MEMES MAGNETS AND MICROCHIPS

NARRATIVE DYNAMICS AROUND COVID-19 VACCINES
Memes, Magnets, and Microchips

Narrative dynamics around COVID-19 vaccines

The Virality Project
www.viralityproject.org

Stanford Internet Observatory
NCoC Algorithmic Transparency Institute
NYU Tandon School of Engineering
NYU Center for Social Media and Politics
UW Center for an Informed Public
Digital Forensic Research Lab
Graphika

2022
© 2022 The Virality Project

This report is made available under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License (international) https://creativecommons.org/licenses/by-nc-nd/4.0/

Version 1.0.2: Bibliography corrections & stylistic updates (August 2023)
Version 1.0.1: Updated contributors (April 2022)
Version 1.0.0: Release (February 2022)

Cover Illustration by Tuan Anh Nguyen

APA Style:

Available at https://purl.stanford.edu/mx395xj8490
Executive Editors

Renée DiResta
Elena Cryst
Lily Meyersohn

Contributors

Taylor Agajanian          Alyssa Kann
Osiris Cruz-Antonio      Kathy Liu
Elena Cryst              Matt Masterson
Renée DiResta             Erin McAweeney
Kaitlyn Dowling          Lily Meyersohn
Laura Edelson             Carly Miller
Kris Fortmann            Chase Small
Toni Friedman            Abigail Tarquino
Cameron Hickey           Julia Thompson
Zoe Huczok               Kyle Weiss
Lindsay Hurley

This report was edited by Eden Beck and designed by David Thiel.

Acknowledgements

This project was supported by a team of researchers and analysts who dedicated their time during the full monitoring period: Sam Bradshaw, Nicole Buckley, Jack Cable, Julienne Ching, Emma Dolan, Paul Duke, Isabella García-Camargo, Josh Goldstein, Jennifer John, Katie Jonsson, Connor Klentschy, Kolina Kolta, Kathy Liu, Pierce Lowary, Damon McCoy, Rachel Moran, Ashwin Ramaswami, Zeve Sanderson, Joey Schafer, Isaac Schaider, Divya Suresh Kumar, Alex Zaheer.

We would like to acknowledge feedback and support from the following colleagues:

Dan Bateyko
Michael Caulfield
Chris Giles
Kate Starbird
David Thiel
Contents

Executive Summary

1 Introduction
   1.1 Persistent Tactics, Actors, and Narratives
   1.1.1 Anti-Vaccine Themes, Tropes, and Narratives
   1.2 Informing the Scope of the Virality Project
      1.2.1 Detecting In-Scope Content
      1.2.2 External Stakeholder Engagement
   1.3 Outputs and Impact
   1.4 Reading this Report

2 Data Collection & Statistics
   2.1 Workflow and Analysis Methodology
      2.1.1 Monitoring, Intake, and Ticketing System
      2.1.2 Engagement Data Collection
   2.2 Engagement: An Imperfect Metric
   2.3 Engagement with Vaccine Misinformation Over Time
   2.4 Engagement by Category
   2.5 Engagement by Tactic
   2.6 Engagement Driven by Recurring Actors
   2.7 Engagement by Community
   2.8 Viral Incidents
   2.9 Conclusion

3 Narratives
   3.1 Narrative Categories
   3.2 Case Studies by Category
      3.2.1 Safety Narratives

1
7
10
13
14
15
17
18
20
27
27
27
30
31
32
32
34
34
36
36
40
42
43
45
45
3.2.2 Distribution Narratives ........................................... 52
3.2.3 Conspiracy Theories ............................................. 58
3.3 Conclusion .......................................................... 61

4  Actors & Tactics ......................................................... 74
4.1 A Cast of Recurring Actors ........................................ 74
   4.1.1 Long-Standing Anti-Vaccine Influencers ................. 75
   4.1.2 Wellness & Lifestyle Influencers .......................... 77
   4.1.3 Pseudomedical Influencers ............................... 79
   4.1.4 Conspiracy Theory Influencers .......................... 82
   4.1.5 Right-Leaning Political Influencers .................... 83
   4.1.6 Medical Freedom Influencers ........................... 86
   4.1.7 Defies Classification ..................................... 88
4.2 Spanish and Chinese Influencers ............................... 90
   4.2.1 Spanish: Doctors for Truth and Natalia Prego Cancelo ... 92
   4.2.2 Chinese: Guo Wengui and Yan Li-Meng ................ 94
4.3 Foreign Actors ..................................................... 96
   4.3.1 Russia .......................................................... 97
   4.3.2 China .......................................................... 100
   4.3.3 Iran ........................................................... 102
4.4 Conclusion .......................................................... 104

5  Platform Policies for Anti-Vaccine Content .................. 122
5.1 Platform Moderation .............................................. 122
   5.1.1 Pre-COVID-19 Vaccine-Related Policies ............... 123
   5.1.2 COVID-19-Related Policies .............................. 124
   5.1.3 Policy Enforcement: Remove, Reduce, Inform ........ 124
5.2 Nuances and Gaps in Vaccine-Related Policies .............. 128
   5.2.1 Personal Stories .......................................... 129
   5.2.2 Medical Freedom Claims ................................ 129
   5.2.3 Sensationalized or Decontextualized Content .......... 130
5.3 Authoritative Sources and the Deplatforming Debate ....... 132
   5.3.1 Elevating Authoritative Sources ....................... 132
   5.3.2 Deplatforming Harmful Sources ...................... 133
5.4 Conclusion .......................................................... 135

6  Recommendations for Stakeholders ............................. 139
6.1 Recommendations for Research Partners ..................... 139
6.2 Recommendations for Public Health Partners ................ 140
6.3 Recommendations for Government Partners .................. 141
6.4 Recommendations to Platforms ................................ 144
6.5 Conclusion .......................................................... 145
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendices</td>
<td>149</td>
</tr>
<tr>
<td>A Research Limitations</td>
<td>149</td>
</tr>
<tr>
<td>B Definitions</td>
<td>152</td>
</tr>
<tr>
<td>C Policy Timeline (Sources)</td>
<td>155</td>
</tr>
<tr>
<td>D Platform Policy Tables (Sources)</td>
<td>159</td>
</tr>
<tr>
<td>Bibliography</td>
<td>179</td>
</tr>
</tbody>
</table>
Executive Summary

Since the start of the global COVID-19 pandemic, global authority figures, institutions, and the public have struggled to share accurate health information. As early as February 2020, the World Health Organization declared an infodemic, a glut of news pouring in and harming peoples’ ability to find trustworthy sources. Mis- and disinformation spread widely. Social media companies, pressured and incentivized to curate the news, sought to surface accurate information about the pandemic, but faced a problem: What should they amplify in the absence of clear scientific consensus? As they pondered, anti-vaccine sentiments grew. Were the vaccines safe? Had our governments overstepped? Misinformation distorted answers to these questions. The global crisis created opportunities for governments, financially motivated actors, and conspiracy theorists to construct or amplify narratives that served their interests. The preexisting US anti-vaccine movement began to insert itself into the national conversation. As far back as January 2020, when talk of a COVID-19 vaccine was purely theoretical, anti-vaccine activists perceived an opportunity to undermine confidence in any and all vaccinations. As vaccinations rolled out in the US in late 2020, anti-vaccine activists became increasingly focused on undermining their uptake.

It was against this backdrop that the Virality Project (VP) came together. VP was formed from a collection of research institutions with experience identifying and understanding the spread of mis- and disinformation: the Stanford Internet Observatory, the University of Washington Center for an Informed Public, the Atlantic Council’s Digital Forensic Research Lab, Graphika, the National Conference on Citi-

---

1See Appendix B: Definitions for definitions of “misinformation” and “disinformation” as used in this report.
2See Appendix B: Definitions for a definition of “narrative” as used in this report.
3See Appendix B: Definitions for definitions of “anti-vaccine” and “anti-vaccine activist” as used in this report.
zenship (NCoC)’s Algorithmic Transparency Institute, and New York University’s Center for Social Media and Politics and Tandon School of Engineering.

Drawing on scholarship documenting the who, what, and how of the anti-vaccine movement, the Virality Project identified four categories of well-established narratives that were considered in-scope for VP research on social media platforms: (1) safety, (2) efficacy and necessity, (3) development and distribution, and (4) conspiracy theory. The VP team developed technology to identify emerging narratives, to understand what communities they appeared within, and to gauge their scope, speed, and spread. In addition, the analysts assessed social media platforms’ published policies to understand how (if at all) platforms might limit or action the spread of misleading vaccine-related content.

Key Takeaways

Anti-vaccine narratives related to the COVID-19 vaccine are not new.

- Long-standing anti-vaccine narratives were adapted and inflected for the COVID-19 vaccine.
- While these narratives traded in COVID-19 specifics, such as the new mRNA vaccines, the global infection rates, and the individual experiences of COVID-19 vaccine recipients, many hinged on repeated storylines that have long been popular within the anti-vaccine community.
- Anti-vaccine narratives within non-English-speaking communities on mainstream platforms in the US were largely translated from English-language content, though with some notable differences and specific influencers.

Vaccine mis- and disinformation was largely driven by a cast of recurring actors.

- Actors within the US include long-standing anti-vaccine influencers and activists, wellness and lifestyle influencers, pseudomedical influencers, conspiracy theory influencers, right-leaning political influencers, and medical freedom influencers.
- Foreign actors in China, Russia, and Iran took a full-spectrum propaganda approach, spanning both media and social media, to influence vaccine conversations in the US and around the world. However, these actors’ reach appeared to be far less than that of domestic actors.

\[iv\]See Appendix B: Definitions for a definition and explanation of our use in this report of the term “pseudomedical,” derived from the concept of “pseudoscience.”
Online engagement with vaccine mis- and disinformation is highly variable and shaped by social media platforms’ structures and policies.

- Engagement is driven by viral content that spreads outside its initial community of focus.
- Conspiracy theory incidents made up a relatively smaller number of incidents but appeared to be more engaging than incidents in other categories.
- While online platforms have made progress in creating and enforcing vaccine-related policies, gaps still exist.

Detection, however, was only part of the work. The Virality Project also sought to relate its findings to the public and to stakeholders in public health, government, and civil society. Outputs focused on creating situational awareness, public communication, and education. The Virality Project’s weekly briefings in particular provided stakeholders responsible for countering emerging vaccine mis- and disinformation with information that helped them allocate their resources and tailor their responses to have the greatest impact.

**Key Recommendations**

**Research Institutions**

Research institutions that study social media can help civil society and government partners better understand the dynamics of emerging narratives through their visibility into public platform data. Doing this most effectively requires two things:

- Lean in to the institution’s research strengths and engage research partners who complement those strengths. Define roles and responsibilities based on competencies; for example, counter-messaging efforts may go to one team while another examines messaging distribution.
- Streamline a tip line process to make it easy for civil society and government partners to share observations. Establish a feedback loop to discuss what types of analysis or tips are most relevant.

**Public Health Partners**

Public health experts know that social media content exacerbates vaccine hesitancy. They are also uniquely suited to provide accurate, trusted information to communities that become concerned about false or misleading claims. They can leverage their position most effectively through strategic communication:

- Focus on countering or addressing misinformation themes and tropes, not attempting to fact-check individual incidents.
Executive Summary

- Be transparent with the public about known unknowns.
- Communicate personal stories about the vaccine’s benefits with support from data; statistics alone are not the most compelling counter-messaging content.

Government

Federal, state, and local leaders need to position themselves to and develop capabilities for communicating to the public when harmful misinformation goes viral. This need extends beyond COVID-19. However, the government must also recognize that it is not always the best messenger for countering misinformation within targeted communities. A networked approach is needed:

- Develop and maintain clear channels of communication that enable federal, state, and local agencies to understand and learn from what might be happening in other regions. Federal Information Sharing and Analysis Centers (ISAC) are one path forward.
- Address vaccine hesitancy in underserved communities through collaboration with trusted voices—and build those relationships before they are needed.
- Implement a Misinformation and Disinformation Center of Excellence housed within the Cybersecurity and Infrastructure Security Agency.³
- Establish a rumor-control mechanism to address nationally trending narratives.

Platforms

Platforms owe the public transparency and accountability as they face the challenges of deciding what to surface, what to curate, and how to minimize the virality of harmful false claims. Tech platform policies against public health misinformation should be clear and precise, and their enforcement should be consistently applied. To these ends, platforms should:

- Consistently enforce policies, particularly against recurring actors.
- Continue to improve data sharing relationships with researchers.

Over the course of its seven months of work, the Virality Project observed narratives that questioned the safety, distribution, and effectiveness of the vaccines. Its work and its researchers appeared in outlets such as the New York Times, the Atlantic, and NBC News and informed the perspectives of reports from the Center for Strategic and International Studies and the Rockefeller Foundation;⁴ the Office of the Surgeon General incorporated VP’s research and perspectives into its own vaccine misinformation strategy.⁵ The Virality Project partnership demonstrated that a focused, well-coordinated collaboration between academia, nonprofits, and
the private sector can provide nimble analysis and support to stakeholders who desperately need it.

The enormity of the challenge demands a whole-of-society response and cross-agency collaboration. Moving forward, there is a need for a non-governmental independent research entity to spearhead VP-style collaboration around emerging mis- and disinformation.

The Virality Project report authors hope that this document offers a view not only of the collective challenge, but of a possible path forward.
Notes


Chapter 1

Introduction

In early 2020, as what was then called the “novel coronavirus” was largely confined to China, researchers at the Stanford Internet Observatory noticed emerging mis- and disinformation narratives surrounding the virus and related containment measures. As media worldwide covered the growing threat of the disease, Chinese state media attempted to downplay the emerging crisis and boasted of how effectively the Chinese Communist Party was handling the situation.1 Social media debates about the origin of the virus were incorporated into geopolitical power games as it spread; a variety of governments leveraged online theories, including bioweapon conspiracy theories, to deflect blame from their own citizens around their handling of the crisis and seize an opportunity to denigrate the response of rivals.2

The disease spread rapidly and globally, capturing mass popular attention worldwide. Few topics have commanded and sustained attention at such a global level, across both traditional media and social media. The public—as well as global authority figures and institutions—struggled to make sense of rapidly evolving theories about the disease, how it spread, and whether it could be treated or prevented. Sustained, widespread attention on the topic afforded a unique opportunity for governments, economically motivated actors, conspiracy theorists, and partisan activists alike to leverage the catastrophe in service to their interests. And so, alongside fear about the burgeoning pandemic came widespread concern about a parallel infodemic, a proliferation of mis- and disinformation that could harm both individual and public health.3

However, it was not always clear what was misinformation; in the case of the novel coronavirus, it was often simply not yet clear what was true or where scientific consensus lay. Social media companies found themselves facing a unique challenge

---

1See Appendix B: Definitions for definitions of “misinformation” and “disinformation” as used in this report.

