Title
A Qualitative Analysis of COVID-19 Pediatric Vaccine Misinformation by Verified Twitter Users of Minority Descent

Author
Le, Nicolette Olivia

Publication Date
2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

A Qualitative Analysis of COVID-19 Pediatric Vaccine Misinformation by Verified Twitter Users of Minority Descent

A thesis submitted in partial satisfaction of the requirements

for the degree Master of Arts

in

Global Health

by

Nicolette Olivia Le

Committee in charge:

Professor Timothy Mackey, Co-Chair
Professor Georgia Robins Sadler, Co-Chair
Professor Rebecca Fielding-Miller

2022

The thesis of Nicolette Olivia Le is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2022

TABLE OF CONTENTS

THESIS APPROVAL PAGE iii

TABLE OF CONTENTS iv

LIST OF FIGURES v

LIST OF TABLES vi

LIST OF ABBREVIATIONS vii

ACKNOWLEDGEMENTS viii

ABSTRACT OF THE THESIS ix

INTRODUCTION 1

METHODS 10

RESULTS 21

DISCUSSION 32

CONCLUSION 32

REFERENCES 40

LIST OF FIGURES

Figure 1: Flowchart of reply scheme classification of tweets from Twitter. The direction of the arrows demonstrates how first-layer are directed at and/or are a response to the “parent” tweet, while second-layer replies are directed at and/or are a response to a first-layer reply. 17

Figure 2: Brief flowchart overview of data collection and qualitative analysis. This study specifically focuses on tweets with explicit misinformation about the COVID-19 pediatric vaccine authored by users of ethnic or racial minority descent and the responses to said misinformation. 20

Figure 3: Percentage of dataset related to top 200 most retweeted tweets calculation............... 23

Figure 5: Top five most common single-string terms present in Twitter user bio of users who authored signal replies to misinformation “parent” tweet by frequency. 28

Figure 4: Relative-scaled word cloud visualization of single-string text from Twitter user bios of users who authored signal replies to misinformation “parent” tweet. Words that appear larger in the word cloud appeared more frequently in bios. 29

Figure 6: Stacked bar graph of self-reported political identification from Twitter user bio and response to misinformation "parent" tweet bar.. 31

Figure 7: 100% Stacked bar graph of self-reported political identification from Twitter user bio and response to misinformation "parent" tweet bar.. 31

.

LIST OF TABLES

Table 1: Examples of content-coded tweets for themes related to vaccine confidence for the COVID-19 pediatric vaccine. Themes were identified after a round of manual inductive content classifcation. Examples are paraphase and/or redatcted to retain anonymity...... 14

Table 2: Examples of content-coded tweets for themes related to vaccine concern for the pediatric COVID-19 vaccine. Themes were identified after a round of manual inductive content classifcation. Examples are paraphase and/or redatcted to retain anonymity...... 15

Table 3: All 196 stop words used in Twitter user bio analysis sorted in alphabetical order. These were omitted from the generation of the word cloud and frequency count. 19

Table 4: Four identified misinformation "parent" tweet authored by Twitter-verified users from minority communities. The table lists the main theme of each tweet, number of replies, and users' information. 25

Table 5: Sentiment analysis of signal replies to the four misinformation "parent" tweets. Counts and percentages of agree, disagree, and undefined sentiments for each misinformation "parent" tweet (individually) and for all four misinformation "parent" tweets (in total) are presented. 26

Table 6: Unique users' self-reported identification of parental/grandparental status and political leaning in Twitter user bio section. Counts and percentages for each misinformation "parent" tweet (individually) and for all four misinformation "parent" tweets (in total) are presented. 27

Table 7: Self-reported political identification in Twitter user bio and sentiment to misinformation "parent" tweet.. 30

LIST OF ABBREVIATIONS

API	Application Programing Interface
BTM	Biterm Topic Model
NPL	Natural Language Processing
RT	Retweets

ACKNOWLEDGEMENTS

I would like to acknowledge and thank my advisor Professor Timothy Mackey for making this work possible. The completion of this study would not have been possible without his continued encouragement and guidance. I would also like to acknowledge Tiana McMann for advice throughout the process. I would also like to thank my committee members, Professor Georgia Robins Sadler and Professor Rebecca Fielding-Miller for their comments, suggestions, and continuous support. Finally, I would like to express my deep appreciation for my family and friends for their unwavering support throughout my undergraduate and graduate studies.

ABSTRACT OF THE THESIS

A Qualitative Analysis of COVID-19 Pediatric Vaccine Misinformation

by Verified Twitter Users of Minority Descent

by

Nicolette Olivia Le

Master of Arts in Global Health

University of California San Diego, 2022

Professor Timothy Mackey, Co-Chair
Professor Georgia Robins Sadler, Co-Chair

The aims of this study were to qualitatively characterize sentiments of Twitter users on the topic of the COVID-19 vaccine for children, specifically in response to tweets with explicit vaccine misinformation that are authored by users who are both verified by Twitter and of minority descent. The investigative approach was through a review of the literature and content classification of tweets collected from the Twitter API. A total of 863,007 tweets were collected.

From which, the 200 most retweeted tweets were subjected to manual content classification to identify four tweets with explicitly vaccine misinformation that are authored by users who are both verified by Twitter and of minority descent. The replies to these four tweets were collected from the Twitter API and subjected to manual content classification to identify themes and the Twitter bios of users who authored these replies were subjected to manual content classification to identify self-reported race, ethnicity, status as a parent or grandparent, and political affiliations. The results of this study provide insight into online sentiments surrounding the COVID-19 pediatric vaccine and specifically organic user reactions to explicit vaccine misinformation. Further studies should examine other themes related to social media-based discussions of misinformation both regarding COVID-19 misinformation and broadly scientific misinformation to better inform public health communication and improve public trust in scientific advancements.

INTRODUCTION

COVID-19 Infodemic

Accompanying the current COVID-19 pandemic has been a massive infodemic, defined by the WHO as "too much information including false or misleading information in digital and physical environments during an outbreak" (*Infodemic*, 2022). Around the world, nations witnessed the "immediate and widespread" demand for and supply of scientific, medical, and technical information relating to the virus (*Infodemic*, 2022; The Lancet Infectious Diseases, 2020; Farooq & Rathore, 2021). This includes theories and assertions with unclear origins, yet-to-be peer-reviewed scientific research (e.g., conclusions made from preprints), and misconstrued or misinterpreted findings from published research studies (Naeem & Bhatti, 2020). Myths about the virus's origins (e.g., 5G), infection prevention (e.g., eating garlic), disease transmission (e.g., through mosquitoes), and disease treatment (e.g., chloroquine) are all commonly disseminated pieces of false or misleading information (Naeem & Bhatti, 2020).

This vast circulation of false information diluted the impact of scientific, medical, and public health efforts to mitigate the spread of the virus, discouraged uptake of COVID-19 vaccines, and creates uncertainty when the public is looking for trusted information (Naeem & Bhatti, 2020; The Lancet Infectious Diseases, 2020). Cuan-Baltazar et al. demonstrated this in their study, which critically analyzed the first 110 websites from a Google search of the term "Wuhan virus" conducted on February 6, 2020 (Cuan-Baltazar et al., 2020). The authors found that less than two percent of the websites had a Health on the Net Foundation Code of Conduct seal, a common indicator a health or medical website's reliability (Cuan-Baltazar et al., 2020). They also found that over 60 percent of websites did not have any of the four categories (e.g, authorship, attribution, disclosure, and currency) of the JAMA benchmark, a common assessment of quality and credibility for health websites (Cuan-Baltazar et al., 2020). Further, none of the websites had high DISCERN scores, an instrument used to evaluate the quality of consumer health information (Cuan-Baltazar et al., 2020).

Exposure to false and misleading information has also been studied. A 2020 survey of adults in the United Kingdom by the government-approved regulatory and competition authority for the industries of broadcasting and telecommunication, Ofcom (Office of Communications) found that 46 percent reported exposure to false or misleading information about the pandemic online (*Covid-19 News and Information: Consumption and Attitude*, 2020). Further, the survey found that 40 percent reported difficulties in recognizing false information about COVID-19 online (*Covid-19 News and Information: Consumption and Attitude*, 2020). Still, over 80 percent reported sharing information about COVID-19 in person and online (*Covid-19 News and Information: Consumption and Attitude*, 2020).

False information can be split into three sub-terms with distinct definitions based on intentionality: misinformation, disinformation, and mal-information (Santos-d'Amorim & Fernandes de Oliveira Miranda, 2021). Misinformation is defined as false information that is disseminated by those who do not intend to mislead (CDC, 2021). For example, an agent of misinformation is someone who shares an unverified claim about COVID-19 that they heard on the news to friends and family. Disinformation is defined as false information that is disseminated by those who intend to mislead (CDC, 2021). For example, an agent of misinformation is someone who knowingly shares doctored news footage to intentionally fuel conspiracy theories. Mal-information is information that is based on fact, but is shared with the intention cause harm (Santos-d'Amorim & Fernandes de Oliveira Miranda, 2021). For example, an agent of mal-information is someone who publishes the private information of a researcher with the intention of inciting harassment.

False, unvetted, and misleading information about COVID-19 is regarded as a major threat to public health and safety (Bin Naeem & Kamel Boulos, 2021). Since the beginning of the pandemic, false information regarding the pandemic's scale, origin, and treatment approaches has been shared through both traditional and social media (Bin Naeem & Kamel Boulos, 2021). According to the Bruno Kessler Foundation, social media platforms, in particular, saw an increase in users by 20 to 80 percent around the world during the pandemic, demonstrating their emergence as channels of "information seeking and sharing" (Bin Naeem & Kamel Boulos, 2021; *COVID-19 and Fake News in the Social Media*, 2020).

As vaccine development began, false information regarding its progress and approval of medical countermeasures also circulated widely (Muric et al., 2021). The U.S. Centers for Disease Control and Prevention (CDC) identifies four prevalent themes among disinformation and misinformation related to the COVID-19 vaccine: "vaccine development", "safety", effectiveness", and COVID-19 denialism" (CDC, 2021). COVID-19 vaccine-related misinformation is regarded to negatively impact vaccine confidence and uptake (Bin Naeem & Kamel Boulos, 2021; CDC, 2021; Coustasse et al., 2021). Loomba et. al found that misinformation was responsible for a six percent decline in intent among those in the UK and US who would "definitely accept a vaccine" (Loomba et al., 2021). Indeed, despite widespread availability of the COVID-19 vaccine and boosters, countries such as the United States and the United Kingdom continue to struggle to achieve high vaccine uptake, with both nations yet to reach the necessary coverage to achieve herd immunity among its population (Loomba et al., 2021). Further, the authors also found that exposure to misinformation had varied affects among different sociodemographic groups.

