-vaccine messages might even worsen misconceptions for people who are already anti-vaccine [14].

A second popular interpretation of anti-vaccination movement comes from the social sciences and qualitative researchers in public health. Some propose to incorporate communications theory and marketing strategies to inject trust and convenience into the vaccine-hesitant [15], while some use social construction theories and postmodernism to explain that anti-vaccination movement is contextual [6]. Under this paradigm, researchers observe that knowledge is perceived as situational and subjective, therefore an individual’s belief about vaccines is actively being generated and interpreted to best suit their personal belief system; often, values are constructed through emotional appeal over evidence [16,17].

Moving along the same line of inquiry, public health professionals recognize that risk-negation increases vaccine-hesitancy [18].
Sociologists and public health scholars employ the concept of “risk society” [19] to explain that as civil society progresses, getting vaccine is beyond the control of technocracy or meritocracy [20-22]. Instead, curbing vaccine hesitancy is not a simple matter of “educate and vaccinate” health service announcements, but rather it requires a targeted and dialogical communication, tailored for the concerned individuals [23].

Finally, many researchers have begun looking at the medium in which misinformation resides, namely, social media. Social media has challenged the models, theories, and processes underpinning decision-making theories [24]. Social media fosters an atmosphere of reciprocal support and mutuality [25], creates virtual networks that allow like-minded people to form clusters/cliques [26], and influence others through information diffusion in a scale-free environment [27]. A large-scale study of 253 million Facebook users has demonstrated that the ecology of social media allows information to propagate through weak ties more than strong ties [28]. Furthermore, comparing the cascade effect between Digg and Twitter, researchers have observed that Digg’s dense network leads to faster but limited spread of information, while Twitter’s lose network created a slower but a wider spread of information. Thus, confirming that network structures can profoundly influence reach [29,30].

In the past two decades, properties of the network theory has been used extensively by social epidemiologists; for example, using the term “social contagion” to explain the spread of non-infectious diseases such as obesity [31], depression, and smoking cessation [32]. Others have looked at the causal pathways of network consolidation on the diffusion of opinions, norms, and practices using web-based stimulator games [33]. Even though social network analysis is at its infancy in public health, researchers are increasingly recognizing that a network’s structural components is an important determinant of health disparities and health beliefs in communities [34–36]. Social media distinguishes itself from traditional health interventions as having scale-free properties that enables us to look at the centrality of the actors (nodes) that can influence opinions and norms [37].

1.2 Social Media and YouTube as an Untapped Opportunity for Public Health Campaigns

It is estimated that 71% of Internet users have searched for information about their health online [38]. For many, social media has also become one of the first places we look for health information, following going to a clinic, or asking advice from family and friends [39,40]. As iterated, the unique function of social media enables instant, free, dialogical interaction, through which communities are created [41]. In the context of public health, there are many possible applications for social media. For example, it can provide users with anonymous information and social support for health conditions, such as sexual and mental health, that are often stigmatized [42]. Social media also creates an environment of “legitimate peripheral learning” for those who wish to partially participate in learning about a certain health information [41,43]. It also enables the exchange of health information without geographical constraints [44]; each of these examples serve as direct and indirect mechanisms to empower the patient and the public.

Despite the benefits of seeking health-information, health care professionals whose daily practice are rooted in the positivist paradigm follow evidenced-based practice. This ideology appears to be at odds with the fast-paced social media platforms, which have no submission threshold where anyone can post and share vaccine information and whose claims might not offer evidence, testing of the hypothesis, or peer-review process, but are still be widely distributed.

As one of the most popular social media platforms, YouTube plays a key role in disseminating public health information (and misinformation). YouTube’s popularity dominates the current video-based social media platform. As of 2016, there are 800 billion active monthly users and 4 billion videos being viewed on YouTube per day, making it one of the widest-reaching, demographically diverse user-generated content platform [45]. It was on YouTube that the North American anti-vaccine movement was able to upload and share conference recordings free of charge to a wider audience. Compared to the past when tapes of proceeding had to be ordered by mail, YouTube allowed non-profit Autism awareness groups to easily host instant discussions, comments, and information-sharing as early as 2005 [46].

In this research, we use a networked mindset to tackle vaccine hesitancy. We examined literature regarding public health promotion that specifically focuses on YouTube as a popular platform for information sharing, and from this research raise the question: What is sentiment towards vaccine on YouTube and are anti-vaccine videos more central, and thus more likely to be found, than pro-vaccine videos? Our objectives are to quantify the current sentiment towards vaccines on YouTube; then, to explore differences and similarities in vaccine sentiment in relation to various video properties, such as video category, view count, and network properties.