2See Appendix B: Definitions for a definition of “narrative” as used in this report.
1. Introduction

as they sought to curate and surface accurate information about the pandemic: Ground truth about COVID-19 was rapidly evolving, and even institutional experts were not always aligned on the facts; however, communicating the most accurate information possible was critically important because of the potential significant impact on individual and community health. As the pandemic spread, infecting millions, people sought out cures even as wildly inaccurate claims about available medications and vitamin supplements went viral. Viral posts that claimed to have the answers to the public’s most pressing questions appeared online; fact-checkers struggled to evaluate them, and platforms wrestled with whether to leave them up or take them down. Social media influencers of varying backgrounds debated the merits and efficacy of masking, providing detailed breakdowns of their analyses in public posts. Institutional health authorities, by contrast, were often reticent and slow to communicate, even as the public sought answers and social media companies looked to them for authoritative information to amplify. A very significant set of challenges presented themselves: What should social media companies surface or amplify in the absence of clear scientific consensus? Which of the myriad influencers—some with massive audiences—were spreading reliable information? This challenge was exacerbated in the United States by the partisan divides that rapidly emerged as states organized responses to the pandemic. In the months before vaccines or treatments emerged, governments worldwide turned to preventative measures such as masking requirements and lockdowns. In the US, these measures were quickly framed as affronts to liberty by facets of the US right-wing political spectrum, turning individual responses to the virus into a function of political identity.

Within this chaotic environment, the long-established US anti-vaccine movement began to insert itself into the national conversation. Movements encompass coordinated, collaborative efforts to bring about social change; the anti-vaccine movement is well-coordinated and expansive domestically and internationally, and it increasingly collaborates with groups ranging from the Nation of Islam to right-wing militias. It has been moderately successful in its efforts to undermine an array of public health efforts nationwide, particularly those that attempt to strengthen childhood vaccine uptake. One of the keys to its success is a strong understanding of the tactics of social media activism. And so, in January 2020, when the virus that caused the pandemic did not yet have a name and talk of a vaccine was purely theoretical, anti-vaccine activists perceived an opportunity to undermine confidence in vaccinations writ large and to begin to frame the outbreak as part of a mass plot by a shadowy global cabal to force-vaccinate the global population. These theories were initially largely confined to insular echo chambers whose members had long focused on vaccines, alternative health, and wellness, as well as the QAnon conspiracy theory community. But on May 4, 2020, the video Plandemic burst onto

See Appendix B for a definition of “anti-vaccine influencer” as used in this report.
See Appendix B for definitions of “anti-vaccine” and “anti-vaccine activist” as used in this report.
the scene. It was a slickly produced litany of false and misleading claims about COVID-19, delivered by a discredited scientist presented as a whistleblower figure; the video made broad conspiratorial allegations of profiteering and cover-ups.\textsuperscript{11} *Plandemic* was successfully propelled across social media in part due to a highly mobilized QAnon online ecosystem.\textsuperscript{12} Before the major social media platforms began to take down the video—which was in violation of their COVID-19 misinformation policies—*Plandemic* amassed tens of millions of views and was shared into a wide variety of communities.\textsuperscript{13} Several prominent anti-vaccine activists began to post regularly about COVID-19; their followings began to increase, despite prior platform efforts to reduce the spread of false and misleading claims from anti-vaccine figures. As the possibility of a vaccine became more of a reality as 2020 progressed, anti-vaccine activists focused on preemptively undermining uptake.\textsuperscript{14} Several of the vaccines in development used relatively novel mRNA technology, which afforded an opportunity to present them as untested, unsafe, rushed, or risky, even to audiences who had taken all previously recommended vaccines.

It was against this backdrop that the Virality Project (VP) came together. A collection of research institutions had previously collaborated through the Election Integrity Partnership (EIP)\textsuperscript{v} to identify and understand the spread of election mis- and disinformation in the US during the 2020 presidential campaign.\textsuperscript{vi} In December 2020, these partners jointly observed that the same tactics used to great effect during the 2020 election were already in use to expand the spread of COVID-19 vaccine mis- and disinformation; for example, false and misleading claims spread quickly through tactics like “friend-of-a-friend” narratives\textsuperscript{15} or by way of highly influential recurring actors,\textsuperscript{vii} such as anti-vaccine activist Robert F. Kennedy Jr. The Project’s broad array of institutions enabled information sharing and rapid response when false and misleading information percolated across social platforms.

The Project’s original framing document articulated the threat:

\begin{quote}
A surge of anti-vaccine disinformation will pose significant challenges to the rollout and public adoption of COVID-19 vaccines in the United States. The anti-vaccine movement has well-developed online networks and expertise in leveraging social channels to spread its messages. For years, its activists have fought against state-level legislation related to school immunizations and have built coalitions with communities: those who believe that vaccination requirements are government overreach on one side, and wellness and parenting influencers on another. The movement is experienced and well funded, and it can easily generate in-the-streets action. It is adept at pushing targeted
\end{quote}

\textsuperscript{v}More information about the Election Integrity Partnership can be found at https://www.eipartnership.net/about.

\textsuperscript{vi}A full report of EIP’s findings is available at https://purl.stanford.edu/tr171zs0069.

\textsuperscript{vii}See Appendix B: Definitions for a definition of “recurring actor” as used in this report.
1. Introduction

messages to specific demographic communities, preying on fear. In the immediate future, it will expend significant efforts to enter mainstream conversation and erode confidence in the COVID-19 vaccines that come to market.

The Virality Project adopted a multistakeholder collaboration with civil society organizations, social media platforms, and government entities to respond to misinformation around the novel vaccines. The research institutions that comprised the Election Integrity Partnership—the Stanford Internet Observatory, the University of Washington’s Center for an Informed Public, the Atlantic Council’s Digital Forensic Research Lab, and Graphika—along with new partners the National Conference on Citizenship (NCoC)’s Algorithmic Transparency Institute and New York University’s Center for Social Media and Politics and Tandon School of Engineering—all elected to participate in this new initiative: the Virality Project. The partners brought together extensive research knowledge on the spread of online misinformation, quantitative social science research tools, and NCoC’s Junkipedia misinformation monitoring platform.

These partnerships allowed the Virality Project to effectively observe online conversations about COVID-19 vaccines for six months. This report details our findings and recommendations.

1.1 Persistent Tactics, Actors, and Narratives

Beliefs about vaccines shape vaccine uptake, and vaccine uptake has profound implications for the health and well-being of all societies, worldwide. In 2019, the World Health Organization listed vaccine hesitancy as a global health threat. While hesitancy dates back to the very first vaccines, the mechanisms by which people get information about immunizations, and the extent to which they trust the public health authorities who promote them, have shifted in the era of blogs and social media platforms. COVID-19, and its vaccines, emerged within this context.

Perhaps the most prominent example of misinformation about vaccines leading to decreased uptake can be found in the dynamics surrounding the measles, mumps, and rubella (MMR) vaccine. A 1998 study by Andrew Wakefield in The Lancet bore much of the responsibility for sparking fears about the safety of the MMR vaccine by alleging a correlation between the vaccine and autism. In the early 2000s, this study propelled the anti-vaccine movement into the public eye, with celebrities and politicians promoting their concerns about the vaccine on television and in the media. This misinformation and the resulting fear have persisted even though the study was retracted in 2010 and Wakefield lost his medical license, and despite dozens of subsequent studies finding no connection between the MMR vaccine and autism. Though the Centers for Disease Control and Prevention (CDC) officially declared measles eliminated in the US in 2000, undervaccinated communities remain
1.1. Persistent Tactics, Actors, and Narratives

at risk of outbreaks from visiting international travelers with the disease—as was seen in the 2018–2019 outbreaks in the Orthodox Jewish community in New York City, when a total of 649 cases of measles were confirmed within an eight-month period. The CDC has suggested that misinformation was a significant factor in contributing to that string of outbreaks.

As public life moved into online spaces in the early aughts, the anti-vaccine movement moved with it. As scientific research moved toward the consensus that vaccines were not causing autism, broadcast media gradually ceased providing coverage of the false claim. However, the internet has no editorial gatekeepers, and the anti-vaccine movement began to produce its own blogs and message board communities to further perpetuate the theory, particularly focused on new parents searching for information. Although it is difficult to measure the direct causal relationship, research has repeatedly suggested a correlation between misinformation spread online and increased levels of vaccine hesitancy. A 2010 study found that viewing a website critical of vaccines for just five to ten minutes could increase the viewer’s perception of vaccine risks and decrease the perception of the risks of vaccine omission. Other scholarship shows that as people increasingly turn to the internet for vaccination advice and are led to anti-vaccine websites, vaccine uptake suffers. The unique social structure of the internet can also help reinforce the spread of anti-vaccine viewpoints: even for many people who may be undecided about vaccination, repeated contact with hesitant peers online may shore up the perceived legitimacy of anti-vaccine beliefs. Although online clusters of anti-vaccine users may be smaller in overall size than pro-vaccine clusters, the online landscape allows anti-vaccine clusters to become highly entangled with clusters of users who are undecided on vaccines, enabling the potential for explosive growth in anti-vaccine views.

Connections made possible through social media platforms further enabled pathways for the spread of vaccine misinformation. Surveys conducted in 2019, during the US’s worst measles outbreak in 25 years, found that users of social media are more likely to be misinformed about vaccines than traditional media users. In the past few years, researchers have found that the use of social media to organize offline action is highly correlated with a belief that vaccines are unsafe—and that the more organization occurs on social media, the more strongly people believe vaccines are unsafe. The new opportunities afforded through social media, such as connecting via Facebook Groups, have enabled users to share, view, and discuss first-person experiences of supposed vaccine side effects—which can then have a disproportionate impact on parents as they try to assess the risks of vaccinating their child. Anecdotes, rather than expert knowledge or findings, are accepted as evidence. These changes, writes medical anthropologist Anna Kata, have ushered in a “new postmodern paradigm of healthcare...where power has shifted from doctors to patients, the legitimacy of science is questioned, and expertise is redefined.”
Public health officials have struggled to adapt to this new communication environment and peer-to-peer content dynamics. Anti-vaccine movement activists, in contrast, have excelled. Many activists use anti-vaccine content purposefully and skillfully, activating fear, uncertainty, and doubt and weaponizing a so-called health “debate” that in actuality has little to do with the debates of scientific and medical communities, which are in widespread agreement on vaccine safety and efficacy. Epidemiologist Tara Smith, writing in a 2017 paper on vaccine rejection and hesitancy, found that scientists (and their slow, measured, and nuanced information) have had difficulty counteracting “media savvy” anti-vaccine activists’ “strict black-and-white thinking.”

Smith, Kata, and other scholars have also documented the unique actors and tactical mechanisms by which anti-vaccine activists have expanded their activities on social media specifically. These mechanisms rely on a set of social media influencers that Smith divides into five distinct groups: The Doctors, The Celebrity, The Organizers, The “Mommy Bloggers,” and The Opportunists. Smith notes that many of these influencers are individuals or groups who benefit (often, financially) from the spread of health misinformation. Indeed, although movements are collaborative efforts, the influencers serve as local celebrities within the community; while some achieve financial gain, others achieve status. These influencers have adopted the best practices of communication in the internet age, and their effectiveness in drawing in online users is made evident by the mass followings they have acquired across social media sites: platforms as varied as Pinterest, Instagram, and YouTube have a significant anti-vaccine presence.

Similarly, anti-vaccine actors use tactics that have persisted over time, many of which have been used in service of spreading mis- and disinformation in contexts beyond the vaccine conversation, for example, the Election Integrity Partnership observed several of these tactics during the lead-up to the 2020 US election.

**Hard-to-Verify Content:** Using content that is difficult to fact-check or verify, such as personal anecdotes and deep fakes, content discussing information from a “friend-of-a-friend,” or opaque scientific information or analysis (either original or doctored in a misleading way).

**Alleged Authoritative Sources:** Using or pointing to information from an alleged public health official, doctor, or other authoritative source.

**Organized Outrage:** Creating events or in-person gatherings, or using or co-opting hashtags.

**Sensationalized/Misleading Headlines:** Using exaggerated, attention-grabbing, or emotionally charged headlines or click-bait.

---

**Note:** Some of these tactics are used by online activists in non-vaccine-related movements and/or without necessarily spreading mis- and disinformation. In this report, we focus on the ways these tactics are leveraged to spread anti-vaccine mis- and disinformation.
1.1. Persistent Tactics, Actors, and Narratives

**Group super-spreader:** An individual account sharing posts into multiple online groups.

**Harassment Campaigns:** Online abuse such as doxing, cyberbullying, hacking, or trolling.

### 1.1.1 Anti-Vaccine Themes, Tropes, and Narratives

The anti-vaccine influencers that Kata and Smith describe owe much of their success to the consistent themes, tropes, and narratives they employ to help sway and eventually build hesitancy in their audiences. These stories are not new, however; they date back to the introduction of the earliest vaccines and have been well documented in academic literature for centuries. In fact, when the smallpox vaccine was developed in 1796, it was met with strong opposition despite the fact that it protected against a disease with a 30% mortality rate. The nineteenth century saw the creation of organizations, such as the Anti-Vaccination Society of America, that spread narratives and conspiracy theories about immunization that are similar in many ways to those spread today about modern vaccines: they have toxic, immoral, or unnatural ingredients; their testing is insufficient; they are used by medical institutions for profit or by the government to invade its citizens’ personal lives.

Claims about smallpox vaccination, which used cowpox for inoculation, suggested that it would create human-cow hybrids or would sicken the vaccinated. More than two hundred years later, the specific claims put forth by the anti-vaccine movement have changed—there are no longer worries about human-cow hybrids—but its overarching themes have remained consistent. In 2017, Smith articulated a comprehensive taxonomy:

- **Toxicity:** Vaccines are unnatural or contain toxic ingredients.
- **Religiosity:** Vaccines contain materials that are objectionable on religious grounds.
- **Liberty:** Individuals have the right to “health freedom”; no government or employer should be able to tell people what to put in their bodies.
- **Distrust of industry:** Vaccines are produced by profit-motivated pharmaceutical companies that have repeatedly concealed harm in pursuit of profit.
- **Safety:** Vaccines are unsafe, undertested, or untested; the risks of vaccination outweigh the benefits.
- **Conspiracy:** Powerful forces conceal information about vaccines from the public or manipulate the people for malicious ends. While the specifics shift over time, common tropes include (but are not limited to):
  - Governments have covered up information proving vaccines are dangerous.
1. Introduction

- Vaccines contain microchips or tracking devices.
- Vaccines are a tool of population control.
- Doctors and politicians who advocate for vaccines have been bought off by “Big Pharma.”