Vaccine Hesitancy and Disparities in Vaccine Uptake

Vaccine hesitancy is a well-studied topic (Jacobson et al., 2015; Kennedy, 2020). In recent years, routine vaccinations for children have met sizable resistance. Though some outright refuse vaccinations for their children, many still believe in the benefits of routine vaccination. (Jacobson et al., 2015). Instead, they disagree with the recommended immunization schedule. Increasingly parents are opting to delay or spread-out routine vaccinations believing that children are receiving too many vaccines at once (McKee & Bohannon, 2016; Smith et al., 2011). The sentiment has also been shared by high profile persons, including former President Donald Trump, who has previously suggested a link between autism and the recommended vaccine schedule (Belluz, 2015). In his words, "[J]ust the other day, 2 years old, beautiful child went to have the vaccine and came back and a week later got a tremendous fever, got very, very sick, now is autistic (Belluz, 2015).

It is important to note that vaccine hesitancy is not strictly partisan issue. A variety of social-cultural factors contribute to individuals being vaccine hesitant. These include falling public trust in the government, concerns over the growth of the pharmaceutical-industrial complex, fear of unethical scientific research, and the growth of interest in alternative medicine (Kennedy, 2020; McKee & Bohannon, 2016). Further contributing is the shift towards shared decision making between patients and clinicians in the medical field (McKee & Bohannon, 2016). This has empowered patients to question the recommendation of their clinician, and, specific to pediatric vaccines, embolden parents to question the recommendations of their child's pediatrician (McKee & Bohannon, 2016). As such, the concern and hesitancy shown towards the COVID-19 vaccine is not a new phenomenon. Still, in the case of infectious disease such as COVID-19, vaccines are one of the most effective public health interventions (Coustasse et al., 2021). It is thus necessary to examine and identify prevalent sentiments among those hesitant to provide support for interventions that can tackle potential misconceptions and address concerns. This is particularly important in communities that are disproportionately affected by disease. In the case of COVID-19, a focus on communities of color that have been disproportionately impacted by the disease is particularly necessary (Razai et al., 2021)

Despite having higher risks for COVID-19, communities of color in the US report lower COVID-19 vaccine confidence and acceptance rates (Malik et al., 2020; Szilagyi et al., 2021). A nationally representative sample of US adults surveyed found that 58 percent intended to receive the COVID-19 vaccine, followed by 32 percent who were unsure, and 11 percent who did not intend to receive the vaccine (K. A. Fisher et al., 2020). The authors also found that Black and Hispanic participants were less likely to respond that they intended to receive the COVID-19 vaccine. Attempts to identify factors associated with vaccine hesitancy among minority communities revealed themes of mistrust due to past unethical research, fears of politicization, and lack of confidence in government or public health authority figures as well as concerns due to structural barriers, such as access, cost, and inequitable distribution (Carson et al., 2021). Similar themes were found by Razai et al., who identified mistrust due to systemic racism, unethical research, and prior negative experiences with healthcare as major themes for vaccine hesitancy among minority communities (Razai et al., 2021).

Specific to the COVID-19 vaccine for children, studies have found that common themes specific to parents' intent to vaccinate their children include concerns for safety and efficacy, lack of trust in the government, and belief that the vaccine is unnecessary for children (Ruggiero et al., 2021; Szilagyi et al., 2021; Teherani et al., 2021). In the weeks before the anticipated FDA approval of the COVID-19 vaccine for children aged 5-11, 41 percent of parents surveyed responded that they planned to vaccinate their child (C. B. Fisher et al., 2021). This was followed by 35 percent who responded otherwise and 25 percent who were unsure. The authors also found that sentiments differed significantly across racial and ethnic demographics, with non-Hispanic Asian parents more likely to have planned to give their child the vaccine, while non-Hispanic Black and non-Hispanic White parents were more likely to be unsure and to report planning to not vaccinate their child, respectively.

Twitter Infovelliance

Established in 2006, Twitter is a social media networking service that is accessible through mobile applications and web browsers. With a focus on microblogging, users' posts are called "tweets". Tweets may be text alone or accompanied by an image, audio clip, or video clip. Text is limited to 280 characters, while audio clips are limited to 140 seconds for most accounts. Tweets are interacted with in three main ways, "liking" to show agreement or enjoyment, "retweeting" to share onto one's own Twitter feed, and "replying" to add a comment to the original tweet. Users who set their account's privacy to "public" are able to be "followed" by anyone who is a registered user of Twitter. All tweets posted or retweeted onto the page of a public Twitter account are visible to anyone viewing the page, including those without a Twitter account. Users who set their account's privacy to "private" must manually approve new "followers". All tweets posted or retweeted onto the page of a private Twitter account are only visible to those followers whom the user has approved. Users discover new tweets on their

“Home timeline,” which displays content suggested by the Twitter algorithm. These include content from within and outside of users’ personal network. In the words of the company, “We select each Tweet using a variety of signals, including how popular it is and how people in your network are interacting with it” (*About Your Home Timeline on Twitter*, n.d.). Essentially, users who interact with certain content or follow others who interact with certain content will be shown more tweets related to that content. This has been identified as a potential way in which echo chambers form, such as in the case of COVID-19 misinformation. Single keywords (e.g., “#COVID-19,” “#Brexit,” etc.) or phrases (e.g., “#MeToo,” “#GirlMedTwitter,” etc.) can be written with the “#” symbol to form a “hashtag.” Clicking on a hashtag in any tweet with show all public tweets that use that hashtag. The main function of hashtags is to index keywords that allow users to follow topics of interest.

Twitter has been identified as a major channel of COVID-19 vaccine misinformation (Bin Naeem & Kamel Boulos, 2021). However, Twitter infovelliance has been employed to gain insight into public sentiments long before the COVID-19 pandemic (Chew & Eysenbach, 2010; Howe et al., 2018). A study focused on the 2009 H1N1 outbreak analyzed 2 million tweets found that the use of “H1N1” significantly increased, demonstrating “gradual adoption of World Health Organization recommended terminology” (Chew & Eysenbach, 2010). The authors’ content analysis found “resource-related” tweets to be the most retweeted. They also found that website belonging to traditional news media rather than government or health organizations were the most commonly linked source (Chew & Eysenbach, 2010).

Efforts to analyze user sentiment on the COVID-19 vaccine have found that positive sentiment is more dominant on Twitter (Hussain et al., 2021; Kwok et al., 2021; Yousefinaghani et al., 2021). However, a study by Bonnevie et al. found an 80 percent increase in vaccine opposition on Twitter when comparing tweets from the months before and after COVID-19 cases increased in the United States (Bonnevie et al., 2021). The authors also found a statistically significant increase in references to public health officials or government authorities figures in vaccine opposition conversation. Relatedly, Muric et al. found that users who interacted with vaccine opposition content tended be conservative or right leaning (Muric et al., 2021). Studies also identifies themes surrounding concerns for children in relation to the COVID-19 vaccine (Osakwe et al., 2021; Scannell et al., 2021). This is consistent with a study which examined tweets mentioned in fact-check claims on Twitter. In this study, the authors found that tweets with misinformation seemed to be more “driven by concerns of potential harm to others” (Shahi et al., 2021).

Shahi et al., also found that Twitter-verified accounts (e.g., celebrities, brands, etc.) also participated in the dissemination of false information. Twitter-verified accounts are those that are determined by Twitter to be “authentic, notable, and active” (*About Verified Accounts*, n.d.). For accounts belonging to individual users, this means that the individual user’s identity has been verified by Twitter with a valid official government-issued form of identification (e.g., driver’s license, passport, etc.), the individual is also considered someone who is prominently recognizable (e.g., widely referenced in news coverage, possess a Wikipedia page, etc.), and the user is active with a good record of adherence to Twitter rules (*About Verified Accounts*, n.d.). Accounts belonging to non-individuals have slightly different qualifiers to demonstrate that they are, as required by Twitter to be considered a verified account “authentic, notable, and active” (*About Verified Accounts*, n.d.).

This study seeks to add to the existing infodemic and vaccine online communication literature by characterizing user sentiment surrounding the COVID-19 pediatric vaccine, specifically responses to vaccine misinformation authored by Twitter-verified users from accounts that identify as minority users.

METHODS

The study was done in two main phases. First, an explorative review of the existing literature on COVID-19 misinformation, disparities in vaccine uptake, and Twitter infovelliance was conducted to contextualize the study. Second, data analysis of tweets collected from the public streaming Twitter Application Programing Interface (API) and relating to the COVID-19 vaccine for children was conducted to examine Twitter-based misinformation authored by Twitter-verified minority users.

Literature Review

A non-systematic literature review was conducted to identify common themes in COVID-19 vaccine misinformation, with a specific focus on themes relevant to or made my those from communities of color on Twitter. The primary purpose of this literature review was to better understand and identify potential gaps within the current literature. A secondary purpose of this literature review was to examine coding schemes from previous studies to help establish a framework for coding the textual data collected for this study. The PubMed (Medline) and JSTOR databases as well as GoogleScholar search engine were searched using terms such as "COVID-19 misinformation"," COVID-19 vaccine disparities", "vaccine hesitancy", "parental attitudes pediatric vaccine", "minority vaccine uptake COVID-19", "Twitter infovelliance", and "content analysis COVID-19 tweets". These search queries were used for contextual information on COVID-19 misinformation, vaccine-specific misinformation, minority and parental attitudes towards the COVID-19 pediatric vaccine, and exploring recent research into Twitter-based dissemination of misinformation. Overall, these searches were also done to better contextualize the relationship between COVID-19 misinformation, vaccine hesitancy, and disparities in vaccine uptake as well as characterize the role of infodemiology (e.g., examining the determinants of health using electronic mediums) and infovelliance within these topics.

Peer-reviewed articles were eligible for inclusion if they were accessible in English-language, regardless of language of origin. The reference lists of eligible articles were reviewed to identify potential eligible articles.