2. RELATED WORK

In this section, we will introduce three areas of research relevant to this paper. First, we discuss the use of sentiment analysis of text, image, and videos on vaccination; second, we briefly go over vaccine awareness interventions and outreach in the community and on social media.

2.1 Vaccine Sentiment Analysis

In response to the rise of anti-vaccine sentiment, the World Health Organization (WHO) organized a Strategic Advisory Group of Experts (SAGE) on Immunization to monitor vaccine confidence, vaccine complacency, and to assess accessibility of vaccines worldwide [47]. While traditional surveys are useful to gauge sentiment in global health settings [48], our online lives are increasingly influencing offline behaviors and decision-making [49].

Researchers have looked at Facebook and Twitter where 30–35% of the content are anti-vaccine [50–52], similar results were found on MySpace and blogs [53–56]. Regarding image-based social media platforms, the only research known to us was by Guidry et al., who have observed that 75% of the vaccine sentiment on Pinterest are negative [57].

In video-based social media, some early research in this area showed that YouTube exhibited 32% negative vaccine sentiment in 2007 [46], the prevalence of negative sentiment continued to rise to 51.7% in 2013 [58], and 65.5% in 2017 [59]. Although the samples of the three studies cited above are somewhat different and they are not setup for direct comparison, the overall results suggest that negative sentiments regarding vaccination on YouTube are potentially on the rise.
2.2 Vaccine Promotion Strategies in the Community, on the Internet, and Social Media

A comprehensive review-of-reviews looked at 15 published literature reviews that analyzed hundreds of interventions addressing vaccine hesitancy. Results showed that it is inconclusive whether any interventions work to address vaccine refusal and vaccine hesitancy [60]. In the review, traditional interventions remain the most prominent: educational (in-person or in-patient), school-based programs, or distributing reminders and standing orders, which are predominantly patriarchal and grounded in the knowledge-deficit framework [4].

Due to the prominence of social media as a go-to source for health information and discussions, public health practitioners have developed and tried a wide array of interventions on social media to promote vaccine awareness. Some interventions include building informational websites, online portals, mobile applications, text-reminders, mailing lists, and advertisement [61,62]. Most interventions were able to monitor changes in attitudes towards vaccines but rarely the actual uptake of vaccines [61]. Furthermore, only one study has ever used YouTube as a source of intervention to observe whether watching anti-vaccine videos can change perceptions towards vaccines. However, the test subjects were first year medical students who were already predominantly pro-vaccine, thus after watching anti-vaccine videos the participants’ views on vaccines did not change [63].

3. METHODS

3.1 Data Collection

3.1.1 Retrieving Vaccine-related Videos

Our data retrieval was based on the “recommended videos” function on YouTube. YouTube relies on a number of parameters to suggest videos tailored to individuals, maximizing viewership, watch-time, and viewer retention. Their proprietary algorithm recommends videos based on related topics, co-viewing information, as well as information about users’ watch and search history, favored videos, likes, and subscriptions [64].

We used a program called Netvizz [65] to collect a YouTube video network. This network consists of nodes that are YouTube videos and connections that link “related” videos according to YouTube’s proprietary algorithm. To build this network, the program starts with the collection of seed videos that correspond to our search query. In our case, the search query was as follows: Immun* OR vaccin* OR vaxx*. In total, we retrieved 250 seed videos that corresponded to our search criteria. The number of seed videos was based on the five iterations of data retrieval using Netvizz. During each iteration, the program retrieved 50 items that corresponded to our search query. Once the seed videos were retrieved, Netvizz then queried each video to retrieve their corresponding lists of “related” videos with a crawl depth of one layer and five iterations. The resulting network consists of 9,489 nodes.

3.1.2 Variables

In order to examine a potential relationship between sentiment expressed in video towards vaccination and video attributes, we have identified seven independent variables, including six that were readily available for retrieval by Netvizz. These include comment count, view count, dislike count, like count, vote/dislike ratio, and video categories. Video categories are predefined by YouTube and allow users to select which category their videos will be filed under.

The following are 13 of those categories: auto vehicle, comedy, education, entertainment, film & animation, gaming, how-to & style, music, news & politics, non-profit & activism, people & blogs, pets & animals, science & technology, travel & events, shows, and movies. The seventh variable was the sentiment of the vaccine-related videos. Since the seed videos were retrieved using English keywords, any non-English titles that might have been recommended by YouTube’s “related” algorithm were ignored during this manual classification. To determine sentiments towards vaccination (either pro-vaccine, anti- or neutral) expressed in a video, one of the authors who specialized in public health watched all 1,984 vaccine-related videos containing in our results network of videos.