1.2 Informing the Scope of the Virality Project

Drawing on this scholarship documenting the who, what, and how of the anti-vaccine movement, the Virality Project determined the scope of its research. The persistence of narratives—the fact that these categories of claims have resonated for so long—suggested a high likelihood that they would appear again in conversations about COVID-19 vaccines, and allowed researchers to understand and classify current narratives. The VP team adapted these well-established themes into four categories of narratives that would be considered in-scope for VP work:

- Safety
- Efficacy and Necessity
- Development and Distribution
- Conspiracy Theory

In parallel, the VP team assessed social media platforms’ published policies to understand how (if at all) platforms might limit or action the spread of misleading vaccine content. Platforms had started adapting their policies to address vaccine misinformation in early 2019, spurred by public outcry, negative press coverage, and government inquiries resulting from measles outbreaks across the world. At that time, many platforms also introduced new policies to address vaccine-related misinformation: Pinterest stopped returning all vaccine-related content, whether for or against vaccines; YouTube demonetized videos that promoted anti-vaccination views; and Facebook/Instagram launched a plan to reduce the spread of vaccine misinformation. These policy shifts rethinking what vaccine-related content was curated or recommended were significant first steps in mitigating the spread of the existing anti-vaccine content and narratives.

The pandemic, however, meant that platforms faced an unprecedented increase in the prevalence of anti-vaccine content and sentiments—months prior to any vaccine approval, anti-vaccine activists had already begun attempting to undermine confidence in what might be developed. These attempts to facilitate hesitancy were created and spread by similar, or the same, actors using existing narratives of earlier anti-vaccine movements. For the Virality Project, this meant that analysts had to

---

ix For limitations in VP’s research, see Appendix A: Research Limitations.
1.2. Informing the Scope of the Virality Project

develop a nuanced and nimble understanding of what content constituted policy violations.

It should be noted that the Project’s scope was limited so as not to engage with the broader range of general COVID-19 misinformation, such as generic conversations about lockdown decisions, mask wearing, false cures, or the severity of the virus itself. Even within this demarcation, we anticipated seeing a variety of tactics, split between domestic activists and state actors with incentives to shape the discourse.

1.2.1 Detecting In-Scope Content

To surface in-scope content, VP’s team of analysts were divided into topical detection teams, referred to as pods (see Table 1.1 on the following page). These pods—some inspired by Smith’s vaccine-specific taxonomy, others designed with common sources or targets of mis- and disinformation in mind—enabled analysts to develop and ensure sustained familiarity with how the COVID-19 vaccine conversation was evolving within particular communities on public platforms. The topically focused communities were defined by the primary type of content they shared: the topics around which the pages or accounts were optimized. Analysts in each pod assessed emerging narratives that were within scope (as defined above), surfacing content both via qualitative observation of the pages and accounts, and by using lists of common terms associated with vaccine hesitancy and long-standing anti-vaccine rhetoric.
1. Introduction

Table 1.1: Descriptions of the pods analysts used to surface tickets.

<table>
<thead>
<tr>
<th>Pod Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conspiracy Theory</td>
<td>Conspiracy theory-oriented groups on social media such as QAnon and adjacent communities.</td>
</tr>
<tr>
<td>Right-Leaning</td>
<td>Right-leaning media, celebrities, influencers, and activists.</td>
</tr>
<tr>
<td>Left-Leaning</td>
<td>Left-leaning media, celebrities, influencers, and activists.</td>
</tr>
<tr>
<td>Minority Communities</td>
<td>Focus on anti-vaccine content targeting African American, Asian American, Hispanic/Latino/a, and Native American communities.</td>
</tr>
<tr>
<td>Non-English Language</td>
<td>Focus on-English-language content targeting US Chinese and Spanish language communities.</td>
</tr>
<tr>
<td>Foreign Influence</td>
<td>Overt foreign state-backed media accounts, specifically from China, Russia, and Iran. Analysts in this pod participated in investigations into covert or obscured attribution activity, such as the Fazze operation, described in Chapter 4: Foreign Actors.</td>
</tr>
<tr>
<td>Lifestyle and Spirituality</td>
<td>Holistic, wellness, parenting, and other lifestyle communities online. This pod also focused on spiritual and religious communities, as analysts observed frequent content overlap between these clusters of accounts.</td>
</tr>
<tr>
<td>General Anti-Vaccine</td>
<td>Anti-vaccine activists and influencers, including those discussed in the Center for Countering Digital Hate's &quot;Disinformation Dozen&quot; Report. This pod surfaced the greatest amount of content compared with the other five pods.</td>
</tr>
</tbody>
</table>

At the beginning of the project, analysts used broad search terms ("vaccine," “jab”) to surface relevant content and incidents (specific events or stories), but gradually began to incorporate a combination of machine learning and hand coding to identify additional recurring narratives relevant to the four in-scope categories. This included terms related to medical freedom under “Vaccine Distribution,” or severe adverse effects and death under “Vaccine Safety,” among others. As narratives and new keywords emerged throughout the analysis period, analysts continually refined their searches.

---

*The Project originally began with two pods specific to Right-Leaning and Conspiracy Theory, but ultimately merged the two research teams because of recurring content overlap between them: QAnon supporters backed right-leaning politicians, and right-leaning media promoted QAnon and adjacent communities’ claims.

See Appendix B: Definitions for a full definition of “incident.”
1.2.2 External Stakeholder Engagement

The Virality Project established a nonpartisan, multi-stakeholder model consisting of health sector leaders, federal health agencies, state and local public health officials, social media platforms, and civil society organizations. These stakeholders provided tips, feedback, and requests to assess specific incidents and narratives, and each entity type brought specific expertise to bear on understanding COVID-19 vaccine hesitancy:

- **Civil society organizations** engaged with VP by both providing tips and receiving VP outputs that enabled them to better understand emerging incidents and narratives that might impact their communities. Community partners were well positioned to observe anti-vaccine misinformation spreading within their communities, as well as to create or deliver effective and accurate counter-messaging to address it. They offered a trusted voice that could also convey up-to-date facts on vaccine safety and availability to counter false or misleading claims.

- **Public health partners** had a deep understanding of the latest research on COVID-19 and the vaccines, as well as prior experience communicating the value of immunization to the communities they serve. What they often lacked was the capability to quantify what false and misleading narratives were going viral, or spreading outside of long-standing anti-vaccine echo chambers, to determine what might benefit from their attention. Partnerships and open communication channels with health practitioners, such as the #ThisIsOurShot (#TIOS) collective and the American Academy of Pediatrics (AAP), facilitated information sharing between VP researchers and frontline health workers. Health practitioners in turn communicated the anti-vaccine and vaccine hesitancy narratives they heard from patients, and offered visibility into the impact the narratives had on vaccination behavior. Doctors also helped the Virality Project understand medical consensus in a rapidly evolving environment by sharing the latest findings on the vaccines (such as the rare link between mRNA vaccines and myocarditis/pericarditis). Partnerships with local health officials through the California Department of Public Health (CDPH) and networked organizations such as the National Association of County Health Officials (NACCHO) offered visibility into geographic communities.

- **Federal government agencies** served as coordinators for national efforts. The Virality Project built strong ties with several federal government agencies, most notably the Office of the Surgeon General (OSG) and the CDC, to facilitate bidirectional situational awareness around emerging narratives. The CDC’s biweekly “COVID-19 State of Vaccine Confidence Insights” reports provided visibility into widespread anti-vaccine and vaccine hesitancy narratives observed by other research efforts.
1. Introduction

- **Platforms** were the final stakeholders in the VP effort. Six social media platforms engaged with VP tickets—Facebook (including Instagram), Twitter, Google (including YouTube), TikTok, Medium, and Pinterest—acknowledging content flagged for review and acting on it in accordance with their policies. On occasion, platforms also provided information on the reach of narratives previously flagged by VP, which provided a feedback loop leveraged to inform the Project’s understanding of policies and ongoing research.

As the effort progressed, input from these partners was crucial in defining the VP’s output formats and in surfacing where the impacts of vaccine mis- and disinformation were being felt offline.

### 1.3 Outputs and Impact

Outputs focused on real-time response, situational awareness, public communication, and education, which took the form of weekly briefings, rapid responses, and blog posts.

The Virality Project delivered 31 weekly briefings focused on increasing situational awareness and enabling the stakeholders working on countering vaccine mis- and disinformation to develop the most effective possible response (see Figure 1.1 on the next page for a sample briefing). Recipients included 88 public health partners across 34 organizations. Briefings discussed incidents observed by analysts throughout the week, contextualized with engagement statistics and analysis of how they fit into long-standing narratives and themes. Incidents were chosen for inclusion in the briefing if they represented a new emerging narrative, showed worrisome or surprising viral spread, weaponized consistent anti-vaccine tropes or tactics, or showed spillover from a long-standing, pre-pandemic anti-vaccine group into mainstream online dialogue.

Briefings and their corresponding weekly emails encouraged a networked response by pointing recipients not only toward other longer-form VP publications, but also to broader informative events and panels by additional organizations working to counter false and misleading information or reduce vaccine hesitancy. Stakeholders received examples of counter-messaging from The Public Good Project’s Stronger campaign, which fights mis- and disinformation about science, medicine and vaccines. Briefings directly informed counter-messaging efforts by public health stakeholders (for example, #ThisIsOurShot; see Figure 1.2 on the facing page) and public health officials (for example, the CDPH), and provided strategic insights to government entities such as the OSG, CDC, and the Department of Health and Human Services. Information and analysis included in the briefings were also cited in media such as the *New York Times*.

Meanwhile, rapid responses and blog posts were used to flesh out analysis of the
1.3. Outputs and Impact

Figure 1.1: Anatomy of a briefing.

Figure 1.2: A Virality Project briefing insight inspired content from the doctor collective #ThisIsOurShot.
incidents, trends, or problems touched on with more brevity in the briefings. Rapid responses, posted on the Virality Project website and to social media, covered events happening in real time, in order to call attention to—and enable public health or government partners to disrupt—false and misleading narratives early in their development. For example, four days after the FDA and CDC recommended pausing the Johnson & Johnson vaccine, VP researchers posted an analysis of the anti-vaccine community’s framing of this incident. Blog posts were written to examine themes and repetitive tactics; some were retrospective, analyzing successful anti-vaccine tactics or conceptual strategies such as content moderation avoidance strategies.

Additionally, our researchers spoke directly with media outlets and served on panels discussing anti-vaccine misinformation. Stanford Internet Observatory and the Virality Project also hosted Surgeon General Vivek Murthy for a seminar on vaccine mis- and disinformation, including the rollout of the Surgeon General’s advisory on health misinformation.

1.4 Reading this Report

This report details our findings and recommendations from observing six months of COVID-19 vaccine discussions online. Each chapter covers a different portion of our findings:

Chapter 2: Data Collection & Statistics outlines the tiered ticket analysis structure and online engagement tracked by VP analysts during its observational period, assessing the tickets across engagement, narrative, and community trends.

Chapter 3: Narratives does a deep dive into significant case studies, highlighting examples of both momentary and persistent narratives across the four in-scope categories tracked by VP analysts.

Chapter 4: Actors & Tactics describes notable influencers in the COVID-19 vaccine conversation and the tactics they employed, based on incidents surfaced in VP tickets and briefings.

Chapter 5: Platform Policies for Anti-Vaccine Content provides an overview of the policy frameworks and moderation approaches that social media platforms took to address COVID-19 vaccine mis- and disinformation, to varying effect.

Chapter 6: Recommendations for Stakeholders highlights takeaways and offers recommendations for stakeholders working in research, public health, and government and at social media platforms.

We hope that readers can use this report to deepen their understanding of the online dynamics surrounding the vaccine rollout in the United States during the COVID-19
pandemic. We additionally hope that the Virality Project’s multistakeholder partnership can serve as an inspiration for a whole-of-society approach toward enabling a more collaborative and effective response to future public health concerns and emerging health threats.
Notes


10. Haelle, “This Is the Moment the Anti-Vaccine Movement Has Been Waiting For.”


27. Annenberg Public Policy Center of the University of Pennsylvania, “Vaccine misinformation and social media: People exposed to vaccine content on social media more likely misinformed than those exposed to it on traditional media,” ScienceDaily, February 17, 2020, https://www.sciencedaily.com/releases/2020/02/200217163004.htm.


33. Centola, “The Complex Contagion Of Doubt In The Anti-Vaccine Movement in Meeting the Challenge of Vaccine Hesitancy.”

34. Tara Smith, “Vaccine Rejection and Hesitancy: A Review and Call to Action,” *Open Forum Infectious Diseases* 4, no. 3 (Summer 2017), https://doi.org/10.1093/ofid/ofx146.

35. T. Smith.

36. T. Smith.


40. T. Smith, “Vaccine Rejection and Hesitancy: A Review and Call to Action.”


42. DiResta and Wardle, “Online Misinformation About Vaccines.”


44. ThisIsOurShot, “This is Our Shot,” accessed October 10, 2021, https://thisisourshot.info/.


49. TIOS, “This is Our Shot.”

There are a few things that have been shown to affect male fertility, but the COVID-19 vaccine is not one of them. Trust science: Wear boxers and get vaccinated. #ThisIsOurShot #VacunateYa #WeCanDoThis,” Twitter, July 1, 2021, https://twitter.com/ThisIsOurShot/status/1410584124314861568.


Data Collection & Statistics

This chapter presents an analysis of the Virality Project’s observed user engagement with vaccine mis- and disinformation on social media through three lenses: narrative categories, tactics, and community focus. We explore specifics related to the content and actors within these three lenses in later chapters, but the focus here is on discussing the engagement: the likes, shares, reactions, and other quantifiable ways in which social media platform users interacted with the material.

As part of the Virality Project, analysts created tickets documenting URLs of in-scope content. In total, 911 tickets were created, tracking both specific pieces of misinformation and broader narratives. At the end of the monitoring period, analysts had created 845 tickets tracking specific vaccine misinformation incidents (events or pieces of content) and 66 tickets tracking broad narratives. The number of incidents monitored and tickets created each week was relatively steady during the period of the project: analysts observed on average 35 incidents per week, with 14 during the quietest week and 59 during the busiest week.

2.1 Workflow and Analysis Methodology

The Virality Project operated with a team of analysts drawn from across the partner organizations. Workflows were designed to detect, analyze, and respond to incidents of COVID-19 vaccine-related disinformation in online ecosystems (see Figure 2.1 on the next page).