Data Collection and Analysis

First-Round of Data Collection

Tweets related to the COVID-19 pediatric vaccine were collected from the public streaming Twitter API using keywords such as “COVID-19” and “pediatric.” From here tweets relevant to the topic of COVID-19 vaccine misinformation were identified using the biterm-topic model (BTM), an unsupervised machine learning approach that analyzes text data using natural language processing (NPL), to determine “highly correlated topic clusters” (Yan et al., 2013). BTM is a word co-occurrence-based topic model that uses “biterms,” which are a combination of two-words or word-word co-occurrence patterns, to split text (Yan et al., 2013). For instance, the text “coronavirus vaccine mandate” has three biterms: “coronavirus vaccine”, “coronavirus mandate” and “vaccine mandate”. BTM is well-suited for sparse text, as opposed to conventional topic models which model topics using document-level word co-occurrence patterns, because it directly models word-co-occurrences (Yan et al., 2013). Thus, the nature of text data on Twitter (i.e., having a 280-character limit) works particularly well for employing BTM without prior coding for content classification of the dataset. This method has been used in prior studies to identify thematic groups for further review on a number of public health issues, including for detection of COVID-19 misinformation (Mackey et al., 2021). This has proven to be useful in prior efforts to characterize unclassified, unstructured, short-form data, such as tweets (Cai et al., 2020; Filippou et al., 2020) BTM analysis was done by setting the BTM topic number (k) as 20 and then choosing clusters with misinformation-related topic clusters, and then extracting the top 200 most retweeted tweets from these clusters. The final dataset for these top 200 retweeted tweets was then content coded for this study's specific aims.

Content Classification

The identified top 200 most retweeted tweets were manually annotated for signal relevant to this study's aims (i.e., relation to the COVID-19 vaccine for children) using a binary coding scheme. Manual qualitative content classification of the identified signal tweets was done inductively, to construct a code book of themes, and deductively, using the codebook to code the same set of signal tweets for themes relating to COVID-19 vaccine confidence and COVID-19 vaccine concern for the pediatric vaccine. Themes related to vaccine confidence include "vaccine is safe" and "vaccine is effective", in which the reply tweet expresses the sentiment that the COVID-19 pediatric vaccine is safe to receive or effective for protection against COVID-19, respectively. Themes related to vaccine concern include "vaccine is unnecessary", "vaccine is experimental", "vaccine is a control tactic", and "vaccine development conspiracy". For these, the reply tweet expresses the sentiment that the COVID-19 pediatric vaccine is not necessary for children to receive for one reason or another, not safe for children to receive due to being or still being experimental, is part of a ploy by the government or other organization to control the public, or is being developed with motives other than protection against COVID-19. See **Table 1** and **Table 2** for examples of tweets in each thematic category. The given examples are paraphrased and/or redacted for anonymity purposes.

Verified User Classification

Signal tweets that were authored by Twitter-verified accounts belonging to individuals, as opposed to belonging to an organization or brand, were identified. The Twitter-verified users who authored these signal tweets were manually coded for the demographic factors of sex, race, ethnicity, and nationality using publicly available metadata and/or self-reported information from the last ten tweets on the user's account. From here the tweets authored by, verified racial and ethnic minority users (i.e., authors of Black, Asian, Pacific Islander/Native Hawaiian, Native American/Alaskan Native, and Hispanic descent) were identified. The data were collected to characterize and aggregate the potential impact of misinformation among Twitter verified users of specific racial and ethnic minority communities. This was only done for accounts that belonged to individuals. Tweets made by Twitter-verified users that were not individuals (e.g., CNN, Fox News, WHO, etc.) were excluded from this process and from analysis for the purposes of this study.

Ethics Approval: This study only conducted analysis of publicly available data and results presented are de-identified to ensure anonymity. Hence, IRB approval was not required for this study.

Table 1: Examples of content-coded tweets for themes related to vaccine confidence for the COVID-19 pediatric vaccine. Themes were identified after a round of manual inductive content classifcation. Examples are paraphase and/or redatcted to retain anonymity.

Theme	Explanation	Examples
Vaccine is safe	Tweet expresses sentiment about the COVID-19 pediatric vaccine being safe for children.	1. BREAKING—lower dose of the Pfizer-BioNTech #COVID19 vaccine — one-third the amount given to adults and teens — is safe and triggered a robust immune response in children 5-11 years old. […] [LINK] 2. BREAKING: Health Canada authorizes the use of the Pfizer-BioNTech COVID-19 vaccine in children 12 to 15 years of age. This is the first COVID-19 vaccine authorized in Canada for use in children.
Vaccine is protective	Tweet expressive sentiment about the COVID-19 pediatric vaccine being protective against COVID-19 in children.	1. Breaking News: A Pfizer-BioNTech trial found the vaccine extremely effective in 12- to 15-year-old [LINK] 2. 📌New CDC report shows surging pediatric #COVID19 hospitalizations […] And it's surging in all ages: 12-17 group, vaccine ineligible 5-11, and in babies/toddlers 0-4. […] [LINK]

Table 2: Examples of content-coded tweets for themes related to vaccine concern for the pediatric COVID-19 vaccine. Themes were identified after a round of manual inductive content classifcation. Examples are paraphase and/or redatcted to retain anonymity.

Theme	Explanation	Example
Vaccine is unnecessary	Tweet expresses sentiment about how the COVID-19 pediatric vaccine is unnecessary for children to receive.	1. Pfizer plans to request Emergency Use Authorization from the FDA based on data from its phase 2/3 trial for children ages 5 to 11, as experts question the company's data and need for kids to be vaccinated against COVID. [LINK] 2. […] why is Pfizer trying to get the FDA to approve their experimental mRNA vaccines for an age-group at virtually no-risk under "emergency use"?
Vaccine is experimental	Tweet expresses sentiment about how the COVID-19 pediatric vaccine is experimental or unsafe for children.	1. The #vaccine leaflets clearly state: No tests done for #infertility […] Now they want to vaccinate #children with an #experimental drug? Which has never been tested for #infertility? 😳 2. Nearly 400 children between the ages of 12 and 17 were diagnosed with heart inflammation after receiving @Pfizer's #Vaccine for #COVID19, according to a study published by @CDCgov. [LINK]
Vaccine is a control tactic	Tweet expresses sentiment about how the COVID-19 pediatric vaccine is a control tactic by an entity.	1. 1937 Nazi Germany heard the same call for the segregation of the unclean, […] Segregation is one step removed from extermination. @picardonhealth #onpoli 2. No entity can force American citizens to take the #COVID19 vaccine. It is still experimental and not approved by the FDA yet. […]
Vaccine development conspiracy	Tweet expresses sentiment questioning vaccine development and motives (i.e., discussing company history, funding sources, etc).	1. […] Pfizer claimed Trovan was "safe," but 181 kids were gravely injured + 11 died. [LINK] 2. Moderna vaccine is not morally produced. Unborn children died in abortions and then their bodies were used as "laboratory specimens". […]

Second Round of Data Collection

In a second round of data collection from the publicly streaming Twitter API these "parent tweets" from verified accounts were scrapped for the first layer of user replies and the publicly available metadata associated with the Twitter user accounts of those who authored replies. First layer user replies are defined as tweets that are intended to be directed at the "parent" tweet, rather than tweets that are intended to be directed at a reply to the "parent" tweet. See **Figure 1** for a visual explanation.

Content Classification of Replies

These first layer user replies were coded with a binary coding scheme to identify signal (i.e., relation to the parent theme). These signal replies were subsequently manually content coded for the sentiment of their response to the "parent tweet" (i.e., agree, disagree, undefined). Replies coded as "agree" expressed support towards the misinformation "parent" tweet. Replies coded as "disagree" expressed opposition towards the misinformation "parent" tweet. Replies coded as "undefined" did not clearly express support or opposition towards the misinformation "parent" tweet. This included things such as one-word exclamations (e.g., "LOL!") and emojis without further context (e.g., 🙄).

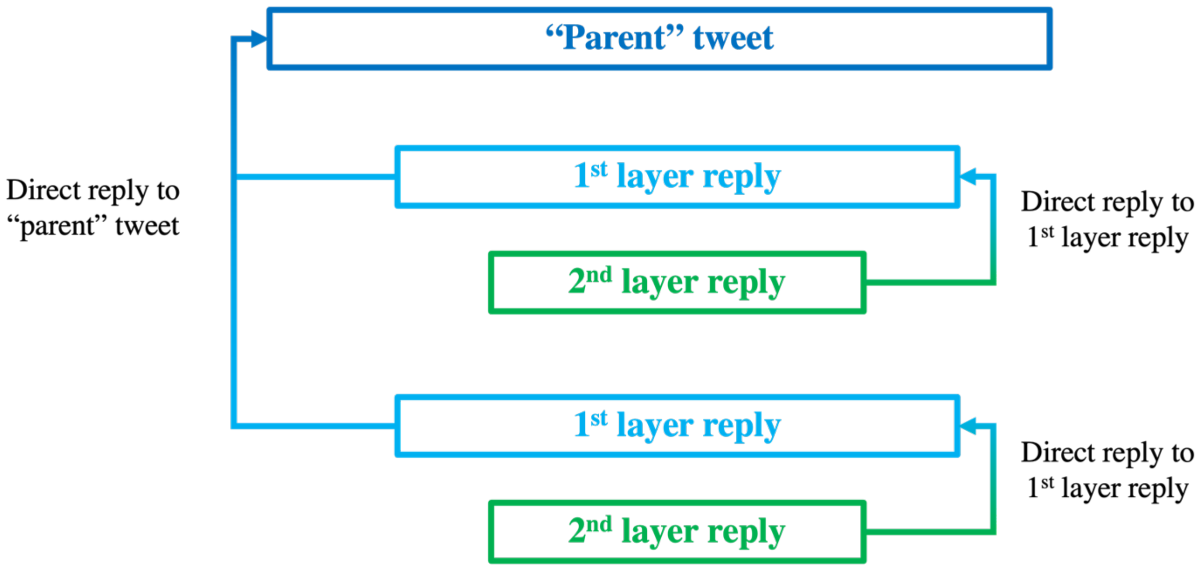

Figure 1: Flowchart of reply scheme classification of tweets from Twitter. The direction of the arrows demonstrates how first-layer are directed at and/or are a response to the "parent" tweet, while second-layer replies are directed at and/or are a response to a first-layer reply.

<u>Content Classification of Replies to Explicit Misinformation</u>

Of the "parent tweets" those with explicit COVID-19 vaccine misinformation, the specific focus of this study, were identified and selected for further analysis. While some sources differentiate between misinformation, disinformation, and mal-information based on the intentions of those who share false information, this study did not attempt to derive the intentions of users who authored tweets with false information. The term "misinformation" is used as an umbrella term to encompass all three subcategories of false information. The first layer replies to the misinformation "parent tweets" were further inductively coded to identify common themes present.