3.2 Data Analysis

3.2.1 Sentiment Analysis of Vaccine-related Videos

Given that the YouTube recommended videos are suggested based on an algorithm designed to optimize viewer retention and viewing time [65], prominent vaccine-related videos and videos related to them, may have distinguishing features in terms of centrality measures. Thus, we explore whether network centrality measures of vaccine-related videos will be different and if it interacts with the independent variables.

Among 9,489 videos collected for this study, only 1,984 videos had titles related to vaccine, immunization, and derivatives of these terms. Using this more targeted sample of 1,984 videos, we manually categorized their sentiment towards vaccination. We identified that 1,290 videos were negative (65.02%), 414 videos were positive (20.87%), 280 videos were neutral (14.11%). This is the first time that a large-scale video-sentiment analysis on vaccination has been achieved.

3.2.2 Measuring Network Centrality

To analyze the network of recommended videos and test whether there is a potential relationship between video’s network position (centrality) and their sentiment, we used Social Network Analysis (SNA). We considered the normalized values of five network measures: betweenness centrality, in-closeness centrality, out-closeness centrality, in-degree centrality, and out-degree centrality. These centrality measures were selected to understand the prominence of a video in the YouTube network.

Betweenness centrality was defined based on Freeman’s betweenness measurement, which is identified by counting the number of shortest paths between any two nodes that pass through a given node, such that videos with higher betweenness would theoretically have more “control” over the network [66-68]. Closeness centrality was calculated based on the reversed shortest distance between a given node and all other nodes in the network. In-closeness centrality considered paths from other videos to a given video, and out-closeness considered paths from a given video to other videos [69]. Finally, degree centrality was calculated based on the number of direct connections to and from other videos. Specifically, in-degree centrality measured the number of incoming direct connections to a video, and out-degree centrality measured the number of outgoing direct connections from a given video. High in-degree videos would be videos that are recommended based on many other videos; and high out-degree videos are those that lead to many other videos based on YouTube’s recommendation algorithm.
We used UCINET6.6 to calculate the network centralities measures of the whole network (9,489 nodes). To measure the variance of centrality measures between pro-vaccine and anti-vaccine videos, we ran T-test on UCINET using 10,000 permutations. Running T-test on UCINET allowed node-level data to be permuted, accounting for the non-independence of node-level videos in the network.

We also measured the effects of video attributes (e.g., video categories, view count, view count, like count, dislike count, comment count) on vaccine sentiment using chi-square test and binomial logistic regression in SPSS23.

4. RESULTS

4.1 Vaccine Sentiment and Centrality

Using the negative sentiment videos as the control group, we ran a two-sample one-tailed T-test. Anti-vaccine (negative sentiment) videos were coded as Group 1 and pro-vaccine (positive sentiment) videos were coded as Group 2. Both anti-vaccine and pro-vaccine videos have the same mean of out-degree and in-degree centralities (p<0.00). See Table 1. At the same time, anti-vaccine videos have higher values of in-closeness (p<0.05) and out-closeness centrality, (p<0.0005) but lower betweenness centrality (p<0.005). This suggests that anti-vaccine videos are easier to reach in the video-related video network, such that if the viewer started with an anti-vaccine video, subsequent recommended videos are more likely to be anti-vaccine in nature.

Table 1. Significance of vaccine sentiment and network centrality measures

| Centrality measures | Mean of Group 1 | Mean of Group 2 | Difference in means (G1 – G2) | Significance |
|---------------------|----------------|----------------|-------------------------------|--------------|
| Out-degree          | 0.004          | 0.004          | 0.000 **                     |              |
| In-degree           | 0.003          | 0.003          | 0.000                         | 0.9735       |
| Out-closeness       | 0.240          | 0.232          | 0.008 **                     |              |
| In-closeness        | 0.183          | 0.179          | 0.004 **                     |              |
| Betweenness         | 0.001          | 0.002          | -0.001 **                    | 0.0021**     |

Anti-vaccine sentiments are coded as Group 1. Pro-vaccine sentiments are coded as Group 2.

**p <0.005    ***p <0.0005

4.2 Video Sentiment and Video Categories

A chi-square test of independence was conducted between vaccine sentiment of the videos and the categories these videos are situated in. No videos directly about vaccines were found in the following YouTube video categories: auto vehicle, film & animation, gaming, how-to & style, pets & animals, travel & events, shows, and movies.