2.1.1 Monitoring, Intake, and Ticketing System

From February to August 2021, VP analysts systematically monitored activity across social media platforms to document emerging narratives and trends in public discourse while also tracking the popularity and spread of older content. Analysts
used both keyword-based queries and cluster-based “community” lists of key public figures and influencers.¹ As noted in Chapter 1, the first iteration of these lists were derived from earlier academic work describing groupings of influencers in the long-established vaccine conversation.¹ Analysts parsed through content on pod-specific lists and from narrative queries three times per week, focusing on major social media platforms, including Facebook, Instagram, Twitter, Reddit, and TikTok. VP also adopted tools to expand visibility into emerging and alternative platforms such as Gab, Parler, Telegram, and Gett. Language specialists observed content in Spanish and Chinese—the two most spoken languages in the US after English. The range of inputs enabled the researchers to capture the breadth of online discussion and to assess whether virality (the spread of content across platforms or communities) or velocity (the rate of engagement with content) metrics suggested that a particular incident, claim, or narrative might be gaining or achieving significant importance or reach, making it potentially relevant for platform partner or external stakeholder awareness.

The Project used the Jira Service Desk software to log mis- and disinformation incidents that were determined to be in scope for specific areas of the public COVID-19-related conversation.¹¹ For each single incident of anti-vaccine mis- or disinformation surfaced during monitoring, an analyst filed a ticket that provided a brief description of the incident, including engagement numbers at the time of creation and links to relevant social media posts (see Figure 2.2 on the facing page for a

---

¹The process by which VP developed its keyword searches is explained in Chapter 3: Narratives.
2.1. Workflow and Analysis Methodology

sample ticket). In the initial iteration of the monitoring process, analysts filed all incidents or content as unique tickets. As the database of tickets grew, VP researchers refined the surfacing and filing process to include narrative tickets that created groupings around specific, recurring sub-narratives (e.g., Vaccine Conspiracy Theories: Magnetism). These narrative tickets allowed analysts to compile ongoing content into the key categories of Safety, Efficacy and Necessity, Development and Distribution, and Conspiracy Theory (defined in Chapter 1: Introduction).

![Sample Issue](image)

Figure 2.2: A sample ticket. Left, analysts completed their assessment and observations of the incident. Right, analysts provided appropriate information to help categorize and prioritize the ticket. Tickets could be linked together around larger incidents or narratives.

---

Atlassian’s Jira Service Manager ticketing platform enabled submissions from external partners and allowed analysts to report incidents to platform partners.
2. Data Collection & Statistics

Tickets also enabled analysts to quickly tag platform or health sector partners to ensure their situational awareness of high-engagement material that appeared to be going viral, so that these partners could determine whether something might merit a rapid public or on-platform response (such as a label).

**Tiered Ticket Analysis**

Once tickets were filed, they went through a tiered analysis process. Analysts were broken down into two tiers. Tier 1 analysts, typically undergraduate and graduate students who were also scheduled for monitoring shifts, were tasked with providing initial ticket analysis. Their analysis consisted of lateral research that used CrowdTangle and Google searches to assess the spread of the incident or content and to surface as many content links as possible to attach to the ticket. Analysts then sent the ticket to an on-call manager for review.

The on-call manager prioritized what tickets were in or out of scope, and which deserved a deeper analysis. The manager sent the tickets slated for deeper analysis to Tier 2, which consisted of upper-level researchers, PhD students, and postdoctoral fellows. Tier 2 analysis provided:

- a wider snapshot of the prevalence of the narrative or incident across different platforms and insight into which accounts were responsible for pushing the content,
- background to the incident or ticket that incorporated domain knowledge, and
- a sense of whether something related to the ticket appeared suspicious—perhaps a fake account or evidence of coordinated inauthentic behavior.

The Tier 2 analyst recommended whether a particular ticket should be made into a Virality Project blog post.

The last step was Manager Review. Managers gave all incident analyses a final review for quality-control purposes, to determine appropriate next steps and to make a final decision about whether tickets should be shared with external stakeholders. Of the 911 incidents monitored, 174 were referred to platforms for potential action.

2.1.2 Engagement Data Collection

As noted above, we use the term *engagement* to refer to user interactions. This takes different forms on different platforms—for example, reactions, comments, shares, or video views—but in general the term covers actions taken by a user on a platform in response to content.

Analysts determined the engagement data associated with each ticket by summing the engagement of each social media link associated with that ticket, weighing all
types of engagement equally. The engagement data or video view data for links associated with each ticket was collected differently depending on the social media platform in question:

- Facebook and Instagram: CrowdTangle API
- Twitter: Twitter API
- YouTube: YouTube API
- TikTok: no relevant API available; data scraped from TikTok using an unofficial open-source TikTok library
- Non-social media links such as news articles: CrowdTangle API to capture total public interactions on Facebook

Engagement over time was measured from the date an analyst created the ticket. In VP researchers’ experience, engagement with most social media content has a half-life of less than one week; generally, most engagement with a post happens within the first week. Therefore, we believe this to be an acceptable simplification and present engagement dates aggregated by week. If analysts identified multiple narratives, tactics, or communities of focus, they assigned engagement to the primary group in question.

2.2 Engagement: An Imperfect Metric

Due to limited transparency from social media platforms, engagement is the closest proxy researchers can use to understand what content users are seeing on social media platforms. Metrics such as impression counts are generally unavailable to outside researchers. As such, user engagement should be understood only as a measure of how much users are interacting with content, and nothing more. However, because social media platforms, including Facebook and Twitter, generally optimize for user engagement—using engagement to rank what content is shown in users’ news feeds—this metric has importance for understanding what kind of content may achieve greater reach via content promotion or curation algorithms, absent deliberate efforts to counter its spread.

We note that the differing engagement levels of different posts we discuss in this section are reflective not only of the content they contain, but also of the existing audiences of the actors that promote them. A post from a Facebook Page with millions of followers will always receive more engagement than a post from a Page with a few hundred.

Platforms’ engagement-based ranking can increase the spread of mis- and disinformation, particularly if it is novel or sensational. Vosoughi, Roy, and Aral show that

---

iii https://github.com/davidteather/TikTok-Api
fake news travels faster on Twitter than factual content,\textsuperscript{4} while Edelson et al. show that misinformation news sources on Facebook are significantly more engaging than factual news sources.\textsuperscript{5} Rathje, Van Bavel, and Van der Linden show that, in addition to negative affect and moral-emotional language, out-group animosity is a powerful driver of user engagement with social media content on both Facebook and Twitter.\textsuperscript{6}

### 2.3 Engagement with Vaccine Misinformation Over Time

Average weekly engagement with content tracked across all Virality Project tickets was 6.7 million. That average declined over time. However, engagement with the content was highly volatile, with median weekly engagement at 4.6 million and a standard deviation of 6.4 million. This high degree of variance makes the overall decline in engagement not statistically significant (see Figure 2.3 on the next page).

We present engagement aggregated by incident (or ticket). Engagement per incident was exponentially distributed, which is typical of engagement on social media. When engagement is considered temporally, there is no consistent pattern of engagement that would indicate a core audience who is always engaging with vaccine misinformation; instead, wide swings week to week may be more indicative of changes in how compelling more recent vaccine misinformation content is.

### 2.4 Engagement by Category

Over the entire monitoring period, analysts classified tickets into the four key categories of narratives.\textsuperscript{4} Of the 911 tickets tracked, roughly a third related to Safety, with Development and Distribution, Conspiracy Theory, and Efficacy and Necessity each taking roughly 20\% (see Figure 2.4 on the facing page). The remaining tickets didn’t fall into any of these narrative categories.

However, these absolute percentages of tracked tickets differed from the distribution of engagement. As seen in Figure 2.5, the leading category for engagement was Conspiracy Theory, followed closely by Safety. Development and Distribution narratives came in third, while Efficacy and Necessity and "other" narrative categories each represented only a small percentage of engagement. When considered on a per-incident basis, Conspiracy Theory narratives had the highest engagement of any category (see Chapter 3 for more on narratives).

There were no strong patterns observed in engagement by category over time: different categories had different engagement levels at different times, but there

\textsuperscript{4}These categories are discussed more fully in Chapter 3: Narratives.
2.4. Engagement by Category

Figure 2.3: Overall engagement with vaccine misinformation content from February 1 to July 31, 2021.

Figure 2.4: Number of tickets per category, out of 911 tickets.
2. Data Collection & Statistics

is no single prevalent trend. Safety had a relatively consistent share of overall engagement over the monitoring period. Conspiracy Theory was the most variable category, with extreme volatility in engagement from week to week.

2.5 Engagement by Tactic

The most commonly employed tactics\(^v\) were Hard-to-Verify Content and Alleged Authoritative Source. The other tactics that were large enough to quantify, including Organized Outrage, Sensationalized/Misleading Headline, and Harassment Campaigns, each took up less than 10% of engagement. These tactics also took up a relatively small share of the number of incidents (see Figure 2.6 on the next page).

Unlike category engagement, the tactics most engaged with changed over time. Notably, while Hard-to-Verify Content drove several high engagement incidents in the first half of the monitoring period, by the end it was driving little meaningful engagement. By contrast, the Alleged Authoritative Source tactic surged late in the monitoring period.

2.6 Engagement Driven by Recurring Actors

Of the engagement captured by tickets, more than a third came from content primarily spread by accounts that demonstrated recurring success making content go viral; we refer to them here as “recurring actors.”\(^vi\) Many, but not all, of these accounts were long-standing anti-vaccine activists (discussed in greater detail in Chapter 4 on page 74). VP analysts tagged Jira tickets involving these actors whose

\(^v\)Descriptions of these tactics can be found in Chapter 1: Introduction.
2.6. Engagement Driven by Recurring Actors

Figure 2.6: Engagement by primary tactic.

Figure 2.7: Engagement by tactic over time.
2. Data Collection & Statistics

Social media content regularly included vaccine mis- and/or disinformation and who had large social media followings, in recognition of the increased likelihood of their content to go viral. Other research teams have additionally noted the important and outsized role that a small number of well-known accounts had in the spread of vaccine mis- and disinformation.\(^7\) Recurring actors drove a majority of engagement in the first half of the study period, but fell off in importance after that, most likely due to platform action against certain users beginning in the late spring of 2021.\(^8\) These recurring actors employed different tactics in their messaging. However, they all had one thing in common: large audiences on social media platforms. Therefore, it is possible that when they created content, or reshared content that originated from other actors, those posts were exposed to a larger audience and had a better chance of going viral.

### 2.7 Engagement by Community

The most engaged community of focus was the broader conspiracy theory community, with 37% of engagement during the monitoring period. This is particularly impressive given that only 20% of incidents appeared to emerge from within the conspiracy theory community. Notably, the anti-vaccine community had only the second largest share of engagement, despite its larger percentage of incidents.

It is difficult to discern exactly why the broader conspiracy theory community generated more engagement with vaccine mis- and disinformation than the anti-vaccine community. First, the relative size of these communities is unclear, and in particular how large they are on the observed platforms. Second, the relative likelihood of these communities to engage with this kind of content, or how likely content aimed at them is to spread, is unclear. However, given that 70% of engagement with the content was not aimed at the existing anti-vaccine community, one thing is clear: engagement with anti-vaccine content was widely promoted to communities outside the anti-vaccine community. Conspiracy theories appear to have been particularly effective at escaping the “bubble” of their established community by going viral.

### 2.8 Viral Incidents

Four distinct weeks during the monitoring period had incidents observed by Virality Project analysts that generated more than 10 million engagements. In general, this was driven by a single incident each week that had very high engagement, although one of these weeks had two separate incidents with high engagement. These specific high engagement incidents are noted in Figure 2.10 on page 38.

\(^7\)See Appendix B: Definitions for a definition of “recurring actors” as used in this report, and Chapter 4: Actors and Tactics for a more detailed discussion.
2.8. Viral Incidents

Figure 2.8: Engagement by community focus.

Figure 2.9: Engagement by community focus over time.
2. Data Collection & Statistics

Figure 2.10: Engagement by category over time, with callouts to four specific incidents.

This figure depicts spikes that correspond to incidents that went viral:

- The Fake Needle conspiracy theory, which purported that the needles used to give vaccine injections to those posting positive images of themselves receiving their shots were in fact retractable, fake needles that were not actually injecting recipients with the vaccine.\(^9\)

- The “18 Reasons” spike was tied to a Chinese-language incident focusing on a list of safety reasons not to get vaccinated.

- The Magnet Challenge multilanguage conspiracy theory claimed that vaccines contained magnets from microchips, which made vaccinated people “magnetic.” Users were challenged to spread copycat videos to “prove” that injection sites had become magnetic.\(^vii\)

- In July, posts went viral expressing outrage at attempts by the Biden administration to engage in vaccine outreach. The graphene oxide conspiracy theory also went viral at this time.

The Fake Needle conspiracy theory was first observed on Facebook on March 24, 2021, and the original post did not have very high engagement. However, after this initial Facebook post, multiple TikTok videos promoting the same content went viral, racking up 26 million views and interactions. The vast majority of the

\(^vii\) The Magnet Challenge is discussed in more detail in Chapter 3: Narratives.
2.8. Viral Incidents

engagement accumulated for this incident happened on TikTok. A separate video promoting the same concept was also active on Instagram, but it only achieved approximately 10,000 interactions. Later, the TikTok accounts promoting these videos were removed along with the video content, though there is no way of knowing if this was due to platform action or was voluntary on the part of the users. The Facebook and Instagram posts are both still available, although the video in the Facebook post has been removed by the original author and the Instagram video has been hidden behind a “false information” label that users must click through to see the content. The content was spread by known recurring actors.10

The second viral incident related to an article entitled “18 Reasons I Won’t Be Getting a COVID Vaccine.” These “18 Reasons” primarily, but not exclusively, relate to vaccine safety narratives. Initially, they were spread in English by known recurring actors, but they were then translated to Chinese. That translated content was posted to Facebook, Instagram, and Twitter in Chinese beginning April 16, 2021, but did not achieve significant engagement there. The translated content was then posted to WeChat, where it quickly went viral, achieving nearly 13 million interactions. The content was removed from WeChat relatively quickly and the accounts that posted it are no longer available, although, again, it is unclear if this was due to platform action or was voluntary on the part of the users. Twitter removed this post but did not shut down the user’s account, and Facebook neither removed the post nor acted against the account.

The third observed viral incident, the “Magnet Challenge,” appeared to begin on May 9, 2021. The Magnet Challenge, based on a conspiracy theory that the vaccine had harmful ingredients like microchips, showed a video of a refrigerator magnet sticking to someone’s arm and challenged users to see if they could stick a magnet to their vaccine site.11 Videos promoting this challenge spread on nearly every monitored platform; however, they only went viral on TikTok. Also notably, the Magnet Challenge was promoted effectively in multiple languages, including Chinese, Spanish, and English. This might be due to the nature of the conspiracy theory—videos demonstrating magnets supposedly sticking to arms require only minimal verbal explanation that can be easily translated into other languages.