User Classification and Analysis of Users Who Replied to Misinformation

The Twitter bio of users who authored signal replies were manually coded for user self-reported identification of race and ethnicity (e.g., “#BlackInSTEM”, “Mexican-American”, etc), status as a parent or grandparent (e.g., “mom”, “dad of 2”), and political leaning (e.g., “MAGA”, “Liberal”, etc.) Data visualization of the text in user bios was conducted utilizing Python’s Matplotlib and WordCloud libraries. Word clouds are graphic visualizations of text used to highlight important keywords and quickly convey crucial information (Dubey, 2020). Word clouds have been used in prior studies to qualitatively analyze data from social media platforms, such as Twitter and Facebook (Dubey, 2020; Mhamdi et al., 2018).

A word cloud and the related frequency of count of strings were generated. Duplicate bios from duplicate users were removed. Bigrams, sequences of two consecutive strings (e.g., “free speech”) were excluded. Numbers, emojis, and symbols were also excluded. Plural words were normalized so that a word with a trailing “s” is counted with the count of the same word without a trailing “s” (e.g., the words “dogs” and “dog” would be counted under “dog”). This did not apply to words that ended in “ss.” Relative scaling was used so that a word that is twice as frequent appeared in the word cloud as twice as large. WordCloud’s default stop words list was used with the addition of stop words added in consideration of links commonly present in user bios was used to filter the text. Stop words are commonly used words in natural language that include words such as “a”, “in”, and “the” that contain little useful qualitative information. A total of 196 strings were in the stop word list. A full list can be found in **Table 3.** See **Figure 2** for a flowchart overview of the data collection and qualitative analysis process starting from data collection.

Table 3: All 196 stop words used in Twitter user bio analysis sorted in alphabetical order. These were omitted from the generation of the word cloud and frequency count.

Stop Words					
a	doing	his	**no**	**t**	we’re
about	don't	how	nor	than	we’ve
above	down	how's	not	that	well
after	during	however	-	that's	were
again	-	http	**of**	the	weren't
against	**each**	https	off	their	what
all	else	-	on	theirs	what's
also	ever	**i**	once	them	when
am	-	i’d	only	themselves	when's
an	**few**	i'll	or	then	where
and	for	i'm	other	there	where's
any	from	i've	otherwise	there's	which
are	further	if	ought	therefore	while
aren't	-	in	our	these	who
as	**get**	into	ours	they	who's
at	-	is	ourselves	they'd	whom
-	**had**	isn't	out	they'll	why
be	hadn't	it	over	they're	why's
because	has	it's	own	they've	with
been	hasn't	its	-	this	won't
before	have	itself	**r**	those	would
being	haven't	-	-	through	wouldn't
below	having	**just**	**s**	to	www
between	he	-	same	too	-
both	he'd	**k**	shall	-	**you**
but	he'll	-	shan't	**under**	you'd
By	he's	**let's**	she	until	you'll
-	hence	like	she'd	up	you're
can	her	-	she'll	-	you've
can't	here	**me**	she's	**very**	your
cannot	here's	mor	should	-	yours
co	hers	most	shouldn't	**was**	yourself
com	herself	mustn't	since	wasn't	yourselves
could	him	my	so	we	
couldn’t	himself	myself	some	we'd	
			such		

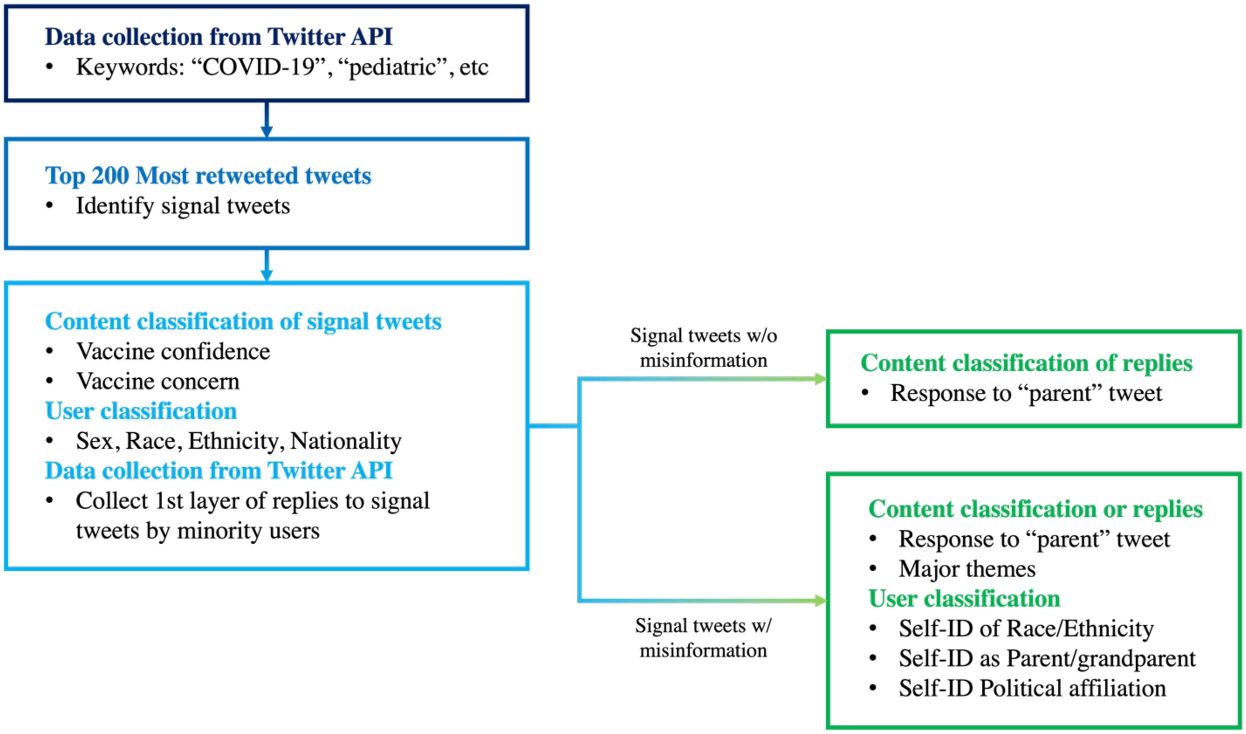

Figure 2: Brief flowchart overview of data collection and qualitative analysis. This study specifically focuses on tweets with explicit misinformation about the COVID-19 pediatric vaccine authored by users of ethnic or racial minority descent and the responses to said misinformation.

RESULTS

Literature Review

<u>COVID-19 Twitter Sentiments by or Related to Racial or Ethnic Minority Users</u>

A review of the literature found a dearth of studies focused specifically on the sentiments made by or related to communities of color on Twitter on the topics of COVID-19 or the COVID-19 vaccine. While there were a large number COVID-19 Twitter sentiment analysis studies found, there were few that examined the tweets made by or related to a racial minority community. Odlum et. al., applied topic modeling to COVID-19 related tweets that were specific to the African American community to identify common topics of discussion pertaining to the lived-experience of African Americans during the pandemic. This was done by identifying tweets that included "publicly open African American Twitter community." These included "#blacktwitter," "staywoke," and "#blacklivesmatter" (Odlum et al., 2020). The authors found sentiments promoting strength, positivity, and cohesion within the community through identified n-grams such as "Black strong," "support black businesses," and "growing up black" (Odlum et al., 2020).

Similar methods have been employed in previous studies to focus the content of tweets specifically from minority online communities or issues that related to minority communities. For example, Cao et. al. examined tweets that included the hashtag "#StopAsianHate" in the two weeks following the signing of the COVID-19 Hate Crimes Act by President Joe Biden. The authors' thematic analysis of 902 eligible tweets revealed themes that discussed the history and racism behind anti-Asian racism as well as how to get involved in the #StopAsianHate movement, "[a]ppreciate the AAPI community's culture, history, and contributions, and help increase visibility on AAPI issues" (Cao et al., 2022).

Themes relating to race, racism, and minority communities were found in other studies that examined tweets at the intersection of topics of COVID-19 vaccines, race, and ethnicity. Criss et. al. found that while racial and ethnic terms were used as descriptors in tweets, some tweets used such terms in a derogatory way (Criss et al., 2021). Among the race and ethnicity-related themes identified were those that encompassed discussions and topics of needing a vaccine to combat racism in addition to COVID-19, conspiracy theories surrounding race extermination, health disparities in communities of color, and vaccine distribution inequity (Criss et al., 2021). The authors also found racially offensive pro-vaccine jokes and humor (Criss et al., 2021). Calac et. al., examined the vaccine-related misinformation in tweets that were regarding the death of Hank Aaron, a notable Black baseball hall of fame player and public figure (Calac et al., 2022). From the sampled tweets using the keywords “Hank Aaron” and “vaccine,” the authors found misinformation in over half of the results. From the direct replies to the misinformation-labeled tweets, the authors found that over three-fourths agreed or had a positive sentiment towards the misinformation presented (Calac et al., 2022). Among the common misinformation themes identified was one regarding “claims that federal officials were targeting Black Americans” (Calac et al., 2022).

Coding Schemes for Tweets and Online Content Relating to the COVID-19 Vaccine

Studies examining online content related to the COVID-19 vaccine produced a variety of qualitative codes for data. Hughes et. al. developed a codebook for anti-vaccine media that captures information about both the content, which the authors call “narrative tropes,” and the dissemination, which the authors call “rhetorical strategies” (Hughes et al., 2021). In an examination of tweets mentioned in fact-check claims, Shahi et al. normalized the verdicts from fact-checking organizations (e.g., Snopes) using a one to four scale (i.e., “1=‘False’, 2=‘Partially

False’, 3=‘True’, 4=‘Others’”) (Shahi et al., 2021). From here, the authors examined tweets in the first two categories. Many studies developed thematic codes associated with a pro- or anti-vaccine status. For example, Criss et. al., identified sub-themes within “vaccine support” and “vaccine opposition” (Criss et al., 2021). The authors also identified themes relating to misinformation, equity, representation, and politics. Another approach involved the WHO’s Strategic Advisory Group of Experts (SAGE) Vaccine Hesitancy Matrix. Calac et al. utilized this in their coding for factors influential to vaccine hesitancy and confidence (Calac et al., 2022). This framework separates factors into three categories. They are “contextual”, “individual and group”, and “vaccine/vaccination-specific” and has been commonly used to guide coding in studies involving vaccine hesitancy and confidence (Calac et al., 2022; de Figueiredo et al., 2020).