Anti-vaccine videos were more likely to be in the “News & Politics” and “People & Blogs” categories; some examples of popular videos belong to the “Alex Jones Show,” and videos related to the premiere of “Vaxxed: From cover up to catastrophe.” Pro-vaccine videos were more likely to be in the “Education” and “Science & Technology” category. There was a statistically significant association between video categories and vaccine sentiment, X^2 (7) = 43.678%, p<0.005. The association was small Cramer’s V = 0.160. See Table 2.

| Sentiment | Comedy | Education | Entertainment | Music | News | Nonprofits & Activism | Science & Technology | Total |
|-----------|--------|-----------|---------------|-------|------|------------------------|----------------------|-------|
| **Negative** | Count | 22 | 265 | 71 | 8 | 385 | 147 | 305 | 87 | 1290 |
| % | 1.7 | 20.5 | 5.5 | 0.6 | 29.8 | 11.4 | 23.6 | 6.7 | 0.005 | 100 |
| Adjusted residual | -1.8 | -3.4 | 0 | -1.2 | 3.5 | 0 | 3.1 | -3.9 | - |
| **Positive** | Count | 13 | 118 | 23 | 5 | 87 | 47 | 68 | 53 | 414 |
| % | 3.1 | 28.5 | 5.6 | 1.2 | 21.0 | 11.4 | 16.4 | 12.8 | 0.001 | 100 |
| Adjusted residual | 1.8 | 3.4 | 0 | 1.2 | -3.5 | 0 | -3.1 | 3.9 | - |
| Total | Count | 35 | 383 | 94 | 13 | 472 | 194 | 373 | 140 | 1704 |

4.3 Vaccine Sentiment and Video Properties

A binomial logistic regression was performed to ascertain the effects of view count, comment count, dislike/like ratio, and like count on the likelihood that vaccine-videos are positive or negative. We tested for linearity of the continuous variables (e.g., view count, like count, dislike count, comment count) via the Box-Tidwell procedure. A Bonferroni correction was applied using all nine terms of the model (including natural log transitions of continuous variables) to adjust the statistical significance to be accepted when p<0.0625. All continuous independent variables were found to be linearly related to the logit of vaccine sentiment.

To test for outliers, we conducted a case-wise diagnostic and identified that there were two studentized residual with a standard deviation of -7.302 and -10.642. After examining the data and its distribution, we determined that they were not outliers and they were kept in the analysis. The logistic regression model was statistically significant X^2 (5) = 100.485 (p<0.0005). The model explained 9.3% of the variance in vaccine sentiment and correctly classified 78.2% of the cases. Of the five variables, dislike count, dislike/like ratio, like count, and view count were statistically significant (see Table 3). The larger the dislike/like ratio, the more a video is disliked in proportion to likes. The ratio is an indicator of controversial content or poor-quality content deemed by the viewers. Interestingly, we found that videos with higher dislike/like ratio have 3.912 times higher odds of being a video that is supportive of vaccines.

Table 3. Logistic regression predicting likelihood of negative vaccine sentiment based on dislike count, dislike/like ratio, like count, view count, and comment counts

| Video attributes | B    | SE   | Wald  | df  | P     | Odds Ratio |
|------------------|------|------|-------|-----|-------|------------|
| Dislike count    | -0.004 | 0.001 | 8.002 | 1  | 0.005** | 0.996      |
| Dislike/like ratio | 1.364 | 0.218 | 39.125 | 1  | 0.001** | 3.912      |
| Like count       | 0.000 | 0.000 | 7.963 | 1  | 0.005** | 1.000      |
| View count       | 0.000 | 0.000 | 4.838 | 1  | 0.028*  | 1.000      |
| Comment count    | 0.000 | 0.000 | 0.004 | 1  | 0.948 | 1.000      |
| Constant         | -1.085 | 0.400 | 7.348 | 1  | 0.007* | 0.338      |
5. DISCUSSION

5.1 Anti-vaccine Sentiment Dominates YouTube

Mounting evidence suggests that imagery and video-based platforms tend to generate more shares than text-based forms or URLs [70,71]. Our research is consistent with previous research, showing that YouTube has been a harbour for anti-vaccine sentiment, which has a higher prevalence of anti-vaccine sentiment than text-based social media platforms observed in the Americas, Southeast Asia, and Eastern Europe [48,72]. Furthermore, our research shows that pro-vaccine videos are significantly positively correlated with more dislikes than likes (dislike/like ratio).