The high-engagement incidents during the fourth weekly spike were related to two incidents. In the first, the Biden administration used the phrase “door-to-door” to describe a push for on-the-ground community-led efforts to persuade more Americans to get vaccinated. Prominent Republican politicians miscast this as a forced vaccination campaign by “Needle Nazis” and a prelude to the government knocking on the door to take away guns.12 The second incident—the graphene oxide incident—claimed that the Pfizer vaccine contained this poisonous compound.
2.9 Conclusion

When we examine online engagement on social media incidents of vaccine mis- and disinformation, as distinct from number of incidents, we see a somewhat different picture of the overall landscape of content. Engagement is highly variable and is heavily driven by viral content that spreads outside its initial community of focus. Conspiracy theory-based incidents are especially notable here, as although these incidents made up a relatively smaller absolute number of incidents, they were much more engaging than other categories. Incidents that promote conspiracy theories clearly have the potential to be highly engaging, and this content must be treated as such in the context of an online landscape increasingly dominated by conspiracy theory communities.

The dynamics surrounding recurring actor accounts and their audiences merit additional research. The decline of content from recurring actors midway through the monitoring period potentially reflects a policy impact, as deplatforming these actors led to an apparent reduction in false or misleading content. However, these actions did not lead to an overall statistically significant decrease in engagement around vaccine mis- and disinformation—at least in part because other successful distributors and unique tactics filled the void left by those actors.

Finally, the dynamics of the spread of anti-vaccine content point to a limitation of tagging and quantitative analysis of processes that involve some degree of subjectivity. For example, the graphene oxide incident illustrates the difficulty of labeling some types of claims when creating tickets. Virality Project analysts did not consider this incident to be in the Conspiracy Theory category, as it primarily centered around vaccine development, distribution, and safety. However, analysts did observe the claims spreading within conspiracy theory groups, and several recurring actors used “Alleged Authoritative Source” dynamics to promote the story.
Notes

1. T. Smith, “Vaccine Rejection and Hesitancy: A Review and Call to Action.”
7. CCDH, “The Disinformation Dozen.”
Vaccines have been the subject of false and misleading claims since their earliest use. Anti-vaccine tropes—similar plot devices and archetypes—and narratives have persisted over time and adapted to the introduction of new vaccines. Recognizing the likelihood that these narratives would once again reappear with slight inflections for COVID-19 vaccines, the Virality Project sought to observe COVID-19 narratives across four key long-standing categories: (1) vaccine safety, (2) efficacy and necessity, (3) distribution and development, and (4) conspiracy theories (see Table 3.1 on the facing page).

In this chapter, we describe prevalent narratives within each of these four categories and present several case studies to illustrate various ways that they gained traction online, shifting alongside the changing landscape of the pandemic and response efforts (such as with the rollout of vaccines to new populations).

We define a narrative as a story that connects a series of related events or experiences, providing compelling interpretations that can help people make sense of these events and experiences. For example, a commonly observed narrative was that individuals are dying after receiving vaccines, so vaccines are unsafe. Here, the narrative uses a series of events masquerading as evidence (people are dying) to imply something about the vaccine (vaccines are unsafe).
3.1 Narrative Categories

The Virality Project followed four categories of narratives that have remained prevalent in vaccine hesitancy conversations over a long period of time, and that were inflected for COVID-19 as the pandemic began and persistent throughout the vaccine rollout:

- **Safety**: Claims that the COVID-19 vaccines cause harm to recipients.
- **Efficacy and Necessity**: Claims that vaccines are not necessary or effective in combating the spread of COVID-19.
- **Vaccine Development and Government Distribution**: Claims that misrepresent vaccine production, distribution plans, and vaccine mandates.
- **Conspiracy Theories**: Claims fueled by distrust of authorities that implicate individuals or institutions to suggest there is malicious intent behind creating or requiring the vaccine.

**Safety**: The safety of vaccines was the most commonly observed category; analysts tagged more than 50% of the 911 total tickets filed as raising vaccine safety concerns.

---

1See Chapter 1: Introduction for a discussion of how VP determined and detected in-scope content.
Vaccine safety-related content often played into fears that the COVID-19 vaccines are new—developed too rapidly or with “unproven” technology—or that their side effects are still uncertain or even harmful. Some content leveraged decontextualized statistics from the US Department of Health and Human Services’ Vaccine Adverse Event Reporting System (VAERS) database. Stories of people dying or developing health problems after receiving a vaccine often gained significant traction, particularly if there was a first-person video in which an individual described a bad experience. Rare incidents documenting verified adverse health events, including blood clotting and heart inflammation, were shared absent context, often in an effort to present them as common and significant risks.

**Efficacy and Necessity:** False and misleading narratives related to efficacy sought to undermine the perceived benefits of vaccines. These narratives included stories of people diagnosed with COVID-19 after being vaccinated—“breakthrough” cases, particularly in the time of the Delta variant—to promote the idea that the vaccines aren’t effective. Later, the idea that natural immunity from infection is superior to immunity from vaccination became a political talking point raised repeatedly by right-leaning political influencers, despite inconclusive scientific evidence.

**Development and Distribution:** This category encompasses narratives about how the vaccines were developed and distributed, with the latter often centering on discussions of vaccine passports and mandates (including months before any state or federal officials began advocating for them). These discussions were often leveraged in combination with false and misleading claims about the safety or efficacy of vaccines to further persuade or entrench audiences, such as in claims that the government was headed toward mandating an unsafe vaccine. Some content in this category also leveraged racial and ethnic minority groups’ distrust of government and public health institutions; some prominent anti-vaccine activists exploited historical concerns about medical racism to foster greater distrust among marginalized communities, such as by alleging that public health institutions were “testing” an unproven medical product on those communities. Additionally, VP analysts observed long-standing mistrust of pharmaceutical companies’ profit motives leveraged to delegitimize vaccines by alleging that their development was a profit-maximizing rush job. These themes were particularly salient within discussions of booster shots, which were framed as an additional profit enterprise.

**Conspiracy Theories:** While many vaccine hesitancy beliefs stem from a degree of distrust in authority or institutions, conspiracy theories are differentiated by the belief that an individual or group is acting (or not acting) deliberately to create or exacerbate the set of circumstances that the theory describes. The narratives in this section generally assign blame to, or express distrust of, notable figures and institutions—from government and Big Tech to the “global elite”—imputing their involvement as an interpretation layer atop other claims. Many of the underlying claims that the conspiracy theories seek to explain emerge from within the other
categories—“vaccines are injuring people, and are therefore unsafe, because of the profit motives of Big Pharma”—which gives them a degree of familiarity or plausibility, and there may be a grain of truth in the underlying set of facts that the conspiracy theory attempts to explain. For instance, VP analysts repeatedly surfaced conspiracy theory claims that Bill Gates was behind a plot to implant vaccinated people with tracking devices or microchips, that the government was covering up vaccine deaths, and that the vaccine had been developed prior to the emergence of COVID-19 (implicit “proof” that the pandemic had been planned). Although many of these claims remained confined to conspiracy theory-focused online communities, analysis revealed that conspiracy theories about the vaccines had high engagement per incident.

### 3.2 Case Studies by Category

The case studies in this chapter illustrate how persistent narratives that predated COVID-19 vaccines were adapted to the pandemic response, and demonstrates how they formed, spread, and were framed both online and offline.

#### 3.2.1 Safety Narratives

**Death and Injury Narratives**

Some of the most consistently observed narratives during the Virality Project’s seven months of operation were stories of post-vaccination adverse events, used to promote the notion that vaccines are unsafe. These narratives included stories of purported vaccine injuries and severe side effects—ranging from rashes, to blood clots, to death—and primarily manifested in two forms: personal anecdotes and distortions of official government statistics.

Personal anecdotes often made their way into mainstream media coverage after gaining traction online. Distortions of official government statistics—most often from VAERS, described in more depth later in this section—were used both to reinforce the personal anecdotes and for focused misinformation solely discussing the statistics. Neither manifestation of these narratives was novel or unique to COVID-19 vaccines; anti-vaccine groups have long used isolated medical incidents and misrepresentations of VAERS data to bolster their claims of widespread vaccine harm.

Anti-vaccine groups on social media often stripped both individual stories and official statistics of important context in their efforts to undermine confidence in the safety of COVID-19 vaccines. Stories of individual tragedy had strong emotional appeal, while misrepresentations of official statistics helped create the impression that individual cases were part of a larger, more frightening pattern. Further, anti-vaccine groups could tailor who they targeted by emphasizing adverse events in
3. Narratives

particular sub-populations, like miscarriages in women or prominent-figure deaths in specific minority communities. And finally, while adverse events stories were most frequently used to raise doubts about the safety of vaccines, they could also be employed to push back against vaccine mandates or to support more conspiratorial claims.

**Individual Accounts of Adverse Events**

Anti-vaccine communities began to spread stories of post-vaccination deaths and side effects as soon as the distribution of COVID-19 vaccines began in December 2020. In the early months of the vaccine rollout, death and injury stories focused on what health officials called Tier 1-priority populations: older adults and medical professionals. When 23 older adults died in a nursing home in Norway in January 2021 after receiving the Pfizer vaccine, the news gained over 200,000 interactions across social media platforms, though Norwegian health officials found no immediate causal link between vaccinations and the deaths; a subsequent investigation published months later would surface a distribution of causes. However, as the news of the cases broke, Robert F. Kennedy Jr. (RFK Jr.) shared an article from his advocacy group about the nursing home deaths that stated that Chinese experts had suggested the vaccines needed to be halted. The incident illustrated two challenges that would recur throughout the pandemic. First, it takes time for health authorities to do a proper investigation into injury claims; the story will go viral on social media far faster than facts can be uncovered. Second, eliminating context about pre-existing conditions or circumstances, while emphasizing deaths, can create a false perception.

A similar story circulated in anti-vaccine groups alongside local news reporting of a grandmother from Virginia who died within hours of receiving the vaccine, leaving out that she had other underlying health conditions. Anti-vaccine influencers drew on these incidents to push the narrative that older adults were too frail for the vaccine, despite their known high risk of severe illness or death from COVID-19 infection (see Figure 3.1 on the next page). In early February 2021, for instance, an alternative health website combined these stories to suggest that COVID-19 vaccines were killing older adults; the Instagram post received over 8,000 interactions.

Stories of adverse reactions to COVID-19 vaccines varied significantly in their degree of verifiability. For example, videos spread on TikTok showed individuals shaking after allegedly receiving the vaccine. In many of these posts, users claimed that they, or someone they knew, experienced a symptom after receiving the vaccine (see Figure 3.2 on the facing page). In some cases, it was difficult or impossible for an outside observer to confirm whether the individual had received the vaccine or whether they truly experienced the symptom they claimed, much less to establish any link between the two events. Other stories contained outright false information.
3.2. Case Studies by Category

Figure 3.1: Left, Robert F. Kennedy Jr. tweets about the post-vaccination death of a grandmother, explicitly sowing distrust in public health institutions. Right, Alex Berenson tweets about the death of a 19-year-old following her second dose of the Moderna vaccine.\(^\text{12}\)

about purported reactions that was nonetheless amplified by anti-vaccine and other influencers.\(^\text{13}\)

Figure 3.2: Left, a TikTok video reshared on Instagram showing a man allegedly experiencing uncontrollable shaking after receiving a vaccine. Right, a post in a public Facebook Group dedicated to raising awareness about alleged vaccine-related injuries, asking people to share their experiences in an effort to lobby a state legislature about vaccine mandates.
3. Narratives

**How Stories of Adverse Reactions Spread: Platforms, Influencers, and the Mainstream Media**

Viral accounts of an individual’s death or injury often achieved significant reach via predictable pathways, leveraging bottom-up dynamics. Influencers with large followings, or long-standing anti-vaccine activist followings, such as RFK Jr. or Erin Elizabeth at Health Nut News, picked up on stories circulating on anti-vaccine websites or from small social media posters. Influencers often inserted their own tone and slant on the story as they promoted it, attracting more attention to the incident (see Figure 3.3). At this point, alternative and/or mainstream media outlets sometimes covered the claims, either credulously or in an attempt to debunk, ultimately allowing the story to reach larger audiences.

![Figure 3.3](image.png)

Figure 3.3: Left, Sara Gonzales of Blaze Media posts a screenshot of her tweet about the post-vaccination death of a 13-year-old boy to Instagram. Right, a Facebook Page posts a meme falsely implying a link between Hank Aaron’s vaccination and his death.

The story of a 13-year-old who died shortly after receiving his second shot illustrates how media and social media intersect in amplification. On June 20, 2021, the boy’s aunt tweeted a thread detailing his death, noting that he had no known medical conditions and calling on the CDC to investigate. While the aunt noted that she is pro-vaccine, and possibly did not intend for her nephew’s story to be used to undermine confidence in vaccine safety, anti-vaccine communities quickly picked up the story and did just that. Right-leaning outlets (like Fox News) and far-right media outlets (like InfoWars, National File, and Breitbart) additionally circulated the story, as well as centrist and left-leaning outlets like ABC, MSN, and Newsweek. The aunt deleted her original thread and stated her dislike of the media representation.
of her story, but the thread had already amassed over 50,000 likes, 35,000 retweets, and nearly 5,000 comments.\(^\text{16}\)

Individual anecdotes have also gained traction through the use of online fundraisers. In February 2021, a registered nurse from North Carolina shared a Facebook video of herself shaking, alleging that her tremors were a neurological effect of receiving the Moderna vaccine.\(^\text{17}\) The video received over 434,000 views. She also set up a GoFundMe campaign to cover her medical expenses and raised over $10,000 (see Figure 3.4).\(^\text{18}\) Others have followed suit; the anti-vaccine advocacy organization Children’s Health Defense reported in July 2021 that over 180 accounts had posted GoFundMe pages seeking help for medical bills and expenses from COVID-19 vaccine injuries.\(^\text{19}\) The personal anecdotes behind these campaigns are just as difficult to verify as those discussed above, but individuals have nonetheless successfully raised thousands of dollars by convincing others of their stories. In August 2021, GoFundMe refunded over $17,000 raised after it was discovered that influencer Dominique de Silva had falsely claimed to have suffered seizures after getting the vaccine and posted an online fundraiser to help cover medical expenses.\(^\text{20}\)

![GoFundMe campaign](https://via.placeholder.com/150)

**Figure 3.4:** A GoFundMe campaign for a nurse who allegedly experienced uncontrollable shaking after receiving the Moderna vaccine.