Overall and Non-Misinformation-Specific Results

A total of 863,007 tweets were collected from the public streaming Twitter API. Of these, 233,612 were related to the top 200 most retweeted tweets, resulting in 27 percent of the total dataset being related to the top 200 most retweeted tweets. See **Figure 3** for the calculation of this percentage.

$$\frac{(233{,}612\ \textit{tweets related to the "top}\ 200\ \textit{most RT tweets"}\ +\ \text{top}\ 200\ \textit{most RT tweets}}{863{,}007\ \textit{total tweets in dataset}} = 27\%$$

Figure 3: Percentage of dataset related to top 200 most retweeted tweets calculation.

Of the top 200 most retweeted tweets, 38% (n = 76) were identified as signal tweets, tweets that contained discussion of or were related to the COVID-19 vaccine topics for pediatrics and children. These 76 were authored by 46 Twitter-verified users. Nearly one-third (n = 24) of these tweets were made by Twitter-verified users of ethnic or racial minority descent. These 24 "parent" tweets had a total of 6,629 first layer replies, with a max of 1,635, a min of 30, and a mean of 271 first layer replies across the 24 parent tweets.

Misinformation-Specific Results

Four "parent" tweets were identified as having explicit misinformation, which was the focus of this study. **Table 2** displays information about the tweets, the number of first layer replies, and information about verified users who authored the misinformation "parent" tweets. For brevity the misinformation "parent" tweets will be referred to by the number in the "Tweet" column (e.g., the first misinformation "parent" tweet will be referred to as "Tweet 1". Three main themes were identified across the four tweets. These are "vaccine development conspiracy" for Tweet 1, "vaccine is experimental", for Tweet 2, and "vaccine is a control tactic" for Tweets 3 and 4.

A total of 314 first layer replies were collected from the four misinformation "parent" tweets. Of these, approximately half (n = 156) were related to the main theme of the misinformation "parent" tweet. These will be referred to as "signal replies". The rest, 158, were unrelated to the main theme of the misinformation "parent" tweet. Unrelated replies include things such as spam, personal attacks on the author without discussion of the misinformation presented, and discussion of unrelated political or current events. The number of signal replies varied from a low of 20 signal replies to a high of 75 signal replies. The percent of signal replies to each misinformation "parent" tweet varied from a low of 21 percent to a high of 67 percent, with a mean of 50 percent.

Tweets 1 and 2 were authored by verified users of the American nationality who are both racially African American and/or Black. Tweets 3 and 4 were authored by verified users of British nationality who are racially Black and Asian, respectively. The audience users who authored the misinformation "parent" tweet varied from a low of 89 thousand followers to a high of 980 thousand followers. The number of total tweets authored by the user also varied with a low of 9 thousand to a high of 71 thousand tweets.

Table 4: Four identified misinformation "parent" tweet authored by Twitter-verified users from minority communities. The table lists the main theme of each tweet, number of replies, and users' information.

Tweet Information				**Users' Information**			
Tweet	Main Theme	Reply Count	Signal Reply Count	Race	Nationality	Follower Count	Tweet Count
1	Vaccine development conspiracy	65	39 (60%)	Black	US	107k	13k
2	Vaccine is experimental	33	22 (67%)	Black	US	980k	71k
3	Vaccine is a control tactic	96	20 (21%)	Black	UK	89k	9k
4	Vaccine is a control tactic	120	75 (63%)	Asian	UK	495k	13k
Total Replies	-	314	156 (50%)	-	-	-	-

The sentiment of the signal replies in relation to the misinformation “parent” tweet was analyzed and revealed varying counts of agree, disagree, and undefined responses for each to the misinformation “parent” tweet. Overall, 72 percent (n = 112) of the signal replies agreed, 23 percent (n = 36) disagreed, and 5 percent (n = 8) had an undefined response to the misinformation “parent” tweet. Tweet 1 and 4 generated a higher percentage of user replies that agreed with the original sentiment, 77 percent (n = 30) and 84 percent (n = 63), respectively, compared to Tweets 2 and 3, 41 percent (n = 9) and 50 percent (n = 10), respectively. **Table 5** displays these data.

Table 5: Sentiment analysis of signal replies to the four misinformation “parent” tweets. Counts and percentages of agree, disagree, and undefined sentiments for each misinformation “parent” tweet (individually) and for all four misinformation “parent” tweets (in total) are presented.

Number of Signal Replies		**Signal Replies’ Response Sentiment**		
Tweet	Total	Agree	Disagree	Undefined
1	**100% (39)**	77% (30)	21% (8)	3% (1)
2	**100% (22)**	41% (9)	45% (10)	14% (3)
3	**100% (20)**	50% (10)	45% (9)	5% (1)
4	**100% (75)**	84% (63)	12% (9)	4 (3)
Total Replies	**100% (156)**	72% (112)	23% (36)	5% (8)

Common themes were identified in signal replies that agreed with the misinformation “parent” tweet. These were, in order from greatest to least, concerns or assertions that the vaccine has or will cause harm to children (n = 51), calls to resist pressures of receiving the vaccine from government or public health officials (n = 29), and conspiracy theories about vaccine manufacturers, public health officials, and the government (n = 21).

The 156 signal replies were authored by 152 unique users. Of these 86% (n = 130) had written something in the optional public bio section of their Twitter user account (i.e., user metadata), while 15% (n = 22) opted to leave that section blank. Analysis of these Twitter user bios revealed that nine percent (n = 13) self-identified as being a parent or grandparent and that nearly half (46 percent, n = 70) self-identified with some type of political leaning. Self-reported ethnic and racial identification was not found. **Table 6** displays these data.

.

Table 6: Unique users' self-reported identification of parental/grandparental status and political leaning in Twitter user bio section. Counts and percentages for each misinformation "parent" tweet (individually) and for all four misinformation "parent" tweets (in total) are presented.

Number of Unique Users		**Parent/Grandparent**		**Political Leaning**	
Tweet	Total	Yes	No	Yes	No
1	**100% (37)**	22% (8)	78% (29)	59% (22)	41% (15)
2	**100% (22)**	0% (0)	100% (22)	36% (8)	64% (14)
3	**100% (20)**	0% (0)	100% (20)	30% (6)	70% (14)
4	**100% (73)**	7% (5)	93% (68)	47% (34)	53% (39)
Total Unique Users	**100% (152)**	9% (13)	91% (139)	46% (70)	54% (82)

Analysis of the text in user bios resulted revealed 871 unique strings, not including the strings included in the stop word list. Eighty percent (n = 699) of these strings had a frequency count of one, 12 percent (n = 107) had a frequency count of two, three percent (n = 25) had a frequency count of 3, two percent (n = 21) had a frequency count of 4, and one percent (n = 12) had a frequency count of five. Strings with frequency counts of six, seven, eight, and ten, in combination, comprised of one percent of the total unique strings (n = 7). These seven terms, in order from most frequent to least frequent, were “love”, “free”, “will”, “lover”, “conservative”, “music”, and “control”. Their frequency counts are displayed in **(Figure 4)**. The resulting relative-scaled word cloud (i.e., a word twice as common will appear twice as large) also highlights these terms **(Figure 5)**.

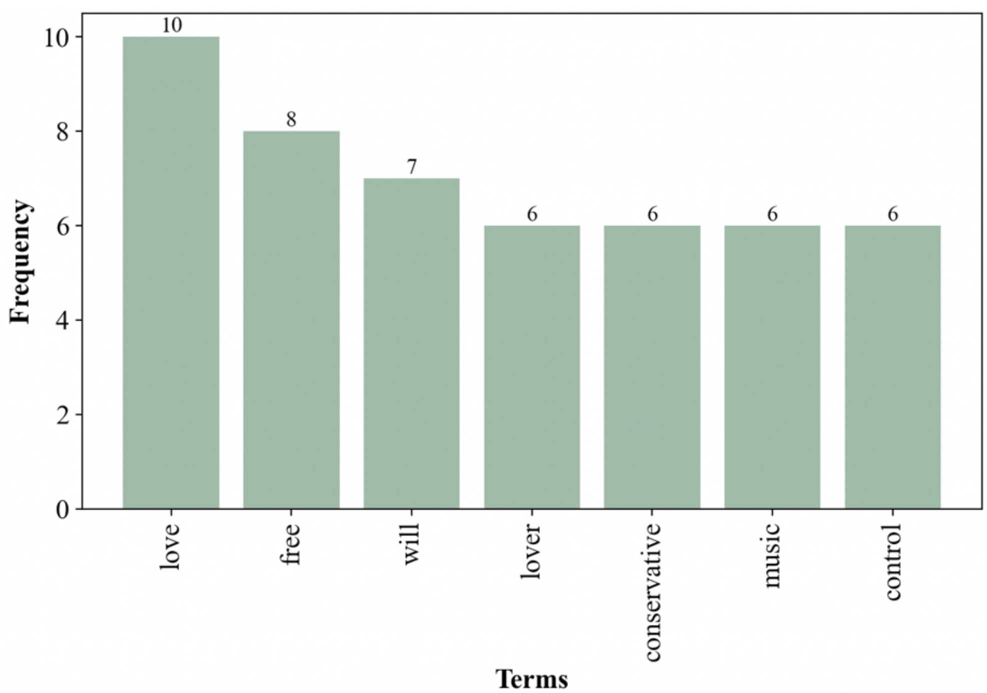

Figure 4: Top five most common single-string terms present in Twitter user bio of users who authored signal replies to misinformation “parent” tweet by frequency.

Figure 5: Relative-scaled word cloud visualization of single-string text from Twitter user bios of users who authored signal replies to misinformation "parent" tweet. Words that appear larger in the word cloud appeared more frequently in bios.

Self-reported political affiliation was examined in relation to the sentiment of the replies to the misinformation "parent" tweet. Of those who self-reported their political identification in Twitter user bio section of their user account, the majority of signal replies were authored by those who self-identified as right leaning (n = 57). This was followed by 11 that were authored by those who self-identified as left leaning and two that were authored by those with a political affiliation that is undefined and not strictly liberal or conservative. Of those who self-identified as left leaning, 27 percent agreed, while 73 percent disagreed. Of those who self-identified as right leaning, 86 percent agreed, while 9 percent disagreed. **Table 7** displays these data. Proportionality of these values were visualized in a simple stacked bar graph **(Figure 6)** and a 100% stacked bar chart **(Figure 7)**.

Table 7: Self-reported political identification in Twitter user bio and sentiment to misinformation "parent" tweet.