To explain this phenomena, we look into past research that showed that negative vaccine sentiment spreads more easily then pro-vaccine messages on Twitter, and pro-vaccine messages are even observed to backfire and become subjects of attacks [73]. Another possible explanation on the prevailing anti-vaccine sentiment is related to the democratic nature of social media; with real-time free-speech comes unverified information, misinformation, and conspiracy theories that are often uncredible. Venkatanaran et al.’s research observed that the freedom of speech is positively correlated to beliefs that vaccines are linked to autism [74].

Our study also adds to the discussion of social media’s structural influence on alternative narratives – the threshold-free access and algorithm catered to user-retention allows for clickbait, false news, and alternative facts to be amplified by botnets and astroturfers [75,76]. Starbird and colleagues studied rumour networks in the U.S. on Twitter and identified the phenomena of “alternative narratives” which was uniquely driven by the public’s tendency to be manipulated by conspiracy theorists, broken epistemologies, and political disinformation (ibid). Whether these false beliefs have an impact on real-life events (e.g., measles outbreaks, US presidential electoral results, and the flat-earth movement, etc.) remains contentious. To an extent, any decision-making is inarguably driven by “affect, cognitive, and conative processes” [77]. Based on our findings, one may argue that recent systematic reviews that have reported suboptimal outcomes of health promotion online may need to factor in the structural properties of social media and the decision-making processes of the public. That is to say, the traditional (yet popular) fact-based one-way information-sharing method needs a drastic makeover.

5.2 The High Closeness Centrality of Anti-Vaccine Videos in the Video Network

This research is the first to inspect the network centrality of pro- and anti-vaccine sentiment on YouTube. We have confirmed that anti-vaccine videos in the vaccine-related video network has higher in-closeness and out-closeness centrality, this would suggest that once a YouTuber watched a video with keywords related to negative sentiment, then it is more likely that subsequent recommended videos will also be anti-vaccine. This result may be indicative that searching for vaccine information online solidifies a person’s sentiment towards vaccines depending on that first keyword that a user search for, which serves to reinforce one’s preconceived opinions of vaccines [78]. Therefore, we propose that popular anti-vaccine recommended videos should be actively monitored. Active monitoring may assist in contributing to developing recommendation algorithms that more closely reflects the needs of the public – that is, to search for credible information, and to easily access credible vaccine information online.

Another useful application of recognizing the higher closeness centrality of anti-vaccine videos is to eliminate the “bubble” effect [79]. In a high-closeness centrality network, pro-vaccine videos are likely overlooked if a user is stuck in the anti-vaccine sentiment bubble. This could potentially lead to “majority illusion” [80] – the theory that people believe that the opinions of “active others” represent the general sentiment on an issue. More vaccine-hesitancy research needs to be done on YouTube to disentangle whether YouTube’s algorithm proliferates a lesser-held belief (anti-vaccine) dominated by those that occupy more active/central nodes could lead inactive others (vaccine hesitant) to change their beliefs.

Although researchers have yet to find online health-seeking behaviors lead to negative health outcomes, social media’s apomediary function (meaning to be removed of gatekeepers of information) indicate that misinformation will be rampant, and therefore requires novel approaches of active infoveillance [81]. One of the findings in our research shows that video categories that are more likely to demonstrate anti-vaccine sentiment rest in the “People & Blogs” and “News & Politics” sections, which reflects the fact that people who are anti-vaccine post heavily in these two categories instead of “Educational” or “Science & Technology” where most pro-vaccine sentiment videos are located. This research echoes the observation that social media’s “openness, participatory, collaborative” functions have brought about apo-mediation. This means that intermediaries (e.g., physicians) are predicted to be removed from the vaccine-information seeking process, and are replaced by YouTube’s content contributors (laymen or professional), and as such, the experiences of the individual is placed above the opinions and expertise of the providers’ [82,83].

5.3 Limitations and Future Research

This research did not measure the duration of the video, the upload year (i.e., how long has the video been uploaded), length of the video, whether it is cross-promoted content on other social media platforms (such as movies and music videos), or is it a viral video cross-posted in other social media platforms (e.g., Instagram, Snapchat, etc.). These are variables that we would recommend looking at in the future. We would also recommend examining the actual content of the video aside from sentiment, adopting a method similar to text-based semantic analysis like LIWC to determine the variables that influence vaccine-hesitancy such as “emotional appeal”, “personal experience,” “conspiracy,” “vaccine safety,” “self-governance,” and keywords are frequently mentioned as reasons for hesitancy [6,47,54].