While most stories of post-vaccination deaths and injuries were largely confined to social media and had low spread, accounts of adverse events among celebrities appeared prone to go viral. This was likely in part because traditional media outlets covered these stories more often. In January, traditional media coverage of the death of retired baseball player Hank Aaron inadvertently fueled the spread of vaccine misinformation and vaccine hesitancy. Aaron died 18 days after receiving the first dose of the Moderna vaccine. His vaccination had been highly publicized, as he hoped it would inspire other Black Americans to do the same.\(^\text{21}\) Media outlets largely did not specify the cause of Aaron’s death in their reporting,\(^\text{22}\) anti-vaccine influencers, meanwhile, including Children’s Health Defense and RFK Jr., took to social media to claim that Aaron’s death was linked to the vaccine. RFK Jr.’s Face-
book and Twitter posts garnered 12,000 and 20,000 interactions respectively, with the caption “Hank Aaron’s tragic death is part of a wave of suspicious deaths among elderly closely following administration of COVID vaccines.” By comparison, subsequent fact-checks by the New York Times and USA Today clarifying that Aaron died of natural causes received only 1,200 interactions on Facebook combined.

### Misrepresentation of Government Statistics

The second major form of death and injury narratives were those misrepresenting statistics on post-vaccination adverse events, most often from VAERS. Notably, anti-vaccine influencers also misleadingly referenced other injury databases outside the United States, such as the UK’s Yellow Card reporting site, EudraVigilance (used in the European Economic Area), and the World Health Organization’s VigiAccess. VAERS is a passive reporting surveillance system, first launched in 1990 and run by the CDC and FDA. The database is meant to serve as an early warning system for detecting possible safety problems in vaccines, but it allows anyone to submit a report of a post-vaccination adverse event, regardless of whether there is a causal link between vaccination and the event. The VAERS website explains that reports may include “incomplete, inaccurate, coincidental and unverified information” and that it should not be used alone “to reach conclusions about the existence, severity, frequency, or rates of problems associated with vaccines.” Nevertheless, anti-vaccine communities have repeatedly cited VAERS data, for many years, as evidence of verified, widespread vaccine-related harms. Beginning in January 2021, this attention largely shifted to COVID-19 vaccines: Children’s Health Defense began publishing near weekly updates on the number of reports submitted to VAERS, which RFK Jr. promoted across his social media accounts. Other accounts that appeared in numerous Virality Project tickets, such as Alex Berenson and Naomi Wolf, routinely shared commentary about VAERS numbers. Another recurring actor was Michael Yeadon, an ex-Pfizer vice president who appeared in multiple VP tickets and briefings. In a June 2021 interview, Yeadon cited VAERS data as evidence that young people were more likely to die from the vaccine than from COVID-19. The interview garnered nearly 350,000 views on Rumble, a video platform that has often been used to spread conspiracy theories or host content that YouTube and other platforms have moderated in some way.

The tactic was used offline as well. In May 2021, Tucker Carlson misrepresented VAERS data on his talk show, decontextualizing it while claiming that 3,362 Americans had died following COVID-19 vaccinations between December 2020 and April 2021, equating to roughly 30 people every day. A few days later, Senator Ron Johnson (R-WI) repeated the claim on Carlson’s show when explaining why he would not get vaccinated. The claims additionally appeared to influence political action offline: In June 2021, Sen. Johnson hosted a live event in Wisconsin to give a platform to individuals who had experienced negative reactions following their
3.2. Case Studies by Category

The event was held the same day that Senators Johnson and Mike Lee (R-UT) formally submitted a letter of inquiry to the CDC and FDA asking for specific details about adverse reactions that occurred within the AstraZeneca vaccine trials. Johnson has repeatedly defended these actions as “asking questions.”

Anti-vaccine activists also developed wrapper sites, which repackaged VAERS data to make it cleaner and more readable for the casual observer than the original VAERS site. This strategy was not new for COVID-19; one such wrapper site, MedAlerts.org, run by the anti-vaccine advocacy group National Vaccine Information Center (NVIC), has existed since at least 2007. In January 2021, however, an Oakland, CA-based woman created a wrapper site, OpenVAERS, that quickly became more popular on social media than the original site. This wrapper site includes a “red box summary” page specifically dedicated to sharing COVID-19 vaccine statistics from VAERS and highlighting particularly troubling reports. By repackaging these statistics and making it easier for users to search reports for specific symptoms, OpenVAERS became a powerful tool for spreading misleading claims using VAERS data. VP analysts repeatedly observed screenshots and other references to OpenVAERS across social media platforms (see Figure 3.5). According to SimilarWeb, more than 1,230,000 people visited OpenVAERS between February and July 2021, compared to 796,300 visits to the official government VAERS website.

The tactical appeal of frightening statistics creates an incentive for anti-vaccine activists to drive up the number of VAERS reports. The numbers presented by both

---

There are limitations to this analysis. Although the higher number of views on OpenVAERS compared to VAERS is consequential, anti-vaccine groups may not be entirely responsible for that traffic flow. OpenVAERS may simply be a more accessible website than VAERS. Further, the traffic data cited here is an estimate and is drawn from the market analysis tool SimilarWeb, which does not open-source its data.
3. Narratives

MedAlerts and OpenVAERS, for example, include foreign reports to VAERS, even though it is a US database and its own downloads exclude these reports by default. For example, the red box summary from OpenVAERS shows that there were 12,366 reported deaths after COVID-19 vaccinations through July 30, 2021, but this is over double the number of reported deaths displayed by default on the CDC WONDER tool for viewing VAERS data.\textsuperscript{40} VP analysts also saw anti-vaccine activists share calls to contribute reports to VAERS, as well as paid advertisements on Facebook that linked directly to the VAERS submission form.\textsuperscript{41}

3.2.2 Distribution Narratives

Government Overreach and Medical Freedom Narratives

One of the primary long-standing themes of anti-vaccine distribution narratives is that mass vaccine distribution constitutes a government overreach. The movement sees vaccine mandates, including, historically, school vaccine requirements, as an assault on “health freedom” or “medical freedom.”\textsuperscript{42}

Dr. Peter Hotez summarized the evolution of the narrative for COVID-19 in the \textit{Journal of Clinical Investigation} in February 2021:

\textit{“Health freedom, also referred to as medical freedom, represents a long-standing and pervasive belief system in American life. Its tenets espouse an aversion to government interference in personal or family health choices, often coupled to the counter promotion of a spectacular or miracle cure. The concept is almost as old as the nation itself, although the newest iteration of health freedom defies vaccinations and federal or state public health interventions directed at COVID-19. ... In 2020, following the emergence of COVID-19, these same health freedom groups expanded their vaccine protests to social distancing, masks, and other prevention measures.”}\textsuperscript{43}

Medical freedom groups and networks were already present (both online and off) in states across the US prior to the pandemic. Many had highly organized online and offline networks, having previously been involved in fights against state-level school vaccine requirements, and they pivoted arguments developed for other vaccines to the COVID-19 vaccine rollout. Other groups emerged on platforms such as Facebook during the pandemic, with names specifically related to COVID-19 or mRNA vaccines, to assist in discoverability; some grew their numbers into the tens of thousands.

During the VP’s period of analysis, narratives about government overreach and medical freedom focused on two areas of controversy: vaccine passports and vaccine mandates. Here, the term “passport” is used to denote any form of digital or physical proof that a person is protected from COVID-19 via vaccine, while “mandate” is used to refer to compulsory vaccination by a school, local government, or employer.
Both passport and mandate debates involve nuanced and complicated questions about people’s individual autonomy and community values. However, the strong opinions and emotional responses these questions raise enable anti-vaccine activists and political influencers to exploit legitimate public health debates—about whether, when, and how to leverage the government's authority to mandate interventions in service to public health—to spread distrust in the institutions and vaccines themselves. As we describe below, anti-vaccine activists have also co-opted historical atrocities perpetrated against marginalized groups, including BIPOC, to make false and harmful claims about the intentions behind and effects of vaccines, passports, and mandates.

**Vaccine Passports**

Discourse around vaccine passports first picked up speed in February 2021, when media in the UK, Israel, and various European countries began to cover possible government plans to digitally certify citizens’ vaccination status. Russian state media outlet RT publicized these debates through a series of Facebook posts that called passports into question, provoking concerns about overreach with headlines like, "Vaccine passports: a route out of lockdown? Or an unnecessary overreach on citizens?" This amplification hinged upon misleading framing that suggested the implementation of any form of vaccine passport would be compulsory. In reality, the plans for many programs were entirely optional. Other framing from domestic right-leaning political actors created a portrait of governments as prying or snooping into citizens’ private matters.

The key aspects of the discourse in both right-leaning political circles and foreign media were not new; they relied on older anti-vaccine tropes and ideas about medical freedom. Activists pushed the idea that through a passport system, governments and "Big Tech" were limiting the public’s freedoms—situating the conversation within a larger set of narratives surrounding pandemic public health regulations like mask mandates, lockdowns, and social distancing. Requiring proof of vaccination status was also framed as a way to systematically disadvantage vulnerable populations. As with the mandate conversation, the passport conversation leveraged real questions about privacy, access, and inequities—but it was also employed to exacerbate vaccine hesitancy and rejection.

Right-leaning influencers and outlets on social media regularly created, and amplified, government overreach and health freedom narratives. In the winter of 2021, commentators at Breitbart focused on the potential threat to individual rights (see example in Figure 3.6 on the following page), while former One America News Network commentator Liz Wheeler called passports an "abuse of power." Headlines sometimes hawked conspiracy theories: one Gateway Pundit headline, “The Great Reset: Big Tech and Big Pharma Join Forces to Build Digital COVID Vaccination Passport,” was a nod to groups such as QAnon, who use “the Great Reset” to refer
to a conspiracy theory that global elites are using the pandemic to take control of the world’s economy and destroy individual freedom. The article alleged collusion between Big Tech and Big Pharma that would threaten “individual rights.”

Figure 3.6: Breitbart headline from January 15, 2021.

Widespread acceptance of this narrative appeared to figure prominently in shaping offline legislation, and several states imposed bans on vaccine passports. The first such vaccine passport ban was issued by Florida Governor Ron DeSantis on April 2, 2021. Coverage of the ban gained over 1,000,000 engagements across social media platforms. After DeSantis’s executive order, several other states followed suit. As of August 13, 2021, 20 states had prohibited proof-of-vaccination requirements; each of these states had a Republican governor. On May 28, 2021, Senator Ted Cruz (R-TX) and a group of Republican senators introduced the first bill at the federal level to ban vaccine passports. As of February 23, 2022, the bill remains in committee.

**Vaccine Mandates**

The backlash to COVID-19 vaccine requirements for employment and other activities parallels the conversation about vaccine passports. It, too, relies on and attempts to exacerbate distrust in public health officials and government institutions.

Two incidents illustrate this dynamic. In July 2021, White House Press Secretary Jen Psaki announced that the Biden administration would begin focusing on door-to-door community-based vaccine outreach. This announcement provoked immediate online outrage, including worrisome jokes with violent undertones that suggested people should respond to official door-knockers with hostility, guns,
or harassment (see Figure 3.7). Conservative commentators also lambasted the outreach program, and *The Alex Jones Show* referred to the push as a government “strike force” sent in to forcibly vaccinate children.

![Figure 3.7: Two interpretations of door-to-door outreach depicted on Twitter.](image)

The account on the left has since been suspended.

The second incident concerned vaccination for children. In May 2021, the FDA approved an emergency use authorization of COVID-19 vaccines for children 12 years old and older. Most states require parental consent for minors to receive the shot, but not all; anti-vaccine activists framed the issue as government overreach into a parent’s right to give consent to their child’s medical treatment. Tennessee’s Department of Health became the focus of outrage after the state’s medical director for vaccine-preventable disease and immunization programs, Dr. Michelle Fiscus, led an outreach program that sent public service advisories to teenagers informing them of their eligibility to receive a COVID-19 vaccine. Major news sources and fact-checkers, including misinformation-fighting organization First Draft, reported that both lawmakers and online anti-vaccine communities used the heated situation to spread misinformation about the safety of vaccines for children. Dr. Fiscus was fired in July 2021 after a meeting in which lawmakers stated that the vaccine’s ingredients and long-term effects are still unknown. Other lawmakers repeated narratives about government overreach, medical freedom, and parental consent. For example, throughout the controversy, Republican legislators said Dr. Fiscus’s work was part of an “agenda” to have children vaccinated “with or without parental consent.”

The situation in Tennessee showcases the complicated dynamic between online narratives and offline impacts: anti-vaccine rhetoric fueled politicized conflict in Tennessee and ultimately led to the removal of a qualified health official.

**Weaponization of Historical Analogies**

Another type of argument that Virality Project analysts observed was the leveraging of historical atrocities in false equivalencies about vaccination requirements. Anti-
vaccine activists have long referenced the Holocaust and the Nuremberg Code (a set of research ethics principles for human experimentation) in opposing immunization requirements for schools and claiming Nazi-like discrimination against unvaccinated people. This is a long-standing anti-vaccine motif: in 2015, Dr. Bob Sears, who had been in opposition to a California bill that would remove certain exemptions to school immunization requirements, used a Facebook post to compare persecution of non-vaccinating parents to the persecution of Jews during the Holocaust. RFK Jr. was also criticized that year for calling vaccination a “holocaust.”

Anti-vaccine activists applied these comparisons and references to COVID-19 vaccines. In March 2021, a tweet thread that garnered 13,000 likes and 7,000 shares on Twitter falsely claimed that Pfizer was found guilty of breaking the Nuremberg Code for unlawful experimentation. Other more recent posts on Instagram and Facebook claimed the vaccine rollout would result in a “Nuremberg 2.0”: a trial that activists hope will take place to prosecute those involved in the COVID-19 response (see Figure 3.8).

Figure 3.8: Conspiracy theory Instagram post about “Nuremberg 2.0” from March 9, 2021.

Similar claims about discrimination were made leveraging a wide range of other historical allusions:

- Michael Yeadon, the ex-Pfizer vice president who repeatedly promoted anti-vaccine tropes, called vaccine passports a “totalitarian” framework for mass deportation.

- Naomi Wolf, a major recurring actor of vaccine misinformation, likened passports to slavery.

- Russian state media compared vaccine cards to “Covid apartheid.”

These rhetorical analogies prey on the public's memories of actual government atrocities, but are not necessarily directed toward the marginalized populations.
who were once harmed by those discriminatory policies or events. This makes the parallels especially offensive to the communities whose histories are exploited for these narratives; the Anti-Defamation League, among other Jewish organizations, has denounced these comparisons as “odious.”