Users' Twitter User Bio		**Sentiment to "Parent" Tweet**		
Self-Reported Political ID	Total	Agree	Disagree	Undefined
Left Leaning	**100% (11)**	27% (3)	73% (8)	0% (0)
Right Leaning	**100% (57)**	86% (49)	9% (5)	5% (3)
Undefined Political Leaning	**100% (2)**	50% (1)	50% (1)	0% (0)
No Political Leaning	**100% (86)**	68% (59)	26% (22)	6% (5)
Total Replies	**100% (156)**	72% (112)	23% (36)	5% (8)

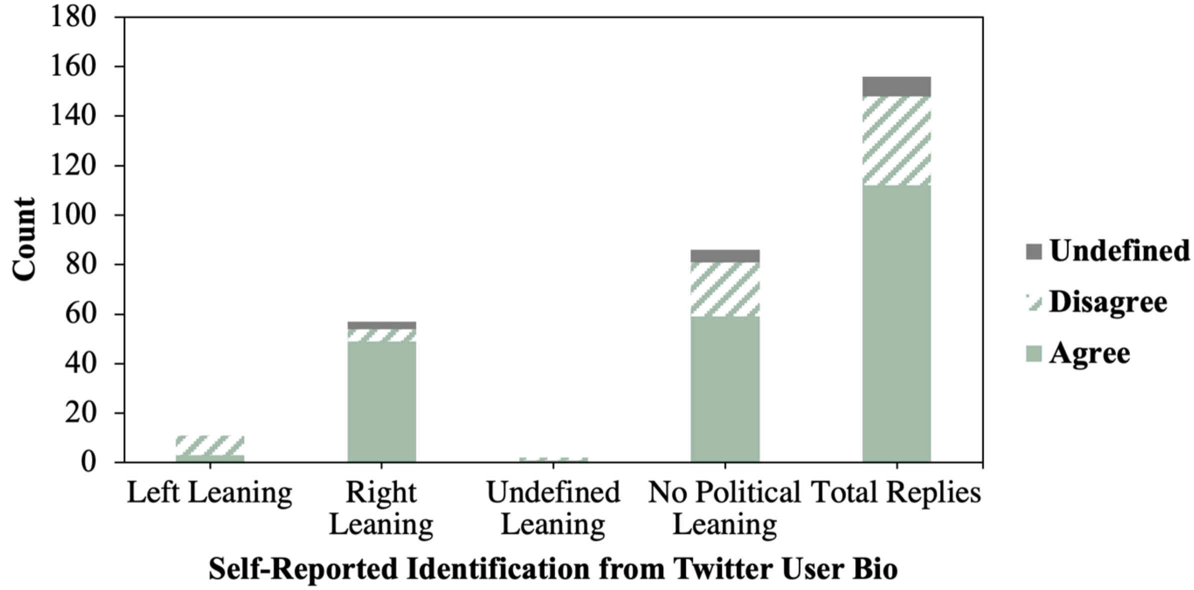

Figure 6: Stacked bar graph of self-reported political identification from Twitter user bio and response to misinformation "parent" tweet bar.

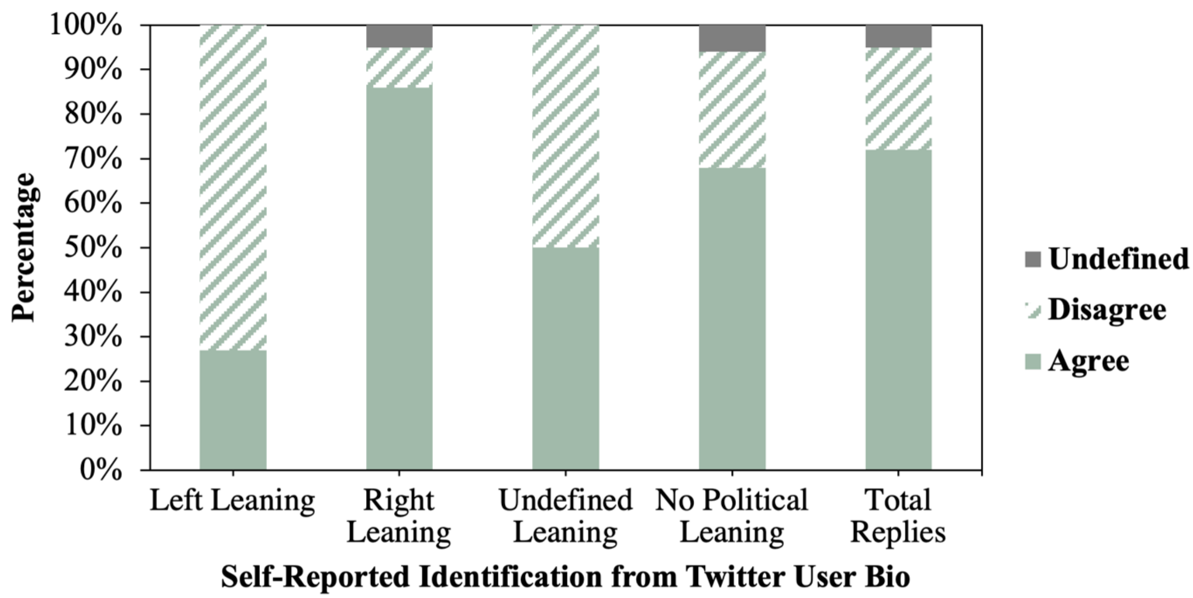

Figure 7: 100% Stacked bar graph of self-reported political identification from Twitter user bio and response to misinformation "parent" tweet bar.

DISCUSSION

Literature Review

The scarcity of COVID-19 Twitter sentiment analysis studies with a specific focus on tweets made by or related-to a racial minority online community of users demonstrates a need for further research that examines online sentiments and its impact both online and offline among minority and groups disproportionately impacted by COVID-19. Though vaccine hesitation and misinformation has been shown to be present in communities of color, there is limited knowledge as to how these communities react and contribute to these narratives themselves online. While it is possible to argue that Twitter is not a platform that facilitates or contains robust discussions by racial and ethnic minority users, a look through hashtags such as #BlackGirlMagic, #VeryAsian, and #ImmiGrad, demonstrate quite the opposite. Studies have previously found nuanced discussions by minority communities on topics such as social inequity and systemic racism on Twitter (Farina et al., 2021; Zakaria et al., 2021). With that being said, it must be acknowledged that there has yet to be an established standard for confirming race or ethnicity from the user profiles of Twitter users. Golder et. al.'s scoping review of the literature on this topic found that a wide range of methodologies (e.g., manual coding, census data linkage, language recognition, machine learning, NLP, etc) using a variety of data (e.g., "names, pictures, information from bios" and "location or content of the tweets") were employed to determine race and ethnicity identification (Golder et al., 2022). The authors also found that self-evaluation of these methods within studies produced a broad range of accuracy, with a low of 45 percent and a high of 93 percent (Golder et al., 2022). Still, as social networking platforms like Twitter continue to grow as channels of information seeking and sharing, it is important to understand how the dissemination of information occurs within marginalized communities and by members of marginalized communities, especially if said information is deemed to be harmful.

Misinformation on Twitter

Twitter has been identified as a major channel of COVID-19 misinformation (Shahi et al., 2021). The Bruno Kessler Foundation reported that in the month of March of 2020 over 40 thousand tweets were posted to Twitter that linked to misinformation regarding the COVID-19 pandemic (*COVID-19 and Fake News in the Social Media*, 2020). The goal of this study was to characterize misinformation authored by Twitter-verified racial minority users and characterize the responses to the misinformation shared.

Of the top 200 most retweeted tweets sampled from the initial corpus of 863,007 tweets, four tweets were identified to (1) be authored by a Twitter-verified user of a racial minority group and (2) contained explicit misinformation. The four Twitter-verified users who authored these tweets had audiences that ranged from a low of 89 thousand followers to a high of 980 thousand followers. Meanwhile, the number of total tweets authored by the user also varied with a low of 9 thousand to a high of 71 thousand tweets. According to data from the Pew Research Center, this likely puts these users into the top ten percent of users on Twitter (Wojcik & Hughes, 2019). This is an important consideration because a larger base of followers may contribute to a wider degree of online influence. This is because the tweets of a user with a larger follower count are shown an disseminated to a much broader group of users compared to the average Twitter user. Additionally, it is estimated that the top ten percent of users on Twitter produce over 80 percent of the content (Wojcik & Hughes, 2019). Thus, a small number of users, those that are in the top ten percent, are more influential in that they have a larger user audience and control more of the content available to all registered users on Twitter as a whole.

Three themes relating to misinformation were identified in these four tweets. These are "vaccine development conspiracy" for Tweet 1, "vaccine is experimental", for Tweet 2, and "vaccine is a control tactic" for Tweets 3 and 4. Interestingly the two tweets authored by the two Twitter-verified users from the UK were both coded as "vaccine is a control tactic". Both tweets implied that approval of the COVID-19 vaccine for children was a part of a larger scheme of control by the government.

Self-Reported Identifications

The Twitter user bio is an unstructured area of the account that presents an optional opportunity for users to display a public summary about themselves. Commonly displayed in user bios are occupations, fields of study, businesses, and other affiliations. This section is located directly under the account username and profile picture. These summaries are meant to be brief and are limited to 160 characters including spaces. Due to this forced brevity, the user must choose what to include or exclude from the bio section of their account. Thus, information that is included may be important the user's self-identity and self-presentation on Twitter.

Nearly half of the users who authored an on-topic reply to the misinformation presented in the "parent" tweet had some type of explicit political affiliation or ideological identification self-reported in the bio section of their account. The majority of these were right-leaning or conservative-leaning. This is inconsistent with the overall ideological demographic of Twitter for both the US and UK. For both nations, it was been found that users on Twitter tend to be more ideologically left-leaning (Sloan et al., 2015; Wojcik & Hughes, 2019). However, this finding is consistent with similar studies focused on Twitter-based COVID-19 misinformation. Muric et al. found that conservative leaning accounts were more likely to interact with misinformation on Twitter (Muric et al., 2021). For users who identified as right-leaning, an overwhelming majority agreed with the misinformation-labeled content presented in the “parent” tweet. Although it was to a lesser degree, the reverse was true for users who identified as left-leaning. Users who did not self-report an explicit political ideology accounted for over half of the responses. Overall, the proportions of agree, disagree, and undefined sentiments in response to the misinformation-labeled “parent” tweet were roughly equivalent between users who did and did not self-report political ideology.