In terms of coding, the process of manual coding sentiments of visual-audio data is labour-intensive and time-consuming. To date, human coding is more accurate than current sentiment analysis software due to our ability to process visual information in conjunction to audio-processing [84]. However, future research may benefit from opensource software like Google’s TensorFlow\(^1\) that has developed complex neural networks that can aid in processing audio-visual content in a timely, iterative manner from multiple social media platforms [85].

6. CONCLUSION

This research confirms that YouTube’s dominant sentiment towards vaccines is anti-vaccine (65.02%), which suggests a

\(^1\) TensorFlow. URL: https://www.tensorflow.org
potentially growing prevalence of anti-vaccine sentiments on imagery/video-based social media platforms. An encouraging step towards addressing vaccine hesitancy from a network perspective includes active monitoring of vaccine sentiment on YouTube and other imagery/video-based social media platforms. In addition, public health institutions and technology companies, such as Google, would benefit from collective action by, for example, fine tuning the recommendation system algorithm to reduce the centrality of anti-vaccine videos and videos containing misinformation at large. Second, YouTube’s recommended videos are likely popular, captivating videos - unlike traditional educational videos that are dominating the YouTube pro-vaccine sphere. Adopting marketing and communications strategies to create engaging vaccine awareness content is another way of breaking the “misinformation bubble” without jeopardizing YouTube’s viewership retention.

It would be useful to apply concepts in political research such as the polarization of public opinion [86] to look for new ways to approach and understand the dichotomized vaccine sentiment. For instance, rethinking the anti-vaccine movement from an activist framework (e.g., The 99% movement, the Arab Spring movement, and the Euromaidan revolution in Ukraine, etc.). Returning to the concept of apo-mediation, our neighbours’ opinion (or in this case, the online YouTube video network’s opinion) is increasingly favorable to a rigid CDC website. To embrace the unique functions of social media’s openness is to embrace that we are experiencing multiple paradigmatic clashes. Living in a post-modern society, the public is free to interpret information and evidence away from science, away from facts, and closer to a level that is relatable to the general populace – one that appeals to emotions, risk-negation, and truth-seeking. As philosopher Bruno Latour noted, scientists addressing scientific evidence needs to address “matters of concern”, in which networks of people form opinions based on not evidence but emotional appeal [89]. This research acknowledges that a “knowledge deficit” oriented intervention may be complimented by incorporating an in-depth understanding of the network in which misinformation resides.