In addition to appropriating specific historical atrocities and applying them to unrelated communities, anti-vaccine activists have also separately weaponized marginalized groups’ preexisting distrust of government to foster vaccine hesitancy specifically among those populations. Although many of the well-known anti-vaccine activists are white, they have often turned to racialized communities, especially Black communities, as the targets of their campaigns, capitalizing on deep-seated concerns about vaccine development and distribution as well as legacies of medical racism in the US.

Some of the most significant promoters of these narratives have been Children’s Health Defense (CHD) and its chairman RFK Jr., who has helped promote the narrative that vaccines are dangerous for Black people through his use of various social media platforms, particularly Twitter. The narrative also gained mainstream traction after CHD announced the premier of its documentary Medical Racism: The New Apartheid. The documentary framed pharmaceutical companies and the World Health Organization as actors in a neocolonial agenda to “enslave” Africans through drug testing, citing controversial claims made about the Bill & Melinda Gates Foundation’s vaccine programs and the history of the slave trade in Africa. The film also amplified a racist stereotype that Black people possess a more robust immune system that produces more antibodies and eliminates the need for vaccines. Although many of the claims made in the documentary were false, the film was legitimized via promotion by the Nation of Islam (NOI), a Black nationalist religious and political organization founded in 1930. Tony Muhammad, an NOI minister who has in the past claimed that childhood vaccines are “genetically modified” to harm children of color, was credited as a producer on the film. In the video, Muhammad makes a parallel between COVID-19 vaccine trials and the 1932–1972 Tuskegee Syphilis Study, an abusive and unethical study that denied lifesaving treatment to hundreds of study subjects, all of whom were Black men.

The documentary’s announcement and trailer did not go nearly as viral as other anti-vaccine content about COVID-19, but still garnered 43,000 views on Instagram and 2,300 engagements on Facebook, where the majority of viewers voiced support for

---

3. Narratives

the film. The film then gained more traction on Telegram channels; the trailer was promoted in six Telegram channels that shared the content with 60,000 members, generating 5,000 views collectively.

Throughout the rollout, VP analysts repeatedly reviewed content that referenced the Tuskegee study. In the month of July 2021 alone, VP analysts observed over 200 posts featuring narratives involving Tuskegee. The legacy of the study was consistently used as evidence to undermine support for vaccines: users claimed that past unethical practices would be repeated in the COVID-19 vaccine rollout, or that the history of Tuskegee reflected establishment public health institutions’ inherent untrustworthiness. These claims sometimes included slightly more conspiratorial speculation about what the vaccine was “really” attempting to do to Black people—from sterilization to infecting them with COVID-19. Others simply speculated that the vaccine wasn’t designed to “work” in Black people.

These claims exploit the existing distrust that marginalized communities often hold for medical and scientific professionals. This pattern is likely to continue as long as neglect and harm in healthcare systems remain common in certain communities. Developing techniques to counter these narratives from a place of deep understanding of historical context must be a focal point for healthcare authorities.

3.2.3 Conspiracy Theories

Conspiracy theories have played an ongoing role in delegitimizing trust in vaccines and public health institutions. Conspiracy theory narratives often leverage persistent anti-vaccine tropes and themes drawn from other categories, such as fears that the vaccines may contain dangerous ingredients, but additionally attempt to assign a motive or blame for the situation to powerful people. Common examples included allegations that cabal-like entities and “higher powers” were controlling the pandemic, or controlling the population through the vaccine rollout. The three case studies below highlight this dynamic.

Mark of the Beast

In the last two decades, the anti-vaccine movement has been defined by questioning the safety of vaccines; religious-based fears have mainly taken a back seat to safety concerns. However, VP analysts saw that the movement against the COVID-19 vaccine has reincorporated religious narratives. One of the most consistent pandemic conspiracy theories relates to the “Mark of the Beast” in reference to Revelation 13:16–17, which describes a “beast” that forces people to receive a mark from the Antichrist and says that people will be unable to “buy or sell unless they had the mark.” Some evangelical Christians believe the COVID-19 vaccine is the prophesied Mark of the Beast. This wide-ranging theory has been used to undermine
3.2. Case Studies by Category

any widely advised COVID-19 mitigation measure, including masks, vaccines, and vaccine passports.80

Political leaders and celebrities have shared content suggesting that the Mark of the Beast is connected to COVID-19 vaccines and vaccine passports. In July 2020, rapper and public figure Kanye West told Forbes magazine in an interview that the COVID-19 vaccine—still in development stages—would be the Mark of the Beast.81 In March 2021 the narrative gained over 200,000 interactions on Facebook in under a week after Congresswoman Marjorie Taylor Greene (R-GA) stated in a livestream video shared across Facebook and Twitter that vaccine passports could be considered the Biden administration’s Mark of the Beast (see Figure 3.9).82

Despite drawing on conspiracy theory framings depicting vaccines as being controlled by a powerful figure, the Mark of the Beast narrative falls within a challenging area for platform enforcement:83 platforms struggle to balance freedom of expression against the leveraging of religious claims as tools to promote vaccine hesitancy and spread vaccine mis- and disinformation.

**Reproductive “Shedding”**

In early 2021, users on Twitter, Facebook, and Reddit reported unverified reproductive side effects, ranging from abnormal menstrual cycles to miscarriages and infertility. Anti-vaccine activists, health and wellness influencers, and pseudomedical influencers alleged that these events were connected to the COVID-19 vaccines and leveraged them to increase vaccine hesitancy and rejection: using misleading VAERS data to advise pregnant women against the vaccine, exploiting a doctor’s miscarriage as “proof” of the vaccine’s danger, and spreading stories of people losing their jobs due to fertility health–related vaccine refusal.84 At the time there was no medical consensus on the vaccine’s effect on reproductive health,85 yet anti-vaccine activists presented the theory as fact and evidence of harm. Research is ongoing, but recent studies have found only small, temporary menstrual cycle changes after
3. Narratives

In April 2021, however, VP analysts identified a new form of reproductive health claim that leveraged a long-standing anti-vaccine trope, “vaccine shedding,” but with a new angle. Vaccine shedding—the potential for a person recently vaccinated with a live virus to “shed” the virus and expose or infect another person—is a real but exceedingly rare phenomenon, only possible with live attenuated vaccines such as the MMR vaccine. Prior to the COVID-19 vaccines, anti-vaccine activists had used shedding narratives to claim that the vaccine itself could spread disease—and, therefore, that vaccines could in fact be responsible for outbreaks. In the new claims, activists claimed not only that shedding was occurring with COVID-19 vaccines (a medical impossibility), but that vaccine shedding could spread harmful side effects, rather than the virus itself.

These claims coalesced into a new narrative combining older, separate narratives about vaccine shedding and reproductive health. On April 15, 2021, the conspiracy theory website InfoWars posted an article claiming it had access to “hundreds of testimonials” from women who reported “bizarre menstrual issues and even miscarriages after being close to [vaccinated] people.” The narrative was also circulated on April 20, 2021, by Naomi Wolf, who tweeted that unvaccinated women were experiencing menstruation changes after being around vaccinated women. Influencers used this conversation to further exploit people’s safety concerns, including about the vaccine’s “spike protein” (mRNA vaccines direct cells to make a harmless part of the coronavirus spike protein). Anti-vaccine activists had previously spread false information that the vaccines contain actual spike proteins and that spike proteins would cause sterilization in women. Here, they took the claims further, suggesting that spike proteins could be shed to unvaccinated people simply via respiration.

The new claims appeared within a broad range of Page and group types. Traditional anti-vaccine communities, where some of these stories originated, were familiar with the more general shedding trope, but shedding was less well known within the conspiracy theory/QAnon and health and wellness groups. The health communities, who are often distrustful of pharmaceutical companies and the mainstream healthcare system, were especially influential in the growth of the narrative. Sharing the claims was often framed as a measure to “protect women” who were depicted as a vulnerable group, which possibly enabled the narrative to attract a larger audience.

None of the shedding claims were true. As the shedding conspiracy theory gained traction on mainstream platforms, traditional media and fact-checking organizations—including The New York Times and NBC News—quickly moved to cover it. Engagement with these fact-checking posts soon dwarfed engagement with misinformation posts, and by May 2021, the narrative had lost traction.

See Appendix B for a definition of “pseudomedical” as used in this report.
3.3. Conclusion

Magnet Challenge

On May 9, 2021, a TikTok video went viral that purported to show a magnet sticking to someone’s arm over their vaccine injection site, implying that the vaccine contains unsafe ingredients that make people magnetic. (There are a number of ways small metallic objects may stick to someone’s skin other than magnetism, including tape, lotion, or body oils.)\(^98\) In just four days, before TikTok removed the video, it received over 9,600,000 views and inspired the creation of hundreds of copycat videos across all major platforms, challenging others to put a magnet on their bodies or over their vaccine injection site. By June 21, 2021, nearly 7,000 tweets with the hashtag #magnetchallenge had been posted.\(^9\) Alarmingy, even videos that appeared to be created satirically were taken seriously, indicated by top-liked comments expressing outrage. On June 8, anti-vaccine influencers, including pseudomedical influencer Dr. Sherri Tenpenny, used this false phenomenon as evidence during hearings in the Ohio state legislature for a house bill on “vaccine discrimination.”\(^99\)

The Magnet Challenge illustrates how novel online tactics like participatory online video “challenges” can be combined with ongoing anti-vaccine tropes about strange reactions and dangerous ingredients to successfully spread misinformation online. Although many creators of these videos seemed to be participating in the trend in jest, the magnetism conspiracy theory was still able to gain traction, likely because it built on ongoing anti-vaccine tropes and prior widespread concerns about microchips in the vaccines.\(^100\) Prior to COVID-19 vaccines, other vaccines have been the subject of toxicity fears and anti-vaccine rhetoric about ingredients such as thimerosal.\(^101\) Thus, seemingly outlandish claims like the Magnet Challenge perhaps resonated within communities that have ongoing fears about what people are putting into their bodies.

3.3 Conclusion

This discussion of prominent anti-vaccine narratives is not meant to be comprehensive. Instead, it offers significant case studies to help readers understand that narratives used to spread and harness COVID-19 vaccine hesitancy, concerns, and misinformation have been both persistent and adaptable throughout the rollout. These narratives are not new. Rather, they hinge on repeated storylines and conspiratorial plots that have long been popular within the anti-vaccine community, adapted to current events or circumstances.

\(^9\)This number was recorded on June 21, 2021, and does not include content that may have been removed prior to that date, including the original viral TikTok video. Chase Small et al., “Made to Stick: Origins and Spread of the Magnetic Vaccine Narrative,” *The Virality Project*, June 4, 2021, https://www.viralityproject.org/rapid-response/made-to-stick-origins-and-spread-of-the-magnetic-vaccine-narrative
3. Narratives

It is precisely this remixing—incorporating a novel angle into otherwise familiar, recognizable tropes—that makes misinformation narratives on social media so difficult to combat. Frequently, the substance of problematic claims is an amalgamation of many bits and pieces of narratives and ideas from across multiple themes. They may draw on a prominent recurring cast of archetypes (e.g., the evil billionaire) within the conspiracy theory spectrum. Often a piece of content will contain three to five different issues rolled into one, making it hard to disentangle the discrete problems or narratives into neat, fact-checkable boxes. As the pandemic and vaccine rollout wore on, this challenge became more prevalent, with a new idea layered on top of existing ones.

The familiarity and repetition of tropes, characters, and narratives may indeed suggest a path for intervention. The Virality Project believes that a deeper understanding of the key narratives outlined here can better facilitate a collective approach to responding to anti-vaccine mis- and disinformation. Prevalence of narratives might indicate where government and public health institutions should focus on building trust in communities in which mass vaccination has proven difficult, controversial, or seemingly impossible. Media can anticipate the transformation or re-uptake of old narratives ahead of the likely emergence of new variants, both in the US and elsewhere, drawing a clearer throughline highlighting recurring tropes. The persistence of key narratives around vaccine safety, efficacy, and development provides opportunities for stakeholders in various sectors to proactively detect, debunk, and even pre-bunk emerging misinformation threats.\(^\text{102}\)
Notes


9. Robert F. Kennedy Jr (@RobertKennedyJr), “#Norway upped the number of deaths under investigation, from 23 last week to 33, while in #Germany, health officials said they are investigating 10 deaths that occurred among elderly patients who received the #COVID #vaccine. #TheDefender https://t.co/YSFxxyIjcK,” Twitter, January 18, 2021, https://web.archive.org/web/20210425131919/https://twitter.com/RobertKennedyJr/status/135127793607948800; Reuters, “Fact Check: Discussion of Deaths in Elderly Vaccine Recipients in Norway Lacks Context.”


13. See, e.g., Alex Berenson (@AlexBerenson), “And there it is: although I must warn I don’t have firsthand knowledge of this interview - can someone provide a direct link if possible?” https://t.co/XaM6szSyV,” Twitter, June 12, 2021, https://web.archive.org/web/20210619061316/https://twitter.com/alexberenson/status/1403800772345745411; Sean Carlin, “Soccer Star’s Collapse Was Unrelated to COVID-19 Vaccine,” FactCheck.Org, June 21, 2021, https://www.factcheck.org/2021/06/scicheck-soccer-stars-collapse-was-unrelated-to-covid-19-vaccine/; Alex Berenson (@AlexBerenson), “Didja ever notice that more than two days after his cardiac arrest, neither Christian Eriksen nor his medical team has said ANYTHING about his vaccination status? all we have are conflicting comments from the team he was NOT playing from at the time,” Twitter, June 14, 2021, https://web.archive.org/web/20210614172542/https://twitter.com/alexberenson/status/140487245608873985.

14. Tami Burages (@tburages), “A week ago today my brother’s 13-year-old son had his 2nd covid shot. Less than 3 days later he died. The initial autopsy results (done Friday) were that his heart was enlarged and there was some fluid surrounding it. He had no known health problems. Was on no medications,” Twitter, June 20, 2021, https://web.archive.org/web/20210710054935/https://twitter.com/tburages/status/1406796255573991424.


16. Tami Burages (@tburages), “I have deleted a tweet thread about the sudden death of my nephew following his 2nd covid shot. Fox News and other far-right disinformation networks are politicizing the death and using my tweet to do it. They are causing my family more pain. If you have any questions, DM me,” Twitter, June 22, 2021, https://web.archive.org/web/202107122233418/https://twitter.com/tburages/status/1407343080026611713.


41. See, e.g., Real Immunity (@realimmunitymovie), “UPDATE- For anyone who has tried to navigate the government VAERS (Vaccine Adverse Event Reporting System) site and has thrown up their hands in frustration, here is an easy way to access an updated summary,” Facebook, July 9, 2021, https://www.facebook.com/695769770631109/posts/1679826732225403.