Though the majority of users with self-reported political ideology were right leaning, it is possible that these results were simply representative of the follower bases of the four Twitter-verified users. This is due to the nature of Twitter because users are shown the tweets of those whom they follow and users are likely to follow those whose tweets they engage with. As Twitter-verified users, the authors of the “parent” misinformation tweets are prominent public figures in their field. While this study did not seek to characterize the profiles and public images of these four users in depth, they include an American congressional candidate, a former American gubernatorial candidate, a British activist, and a British politician. All four are involved in politics in some manner. This may explain why self-reporting of political ideology was common in the user bios of those who replied. This may also explain the sizable presence of replies that disagreed with the original misinformation-labeled content presented as political public figures are among those who are followed both by those who agree and disagree their ideological ideas and platforms.

In contrast with the self-reporting of political ideology, self-reporting of users' status as a parent or grandparent were less common. Just under ten percent of users who authored an on-topic reply to the misinformation presented in the "parent" tweet had some type of explicit self-identification as a parent or grandparent. The word cloud visualization shows that the most common parental identifier was "mom". Indeed, the overwhelming majority of self-reported parental status were those who reported being mothers. This is inconsistent with data showing that, among parents who use a variety of social media, there is no significant difference between mothers and fathers (Duggan et al., 2015). However, the nature of the study's specific focus on the pediatric COVID-19 vaccine, may help explain why the findings are as such. As of 2020, mothers are still the primary caretaker for children, even in households with two working parents (Guy & Arthur, 2020).

While the four Twitter-verified users who authored the misinformation-labeled parent tweet were racial minorities, racial and ethnic identification was not available for users who replied to those tweets. Self-reporting of racial and ethnic identities was not present in user bios of those who authored on-topic replies. As mistrust in institutions and organizations has been a key theme identified in vaccine hesitant minority communities, it may play a role in whether a minority user includes such information in the bio section of their Twitter account. Overall, reliance on self-reported identifications in Twitter user bios limits the study's ability to determine the true proportions of users' identities. Further, the use of a word cloud to process present in user bios and visualize the text as most common single strings does not capture context.

Themes Among Replies to Misinformation

Among the replies that agreed with the vaccine misinformation presented, the most common themes identified include (1) concerns or assertions that the vaccine has or will cause harm to children, (2) calls for the resistance to pressures to receive the vaccine from government or public health officials, and (3) conspiracy theories about vaccine manufacturers, public health officials, and the government. The themes identified among the reply tweets were consistent with themes identified in previous studies that surveyed parents' attitudes and intentions towards the pediatric COVID-19 vaccine (C. B. Fisher et al., 2021; Ruggiero et al., 2021; Szilagyi et al., 2021). There was limited discourse on vaccine efficacy among those who agreed with the misinformation-labeled "parent" tweet. However, efficacy was a more common theme in replies that disagreed. This may indicate that efficacy is not a top-of-mind concern for those who have concerns about the COVID-19 pediatric vaccine as identified in this study. Additionally, one common myth regarding COVID-19 in children that has been identified is that children cannot contract COVID-19 and/or are not harmed by COVID-19 (C. B. Fisher et al., 2021). Thus, the efficacy of the COVID-19 pediatric vaccine may not be relevant to those who subscribe to this belief. However, for those who do believe that children can contract and/or can be harmed by COVID-19, the efficacy of the pediatric COVID-19 vaccine may be a more relevant point to address in targeted health promotion activity.

Limitations

As mentioned, and discussed, the major limitations of this study were the lack of a standard to extract user ethnic and racial identities from Twitter accounts and the reliance on information self-reported in user bios. This limits the study's ability to accurately estimate the true proportions of ethnic and racial minorities, parents, grandparents, and users' political stance within the replies to the four "parent" misinformation-labeled tweets. Other limitations include the sample size, the lack of a control group, and the choice to use "misinformation" as an umbrella term that encompasses all false information regardless of intent and impact.

Though a corpus of 863,007 tweets were collected from the Twitter API, only four tweets within the top 200 most retweet tweets were identified to have explicit false information about the pediatric COVID-19 vaccine and were authored by Twitter-verified users who were from ethnic and/or minority descent. The small sample of size of four users further limits the generalizability of these findings, which makes it difficult to determine if the findings are representative of all misinformation produced and propagated through tweets authored by Twitter-verified users of minority descent. Further, the lack of a control group to use in comparison to the findings limits the study's ability to comprehend if and how the misinformation and the response to misinformation presented by these users is different from the misinformation and the response to misinformation presented overall. For example, it is unclear how the tweets and responses in this sample compare to the tweets and responses authored by Twitter-verified White users, which warrants further study.

Finally, though false information can be classified into the three terms of "misinformation," "disinformation," and "mal-information," based on intent, this study did not differentiate between the three categories, opting to call all false information, regardless of intent, "misinformation." This lack of differentiation limits the study's ability to identify potential differences in the rhetoric and impact associated with the nuances of each category of false information.

CONCLUSION

This study examined Twitter-based misinformation regarding the COVID-19 pediatric vaccine that was disseminated by verified users who were from minority communities. The themes identified in misinformation presented in these tweets were consistent with those found in the literature. Analysis of the user responses to the misinformation revealed concerns for safety as the top priority among those who agreed with the sentiment of the original tweet. Characterizing online sentiments regarding the COVID-19 pediatric vaccine and the related misinformation, as was done in this study, may help inform current and future public health interventions and health communication and promotion activities that address vaccine hesitancy. Currently, there little research into the online sentiments of minority communities on this topic. Further research is needed to examine the specific content and rhetoric among misinformation that is created and shared minority communities. Research is also needed to examine how the misinformation-related sentiments within minority communities' online discussions compare to the misinformation-related sentiments found in online discussions as a whole. Targeted health promotion strategies towards vaccine hesitant communities and communities disproportionately affected by infectious, preventable disease, such as racial and ethnic minority communities in the case of COVID-19, would greatly benefit from a better understanding of the dissemination of misinformation that occurs within these communities online.

REFERENCES

About Verified Accounts. (n.d.). Twitter Help Center. Retrieved May 22, 2022, from About Verified Accounts

About your Home timeline on Twitter. (n.d.). Twitter Help Center. Retrieved May 19, 2022, from https://help.twitter.com/en/using-twitter/twitter-timeline

Belluz, J. (2015, September 16). Donald Trump believes vaccines cause autism. Here's the evidence that proves him wrong. *VOX*. https://www.vox.com/2015/9/16/9342825/donald-trump-vaccines-autism

Bin Naeem, S., & Kamel Boulos, M. N. (2021). COVID-19 Misinformation Online and Health Literacy: A Brief Overview. *International Journal of Environmental Research and Public Health*, *18*(15), 8091. https://doi.org/10.3390/ijerph18158091

Bonnevie, E., Gallegos-Jeffrey, A., Goldbarg, J., Byrd, B., & Smyser, J. (2021). Quantifying the rise of vaccine opposition on Twitter during the COVID-19 pandemic. *Journal of Communication in Healthcare*, *14*(1), 12–19. https://doi.org/10.1080/17538068.2020.1858222

Cai, M., Shah, N., Li, J., Chen, W.-H., Cuomo, R. E., Obradovich, N., & Mackey, T. K. (2020). Identification and characterization of tweets related to the 2015 Indiana HIV outbreak: A retrospective infoveillance study. *PLOS ONE*, *15*(8), e0235150. https://doi.org/10.1371/journal.pone.0235150

Calac, A. J., Haupt, M. R., Li, Z., & Mackey, T. (2022). Spread of COVID-19 Vaccine Misinformation in the Ninth Inning: Retrospective Observational Infodemic Study. *JMIR Infodemiology*, *2*(1), e33587. https://doi.org/10.2196/33587

Cao, J., Lee, C., Sun, W., & De Gagne, J. C. (2022). The #StopAsianHate Movement on Twitter: A Qualitative Descriptive Study. *International Journal of Environmental Research and Public Health*, *19*(7), 3757. https://doi.org/10.3390/ijerph19073757

Carson, S. L., Casillas, A., Castellon-Lopez, Y., Mansfield, L. N., Morris, D., Barron, J., Ntekume, E., Landovitz, R., Vassar, S. D., Norris, K. C., Dubinett, S. M., Garrison, N. A., & Brown, A. F. (2021). COVID-19 Vaccine Decision-making Factors in Racial and Ethnic Minority Communities in Los Angeles, California. *JAMA Network Open*, *4*(9), e2127582. https://doi.org/10.1001/jamanetworkopen.2021.27582

CDC. (2021, November 3). *How to Address COVID-19 Vaccine Misinformation*. Centers for Disease Control and Prevention. https://www.cdc.gov/vaccines/covid-19/health-departments/addressing-vaccine-misinformation.html

Chew, C., & Eysenbach, G. (2010). Pandemics in the Age of Twitter: Content Analysis of Tweets during the 2009 H1N1 Outbreak. *PLoS ONE*, *5*(11), e14118. https://doi.org/10.1371/journal.pone.0014118

Coustasse, A., Kimble, C., & Maxik, K. (2021). COVID-19 and Vaccine Hesitancy: A Challenge the United States Must Overcome. *Journal of Ambulatory Care Management*, *44*(1), 71–75. https://doi.org/10.1097/JAC.0000000000000360

COVID-19 and Fake News in the Social Media. (2020, March 10). Bruno Kessler Foundation. https://www.fbk. eu/en/press-releases/covid-19-and-fake-news-in-the-social-media/

Covid-19 news and information: Consumption and attitude. (2020). Ofcom.

Criss, S., Nguyen, T. T., Norton, S., Virani, I., Titherington, E., Tillmanns, E. L., Kinnane, C., Maiolo, G., Kirby, A. B., & Gee, G. C. (2021). Advocacy, Hesitancy, and Equity: Exploring U.S. Race-Related Discussions of the COVID-19 Vaccine on Twitter.