7. REFERENCES

[1] CDC 1999. Ten great public health achievements–United States, 1900-1999. MMWR. Morbidity and mortality weekly report. 48, 12 (1999), 241.
[2] Andre, F.E. et al. 2008. Vaccination greatly reduces disease, disability, death and inequity worldwide. Bulletin of the World Health Organization. 86, 2 (2008), 140–146.
[3] Blume, S. 2006. Anti-vaccination movements and their interpretations. Social Science & Medicine. 62, 3 (Feb. 2006), 628–642.
[4] Dubé, E., Vivion, M., and MacDonald N. E. 2015. Vaccine hesitancy, vaccine refusal and the anti-vaccine movement: influence, impact and implications. Expert Rev. Vaccines 14, 1 (Jan. 2015), 99–117.
[5] Bolton, K., Memory, K., and McMillan, C. 2015. Herd immunity: Does social media affect adherence to the CDC childhood vaccination schedule? Pursuit-The Journal of Undergraduate Research at the University of Tennessee. 6, 1 (2015), 5.
[6] Kata, A. 2010. A postmodern Pandora’s box: Anti-vaccination misinformation on the Internet. Vaccine. 28, 7 (Feb. 2010), 1709–1716.
[7] Kata, A. 2012. Anti-vaccine activists, Web 2.0, and the postmodern paradigm – An overview of tactics and tropes used online by the anti-vaccination movement. Vaccine. 30, 25 (May 2012), 3778–3789.
[8] Omer, S.B., Salmon, D. A., Orenstein, W. A., deHart, M. P., and Halsey, N. 2009. Vaccine refusal, mandatory immunization, and the risks of vaccine-preventable diseases. New England Journal of Medicine. 360, 19 (2009), 1981–1988.
[9] Bloom, B.R., Marcuse, E., and Mnookin, S. Addressing vaccine hesitancy. Science. 344, 6182 (Apr. 2014), 339–339.
[10] Larru, B. and Offit, P. 2014. Communicating vaccine science to the public. Journal of Infection. 69, (Nov. 2014), S2–S4.
[11] Feinberg, Y. et al. 2015. Understanding public perceptions of the HPV vaccination based on online comments to Canadian News Articles. PLOS ONE 10, 6 (Jun. 2015), e0129587.
[12] Poland, G.A. and Jacobson, R.M. 2011. The age-old struggle against the Antivaccinationists. New England Journal of Medicine. 364, 2 (Jan. 2011), 97–99.
[13] Sadaf, A., Richards, J.L., Glanz, J., Salmon, D.A., and Omer, S.B. 2013. A systematic review of interventions for reducing parental vaccine refusal and vaccine hesitancy. Vaccine. 31, 40 (Sep. 2013), 4293–4304.
[14] Nyhan, B., Reifler, J., Richey, S., and Freed, G.L. 2014. Effective messages in vaccine promotion: A randomized trial. Pediatrics. 133, 4 (Apr. 2014), e835–e842.
[15] Nowak, G.J., Gellin, B.G., MacDonald, N.E., and Butler, R. 2015. Addressing vaccine hesitancy: The potential value of commercial and social marketing principles and practices. Vaccine. 33, 34 (Aug. 2015), 4204–4211.
[16] Betsch, C., Ulshofer, C., Renkewitz, F., and Betsch, T. 2011. The influence of narrative v. statistical information on perceiving vaccination risks. Medical Decision Making. 31, 5 (Sep. 2011), 742–753.
[17] Coleman, R. and Wu, H.D. 2010. Proposing emotion as a dimension of affective agenda setting: Separating affect into two components and comparing their second-level effects. Journalism & Mass Communication Quarterly. 87, 2 (Jun. 2010), 315–327.
[18] Betsch, C. and Sachse, K. 2013. Debunking vaccination myths: Strong risk negations can increase perceived vaccination risks. Health Psychology. 32, 2 (2013), 146–155.
[19] Beck, U. 1992. Risk society: Towards a new modernity. Sage.
[20] Lupton, D. 1993. Risk as moral danger: the social and political functions of risk discourse in public health. International journal of health services. 23, 3 (1993), 425–435.
[21] Petersen, A., Petersen, A., and Bunton, R. 1997. Risk, governance and the new public health. Foucault, health and medicine (1997), 189–206.
[22] Vasconcellos-Silva, P.R., Castiel, L.D., and Griepp, R.H. 2015. The media-driven risk society, the risk of antivaccination movement. Ciência & Saúde Coletiva. 20, 2 (Feb. 2015), 607–616.
[23] Kahan, D.M. 2014. Vaccine risk perceptions and Ad Hoc risk communication: An empirical assessment. (2014).
[24] Sadovykh, V. and Sundaram, D. 2017. Decision making processes in online social networks: A comparative analysis of health and financial online social networks. Proceedings of the 50th Hawaii International Conference on System Sciences (2017).
[58] Briones, R., Nan, X., Madden, K., and Waks, L. 2012. When vaccines go viral: An analysis of HPV vaccine coverage on YouTube. *Health Communication*. 27, 5 (Jul. 2012), 478–485.

[59] Basch, C.H., Zybert, P., Reeves, R., and Basch, C.E., What do popular YouTube™ videos say about vaccines? *Child: Care, Health and Development*. (Jan. 2017) DOI: 10.1111/ech.12442

[60] Dubé, E., Gagnon, D., and MacDonald, N.E., 2015. Strategies intended to address vaccine hesitancy: Review of published reviews. *Vaccine*. (Apr. 2015).

[61] Odone, A. et al. 2015. Effectiveness of interventions that apply new media to improve vaccine uptake and vaccine coverage: a systematic review. *Human vaccines & Immunotherapeutics*. 11, 1 (2015), 72–82.

[62] Wilson, K., Atkinson, K., and Deeks, S. 2014. Opportunities for utilizing new technologies to increase vaccine confidence. *Expert review of vaccines*. 13, 8 (2014), 969–977.

[63] Robichaud, P. et al. 2012. Vaccine-critical videos on YouTube and their impact on medical students’ attitudes about seasonal influenza immunization: a pre and post study. *Vaccine*. 30, 25 (2012), 3763–3770.

[64] Bendersky, M., Garcia-Pueyo, L., Harmsen, J., Josifovski, V., and Lepikhin, D. 2014. Up next: retrieval methods for large scale related video suggestion. *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining* (Aug. 2014), 1769–1778.

[65] Rieder, B. 2015. YouTube data tools. computer software. vers. 1.0.,” (May 2015). Available: https://tools.digitalmethods.net/netvizz/youtube.

[66] Hanneman, R.A. and Riddle, M. 2005. *Introduction to social network methods*. University of California Riverside.

[67] Muller-Prothmann, T. 2008. Use and methods of social network analysis in knowledge management. *Knowledge Management: Concepts, Methodologies, Tools, and Applications*. IGI Global. 1096–1106

[68] Haythornthwaite 1996. Social network analysis: An approach and technique for the study of information exchange. *ALISE conference*. (1996).

[69] McFarland, D., Messing, S., Nowak, M., and Westwood S.J. 2010. Social network analysis labs in R. [Online]. Available: https://sna.stanford.edu/lab.php?l=4. (2010).

[70] Bakhshi, S., Shamma, D.A., and Gilbert, E. 2014. Faces engage us: photos with faces attract more likes and comments on Instagram. (2014), 965–974.

[71] Strekalova, Y.A. and Krieger, J.I. 2015. A picture really is worth a thousand words: Public engagement with the National Cancer Institute on social media. *Journal of Cancer Education*. (Sep. 2015).

[72] The United Nations Children’s Fund. 2013. Tracking anti-vaccine sentiment in Eastern European social media networks. Working paper, UNICEF. (2013).

[73] Salathe M., and Kennedy, B., 2013. On Twitter, anti-vaccination sentiments spread more easily than pro-vaccination sentiments. [Online]. Available: http://science.psu.edu/news-and-events/2013-news/Salathe4-2013.

[74] Venkatraman, A., Garg, N., and Kumar, N., 2015. Greater freedom of speech on Web 2.0 correlates with dominance of views linking vaccines to autism. *Vaccine*. 33, 12 (Mar. 2015), 1422–1425.

[75] Starbird, K. 2017. Information Wars: A window into the alternative media ecosystem. *Medium*. (Mar. 2017).

[76] Harris, J.K., Moreland-Russell, S., Choucair, B., Mansour, R., Staub, M., and Simmons, K. 2014. Tweeting for and against public health policy: Response to the chicago department of public health’s electronic cigarette twitter campaign. *Journal of Medical Internet Research*. 16, 10 (Oct. 2014), e238.

[77] Papacharissi, Z. 2015. *Affective publics: Sentiment, technology, and politics*. Oxford University Press.

[78] Yom-Tov, E. and Fernandez-Luque, L. 2014. Information is in the eye of the beholder: Seeking information on the MMR vaccine through an internet search engine. *AMIA Annual Symposium Proceedings* (2014), 1238.

[79] Resnick, P., Garrett, R.K., Kriplean, T., Munson, S.A., and Stroud, N.J. 2013. Bursting your (filter) bubble: strategies for promoting diverse exposure. *Proceedings of the 2013 conference on Computer supported cooperative work companion* (2013), 95–100.

[80] Lerman, K., Yan, X., and Wu, X.-Z. 2016. The “majority illusion” in social networks. *PLOS ONE*. 11, 2 (Feb. 2016), e0147617.

[81] Eysenbach, G. 2002. *Infodemiology: the epidemiology of (mis) information*. *The American journal of medicine*. 113, 9 (2002), 763–765.

[82] Eisenstein, M. 2014. Public health: an injection of trust. *Nature*. 507, 7490 (2014), SI7–S19.

[83] Hesse, B.W. et al. 2005. Trust and sources of health information: the impact of the Internet and its implications for health care providers: findings from the first Health Information National Trends Survey. *Archives of Internal Medicine*. 165, 22 (2005), 2618–2624.

[84] Wöllmer, M. et al. 2013. YouTube movie reviews: Sentiment analysis in an audio-visual context. *IEEE Intelligent Systems*. 28, 3 (May 2013), 46–53.

[85] Abadi M. et al. 2016. “TensorFlow: large-scale machine learning on heterogeneous distributed systems” *ArXiv160304467 Cs*, Mar. 2016.

[86] Conover, M., Ratkiewicz, J., Francisco, M.R., Gonçalves, B., Menczer, F., and Flammini A., 2011. Political polarization on twitter. *ICWSM*. 133, (2011), 89–96.

[87] Dvir-Gvirsman, S. 2016. Media audience homophily: Partisan websites, audience identity and polarization processes. *new media & society* (2016).

[88] Gruzd, A. and Roy, J. 2014. Investigating political polarization on Twitter: A Canadian perspective. *Policy & Internet*. 6, 1 (2014), 28–45.

[89] Latour, B. 2004. Why has critique run out of steam? From matters of fact to matters of concern. *Critical inquiry*. 30, 2 (2004), 225–248.