International Journal of Environmental Research and Public Health, *18*(11), 5693. https://doi.org/10.3390/ijerph18115693

Cuan-Baltazar, J. Y., Muñoz-Perez, M. J., Robledo-Vega, C., Pérez-Zepeda, M. F., & Soto-Vega, E. (2020). Misinformation of COVID-19 on the Internet: Infodemiology Study. *JMIR Public Health and Surveillance*, *6*(2), e18444. https://doi.org/10.2196/18444

de Figueiredo, A., Simas, C., Karafillakis, E., Paterson, P., & Larson, H. J. (2020). Mapping global trends in vaccine confidence and investigating barriers to vaccine uptake: A large-scale retrospective temporal modelling study. *The Lancet*, *396*(10255), 898–908. https://doi.org/10.1016/S0140-6736(20)31558-0

Dubey, A. D. (2020). Twitter Sentiment Analysis during COVID19 Outbreak. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3572023

Duggan, M., Lenhart, A., Lampe, C., & Ellison, N. B. (2015, 16). *Parents and Social Media*. Pew Research Center. https://www.pewresearch.org/internet/2015/07/16/parents-and-social-media/

Farina, A. S. J., Klumpner, S., Alvarez, A. R. G., Azhar, S., & Nguyen, C. M. (2021). Experiences of racist encounters among Asian Americans: Analysis of #thisis2016. *Journal of Ethnic & Cultural Diversity in Social Work*, 1–12. https://doi.org/10.1080/15313204.2021.1984356

Farooq, F., & Rathore, F. A. (2021). COVID-19 Vaccination and the Challenge of Infodemic and Disinformation. *Journal of Korean Medical Science*, *36*(10), e78. https://doi.org/10.3346/jkms.2021.36.e78

Filippou, I., Gozluklu, A. E., T. Nguyen, M., & Viswanath-Natraj, G. (2020). The Information Content of Trump Tweets and the Currency Market. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3754991

Fisher, C. B., Gray, A., & Sheck, I. (2021). COVID-19 Pediatric Vaccine Hesitancy among Racially Diverse Parents in the United States. *Vaccines*, *10*(1), 31. https://doi.org/10.3390/vaccines10010031

Fisher, K. A., Bloomstone, S. J., Walder, J., Crawford, S., Fouayzi, H., & Mazor, K. M. (2020). Attitudes Toward a Potential SARS-CoV-2 Vaccine: A Survey of U.S. Adults. *Annals of Internal Medicine*, *173*(12), 964–973. https://doi.org/10.7326/M20-3569

Golder, S., Stevens, R., O'Connor, K., James, R., & Gonzalez-Hernandez, G. (2022). Methods to Establish Race or Ethnicity of Twitter Users: Scoping Review. *Journal of Medical Internet Research*, *24*(4), e35788. https://doi.org/10.2196/35788

Guy, B., & Arthur, B. (2020). Academic motherhood during COVID-19: Navigating our dual roles as educators and mothers. *Gender, Work & Organization*, *27*(5), 887–899. https://doi.org/10.1111/gwao.12493

Howe, N., Giles, E., Newbury-Birch, D., & McColl, E. (2018). Systematic review of participants' attitudes towards data sharing: A thematic synthesis. *Journal of Health Services Research & Policy*, *23*(2), 123–133. https://doi.org/10.1177/1355819617751555

Hughes, B., Miller-Idriss, C., Piltch-Loeb, R., Goldberg, B., White, K., Criezis, M., & Savoia, E. (2021). Development of a Codebook of Online Anti-Vaccination Rhetoric to Manage COVID-19 Vaccine Misinformation. *International Journal of Environmental Research and Public Health*, *18*(14), 7556. https://doi.org/10.3390/ijerph18147556

Hussain, A., Tahir, A., Hussain, Z., Sheikh, Z., Gogate, M., Dashtipour, K., Ali, A., & Sheikh, A. (2021). Artificial Intelligence–Enabled Analysis of Public Attitudes on Facebook and Twitter Toward COVID-19 Vaccines in the United Kingdom and the United States: Observational Study. *Journal of Medical Internet Research*, *23*(4), e26627. https://doi.org/10.2196/26627

Infodemic. (2022). World Health Organization. https://www.who.int/health-topics/infodemic

Jacobson, R. M., St. Sauver, J. L., & Finney Rutten, L. J. (2015). Vaccine Hesitancy. *Mayo Clinic Proceedings*, *90*(11), 1562–1568. https://doi.org/10.1016/j.mayocp.2015.09.006

Kennedy, J. (2020). Vaccine Hesitancy: A Growing Concern. *Pediatric Drugs*, *22*(2), 105–111. https://doi.org/10.1007/s40272-020-00385-4

Kwok, S. W. H., Vadde, S. K., & Wang, G. (2021). Tweet Topics and Sentiments Relating to COVID-19 Vaccination Among Australian Twitter Users: Machine Learning Analysis. *Journal of Medical Internet Research*, *23*(5), e26953. https://doi.org/10.2196/26953

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

Mackey, T. K., Purushothaman, V., Haupt, M., Nali, M. C., & Li, J. (2021). Application of unsupervised machine learning to identify and characterise hydroxychloroquine misinformation on Twitter. *The Lancet Digital Health*, *3*(2), e72–e75. https://doi.org/10.1016/S2589-7500(20)30318-6

Malik, A. A., McFadden, S. M., Elharake, J., & Omer, S. B. (2020). Determinants of COVID-19 vaccine acceptance in the US. *EClinicalMedicine*, *26*, 100495. https://doi.org/10.1016/j.eclinm.2020.100495

McKee, C., & Bohannon, K. (2016). Exploring the Reasons Behind Parental Refusal of Vaccines. *The Journal of Pediatric Pharmacology and Therapeutics*, *21*(2), 104–109. https://doi.org/10.5863/1551-6776-21.2.104

Mhamdi, C., Al-Emran, M., & Salloum, S. A. (2018). Text Mining and Analytics: A Case Study from News Channels Posts on Facebook. In K. Shaalan, A. E. Hassanien, & F. Tolba (Eds.), *Intelligent Natural Language Processing: Trends and Applications* (Vol. 740, pp. 399–415). Springer International Publishing. https://doi.org/10.1007/978-3-319-67056-0_19

Muric, G., Wu, Y., & Ferrara, E. (2021). COVID-19 Vaccine Hesitancy on Social Media: Building a Public Twitter Data Set of Antivaccine Content, Vaccine Misinformation, and Conspiracies. *JMIR Public Health and Surveillance*, *7*(11), e30642. https://doi.org/10.2196/30642

Naeem, S. B., & Bhatti, R. (2020). The Covid-19 'infodemic': A new front for information professionals. *Health Information & Libraries Journal*, *37*(3), 233–239. https://doi.org/10.1111/hir.12311

Odlum, M., Cho, H., Broadwell, P., Davis, N., Patrao, M., Schauer, D., Bales, M. E., Alcantara, C., & Yoon, S. (2020). *Application of Topic Modeling to Tweets to Learn Insights on the African American Lived Experience of COVID-19*. 6. https://doi.org/10.3233/SHTI200484

Osakwe, Z. T., Ikhapoh, I., Arora, B. K., & Bubu, O. M. (2021). Identifying public concerns and reactions during the COVID-19 pandemic on Twitter: A text-mining analysis. *Public Health Nursing*, *38*(2), 145–151. https://doi.org/10.1111/phn.12843

Razai, M. S., Osama, T., McKechnie, D. G. J., & Majeed, A. (2021). Covid-19 vaccine hesitancy among ethnic minority groups. *BMJ*, n513. https://doi.org/10.1136/bmj.n513

Ruggiero, K. M., Wong, J., Sweeney, C. F., Avola, A., Auger, A., Macaluso, M., & Reidy, P. (2021). Parents' Intentions to Vaccinate Their Children Against COVID-19. *Journal of Pediatric Health Care*, *35*(5), 509–517. https://doi.org/10.1016/j.pedhc.2021.04.005

Santos-d'Amorim, K., & Fernandes de Oliveira Miranda, M. (2021). Informação incorreta, desinformação e má informação: Esclarecendo definições e exemplos em tempos de desinfodemia. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência Da Informação*, *26*, 01–23. https://doi.org/10.5007/1518-2924.2021.e76900

Scannell, D., Desens, L., Guadagno, M., Tra, Y., Acker, E., Sheridan, K., Rosner, M., Mathieu, J., & Fulk, M. (2021). COVID-19 Vaccine Discourse on Twitter: A Content Analysis of Persuasion Techniques, Sentiment and Mis/Disinformation. *Journal of Health Communication*, *26*(7), 443–459. https://doi.org/10.1080/10810730.2021.1955050

Shahi, G. K., Dirkson, A., & Majchrzak, T. A. (2021). An exploratory study of COVID-19 misinformation on Twitter. *Online Social Networks and Media*, *22*, 100104. https://doi.org/10.1016/j.osnem.2020.100104

Sloan, L., Morgan, J., Burnap, P., & Williams, M. (2015). Who Tweets? Deriving the Demographic Characteristics of Age, Occupation and Social Class from Twitter User Meta-Data. *PLOS ONE*, *10*(3), e0115545. https://doi.org/10.1371/journal.pone.0115545

Smith, P. J., Humiston, S. G., Marcuse, E. K., Zhao, Z., Dorell, C. G., Howes, C., & Hibbs, B. (2011). Parental Delay or Refusal of Vaccine Doses, Childhood Vaccination Coverage at 24 Months of Age, and the Health Belief Model. *Public Health Reports*, *126*(2_suppl), 135–146. https://doi.org/10.1177/00333549111260S215

Szilagyi, P. G., Shah, M. D., Delgado, J. R., Thomas, K., Vizueta, N., Cui, Y., Vangala, S., Shetgiri, R., & Kapteyn, A. (2021). Parents' Intentions and Perceptions About COVID-19 Vaccination for Their Children: Results From a National Survey. *Pediatrics*, *148*(4), e2021052335. https://doi.org/10.1542/peds.2021-052335

Teherani, M., Banskota, S., Camacho-Gonzalez, A., Smith, A. G. C., Anderson, E. J., Kao, C. M., Crepy D'Orleans, C., Shane, A. L., Lu, A., & Jaggi, P. (2021). Intent to Vaccinate SARS-CoV-2 Infected Children in US Households: A Survey. *Vaccines*, *9*(9), 1049. https://doi.org/10.3390/vaccines9091049

The Lancet Infectious Diseases. (2020). The COVID-19 infodemic. *The Lancet Infectious Diseases*, *20*(8), 875. https://doi.org/10.1016/S1473-3099(20)30565-X

Wojcik, S., & Hughes, A. (2019). *Sizing Up Twitter Users* (p. 23). Pew Research Center. https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users/

Yan, X., Guo, J., Lan, Y., & Cheng, X. (2013). A biterm topic model for short texts. *Proceedings of the 22nd International Conference on World Wide Web - WWW '13*, 1445–1456. https://doi.org/10.1145/2488388.2488514

Yousefinaghani, S., Dara, R., Mubareka, S., Papadopoulos, A., & Sharif, S. (2021). An analysis of COVID-19 vaccine sentiments and opinions on Twitter. *International Journal of Infectious Diseases*, *108*, 256–262. https://doi.org/10.1016/j.ijid.2021.05.059

Zakaria, R., Herwanis, D., & Kinanti, S. (2021). Hashtag Black Lives Matter's Tweets as Education Media Messages. *PIONEER: Journal of Language and Literature*, *13*(2), 337. https://doi.org/10.36841/pioneer.v13i2.1325