59. Jacqueline Howard and Virginia Langmaid, “Most states require parental consent for Covid-19 shots for ages 12 to 15, but not all. Here’s the list,” CNN, May 18, 2021, https:


76. See, e.g., Freedom From Mental Slavery, “Tuskegee Study: Black people thought they were receiving free healthcare but instead were injected with Syphilis. This was conducted by the U.S. Public Health Service and happened over a period of 40 years between 1932 and 1972. Let us never forget,” Facebook, July 25, 2021, https://www.facebook.com/161399403976816/posts/4137971696319547; Ishmael Mayhew (@ishcontent), “What part did we misunderstand about it? The Tuskegee Experiment went on for 40 years. Define not that long ago? It sounds like gaslighting to me. This was REAL and it impacted generations. We not about to skim over that. You shouldnt need to do that to defend the vaccine,” Twitter, July 27, 2021, https://web.archive.org/web/20210727195548/https://twitter.com/ishcontent/status/1420100592131887115; KING KY (@DaddyKy_), “Yeaaa it’s time to get @POTUS on up out of the house. You can’t force ppl to get a vaccine. We remember the Tuskegee Syphilis Study sir.. Why they hell should we trust what the government say when all y’all ever done was let us down?? Now you want us to have faith in y’all? 😅,” Twitter, July 27, 2021, https://web.archive.org/web/20210831050326/https://twitter.com/ThatGuyOGPrime/status/1420042590821527552; The Chad (@theChadFactor), "Daily reminder I don’t care if half the world is vaccinated I’m not taking the vaccine. I don’t even take flu shots. A lot of you people need to research Tuskegee seriously," Twitter, July 27, 2021, https://web.archive.org/web/20210727114297/https://twitter.com/TheChadFactor/status/1420130427709366288.


78. Rv. 13:16–17 NIV.


This chapter discusses notable actors and tactics that repeatedly appeared in incidents of false and misleading claims in the Virality Project’s observation of the COVID-19 vaccine conversation. These short descriptions shed light on some of the dynamics leveraged to produce, spread, and sustain vaccine mis- and disinformation on social media from February to August 2021.

As described in Chapter 1: Introduction, the VP consortium was informed by Tara Smith’s taxonomy of particular types of actors who have long been active in vaccine conversations. At the conclusion of the Virality Project’s observation period, as we reviewed the corpus of incidents in the reports database and weekly briefings, some additional actor types emerged, which we discuss in this chapter.

While no perfect system of identification or classification exists—categories are porous, and some actors, including high-profile ones, did not squarely fit into any one category—examining recurring dynamics or prevailing narratives with a lens focused on the actor type involved (rather than substance of the claims) provides a glimpse into how different types of influencers worked to appeal to their audiences.

### 4.1 A Cast of Recurring Actors

**Long-standing anti-vaccine influencers:** These actors have a long-term primary focus on vaccines and have been building anti-vaccine audiences and content repositories for years prior to the COVID-19 vaccines. Many can be classified not only as influencers but also as activists, given their involvement in political organizing and legislative efforts.

**Wellness and lifestyle influencers:** This group of holistic health and wellness advocates often express distrust of traditional medicine. It includes parenting-
advice bloggers, yoga instructors, dieticians, health coaches, “natural” supplement purveyors, and general new-age consultants.

**Pseudomedical influencers:** These individuals have scientific or medical backgrounds, and leverage their actual or perceived credentials and expertise while undermining vaccines.

**Conspiracy theory influencers:** These individuals have built their following by discussing and amplifying a range of conspiracy theories, some related to pseudoscience, others related to government control; they frame COVID-19 vaccines as part of a larger nefarious agenda by an overarching, shadowy enemy.

**Right-leaning political influencers:** This group encompasses both mainstream and alternative right-leaning media figures, politicians, and publications.

**Medical freedom influencers:** Actors in this group are averse to government interference in individuals’ personal lives. While they explicitly advocate for “health freedom” or “vaccine choice,” these actors often propagate vaccine doubt by contextualizing the choice with misleading claims of vaccines’ adverse medical consequences.

### 4.1.1 Long-Standing Anti-Vaccine Influencers

Long-standing anti-vaccine influencers (LAVIs) are those individuals whose primary content output is in opposition to vaccination and who have been prominent parties in the vaccine conversation prior to the emergence of COVID-19. They have developed their audiences and refined their anti-vaccine messages for years. Before turning their attention to the COVID-19 vaccines, they primarily focused on childhood immunizations. LAVIs had heavy overlap with the “Recurring Actors” discussed in Chapter 2: Data Collection and Statistics, who generated a large percentage of observed incidents, though some recurring actors belong to other categories. LAVIs’ posts frequently garnered high levels of engagement.

Within this group, the actor associated with the highest number of incidents (specific events or stories reported on a ticket) was Robert F. Kennedy Jr. (RFK Jr.)—though others, including Del Bigtree and Sherri Tenpenny, appeared in incidents across the entire duration of VP’s research. In her 2017 paper, Tara Smith identified RFK Jr. as a movement Organizer;^2^ over a period of years, he has conducted much of his online activism and built his audience through the use of mainstream social

---

^1^See Appendix B: Definitions for the full definition and explanation of our use in this report of the term “pseudomedical,” derived from the concept of “pseudoscience.”

^2^While there was a research pod dedicated to looking at left-leaning political influencers, this chapter focuses on entities that appeared in multiple tickets. Left-leaning political influencers did not arise in a threshold number of incidents to merit inclusion in this chapter.
media platforms such as Facebook and Twitter, as well as his own website and accompanying email list. He had for many years additionally incorporated Instagram into a cross-platform communication strategy, but lost his approximately 799,000-follower Instagram account in February 2021 as the platform began to crack down on anti-vaccine misinformation. Del Bigtree’s Instagram account, with 162,000 followers, was also removed.iii

That deplatforming did not lead to the silencing of their voices on social media, however. Prominent and long-standing anti-vaccine activists and organizers often establish nonprofits, media organizations, or other entities through which they can further legitimize their operations; they also use these organizations as a vehicle for fundraising and donations. RFK Jr. established the nonprofit Children’s Health Defense (CHD) to gain legitimacy;3 Bigtree created Informed Consent Action Network (iCAN) as well as his (now deplatformed) video program The HighWire.4 For a time, this split-entity strategy created redundancy: if an individual activist or entity page was actioned by platforms for violating health misinformation policies, the affiliated entity was able to continue posting.

RFK Jr. and CHD appeared in the VP’s list of actors associated with the top number of incidents, at fifth place and second place, respectively. RFK Jr.’s anti-vaccine posts were also repeatedly among the highest performing of the week by engagement. On Instagram, Facebook, and Twitter, he amplified other long-standing anti-vaccine influencers; drew attention to both individual anecdotes and broader statistical misinformation about vaccine injury (such as misinterpretations of VAERS); called vaccine trials into question; spread conspiracy theories about Bill Gates, Anthony Fauci, and others; and produced a highly slanted documentary on medical racism. His posts regularly garnered over 1,000 interactions and were widely shared across the network of LAVIs. He used his online presence to direct users to buy and leave reviews of his book The Real Anthony Fauci: Bill Gates, Big Pharma, and the War on Democracy and Public Health, which in mid-November 2021 hit number one on Amazon’s bestseller list.5 This harnessing of his long-standing followers further boosts discoverability through search result placement, increases positive reviews of the book, and broadens his audience.

RFK Jr.’s activism is especially pernicious because, like other long-standing influencers, he has a large and committed following and has become somewhat of a household name in the US. He is thus able to easily bring on board other influencers to collaborate toward larger aims offline, including in the legal space. In a viral incident on May 16, 2021, RFK Jr. co-filed a citizen petition on behalf of CHD to ask the Food and Drug Administration to immediately revoke the Emergency Use Authorizations (EUAs) for all COVID-19 vaccines.6 The petition also called for the FDA to ban all under-18 participants in clinical trials for COVID-19 vaccines and for EUAs permitting the vaccination of children under 18 to be revoked. At the time,

---

iiiSee Chapter 5: Platform Policies for more information on specific changes to the policies.
influencers in other anti-vaccine communities, like the health freedom-focused Millions Against Medical Mandates, vocally supported the petition.7

RFK Jr.’s amplification of the CHD article about the Citizen Petition garnered 2,700 engagements; the petition itself garnered over 30,000 public comments. Yet this case is only one example of this type of activity from the organization. Individually, this kind of legal case may not have much impact. However, over time, the accumulation of these types of suits by anti-vaccine organizations—including one filed by America’s Frontline Doctors against the US Department of Health and Human Services similarly attempting to invalidate the vaccines’ EUAs8—gives the actions of anti-vaccine groups an appearance of credibility, allowing them to use these legal filings as evidence of the legitimacy of their claims.9

4.1.2 Wellness & Lifestyle Influencers

Wellness and Lifestyle Influencers (W&LIs) focus on alternative health and well-being content such as holistic health and wellness, parenting advice, and “clean” diets and food. These accounts have high overlap with Tara Smith’s “Mommy Bloggers.”10 They describe themselves in their profiles as yoga instructors, dieticians, health coaches, and general new-age consultants. During the observation period, VP analysts found that misinformation coming from this community most commonly appeared and spread on Instagram and in Facebook Groups, where influencers are often savvy and experienced in evading automated AI platform moderation. They employ obfuscation tactics, such as posting text in an image to evade automated word detection, as seen in Figure 4.1 on the following page. This tactic is particularly prevalent on image- and video-based platforms such as Instagram and TikTok.11

W&LI posts were regularly among the top engaged-with COVID-19 vaccine misinformation posts on social media during VP’s period of observation. As an example, Erin Elizabeth at Health Nut News was a recurring W&LI who regularly authored viral false and misleading content.13 Erin Elizabeth runs a website featuring her blog, newsletter, and online store. Her content is known to veer into conspiracy theory.14 Prior to COVID-19, her social media profiles featured recurring allegations that holistic doctors were being murdered and that Big Pharma was behind the incidents.15 As the COVID-19 vaccine rollout progressed, her Facebook posts shared stories of possible vaccine injury or death and promoted health freedom. Several of her posts were the most engaged-with posts in VP’s weekly observations.

While some W&LIs engaged with explicit anti-vaccine content prior to the pandemic, the movement is fundamentally distrustful of the medical establishment,16 and as a result is a particularly receptive audience for COVID-19-related conspiracy theory and anti-vaccine content. For example, VP analysts witnessed an evolving, ongoing relationship between QAnon-related wellness influencers, medical freedom influencers, and long-standing anti-vaccine influencers at various gatherings.
4. Actors & Tactics

![Image of Joseph Yi using the term "MaXXXine"
for vaccine](https://via.placeholder.com/150)

Figure 4.1: Anti-vaccine wellness influencer Joseph Yi uses “MaXXXine” as code for “vaccine” in a video and encourages his followers to visit their free “Censorsh!t” Telegram channel.12

... and events used to amplify each other’s platforms.17

W&LLIs often have financial incentives in sowing distrust in traditional medicine and vaccines, particularly those who sell their own wellness products and services. These accounts commonly misled their followers about the seriousness of COVID-19 or the pandemic overall, and simultaneously promoted supplements and devices, claiming they carry health benefits.18 Many of these influencers already had an established online brand that they promoted. Pivoting their marketing to include COVID-19 and spreading anti-COVID-19 vaccine content provided an opportunity to broaden their audience.

For example, the mother-son duo Ali and Alec Zeck, and the group Health Freedom for Humanity (HFFH), gained significant popularity during the vaccine rollout before their Facebook and Instagram accounts were both taken down. The duo’s content

---

4.1.3 Pseudomedical Influencers

Pseudomedical influencers (PMIs) are individuals with medical backgrounds who leverage their credentials as badges of authority while undermining vaccines, often using pseudomedical rhetoric. They often prominently display their title (e.g., Dr.) or degree (e.g., RN) in their social media usernames and present themselves as authoritative experts who are simply deeply skeptical of vaccines. Many in fact come from an alternative-health background, such as naturopaths, chiropractors, and homeopathists, or are MDs or DOs who do not have expertise in immunology or infectious disease, or who outwardly express viewpoints contrary to the medical establishment. PMIs are analogous to The Doctors in Tara Smith’s 2017 framework for anti-vaccine influencers. In VP observations, members of this category commonly had a well-documented history of spreading vaccine misinformation prior to the pandemic, though several rose to prominence by spreading misinformation specifically about COVID-19 vaccines. Other influencers from this group, such as Judy Mikovits and Robert Malone, were former virologists who had papers retracted or have had prior professional fallings-out with the professionals and institutions who are now developing or distributing COVID-19 vaccines.

PMIs rely on a host of tactics to spread vaccine misinformation as they leverage their alleged authoritative position as medical practitioners. Their tactics are effective: 21% of engagement across all tickets and topics was associated with this tactic. Dr. Robert Malone, who falsely claims to have invented mRNA vaccines and now spreads misinformation about the danger of such vaccines, has been featured on several popular shows that spread anti-vaccine misinformation, including The
4. Actors & Tactics

*HighWire* with Del Bigtree and *Tucker Carlson Tonight* on Fox News. On these shows, Dr. Malone can use his title and self-styled claim to expertise (“mRNA inventor” or “inventor of mRNA vaccines”) to lend extra weight to, for instance, decrying the use of mRNA vaccines in children or making misleading claims about the harmful effects of vaccine spike proteins. Few in the audience are likely to dig into the facts of the self-styled claim: while Malone played a role in early mRNA research, he has not played a significant role in its application to vaccines. In-depth coverage of the research behind mRNA vaccines and its pioneers do not mention Malone.

Joseph Mercola, an osteopathic physician cited in 30 VP incidents, was another prominent figure in this group. Mercola has a sprawling platform presence. He has multiple English and Spanish accounts on WhatsApp, Telegram, BitChute, and Facebook and close to 4 million followers and subscribers across his social media accounts. His 1.7 million-follower Facebook Page, shown in Figure 4.2 on the next page, has advertised his book *The Truth About COVID-19*, which has prominent placement on Amazon in COVID-19 book search results. According to the *New York Times*, he has also set up companies like Mercola.com Health Resources and Mercola Consulting Services to churn out content in several languages across websites and social media. While his posts on mainstream media are often framed as “just asking questions,” links to his articles direct his large online audience to his blog. On August 4, 2021, Mercola announced that he was deleting 15,000 past posts from this site, and though he would still upload daily articles, he would remove these writings every 48 hours to work around platform moderation. In one notable incident involving Mercola, he argued in a video posted on BitChute and spread via Facebook on July 10, 2021, that COVID-19 vaccines would “destroy the lives of millions,” describing the vaccine as “experimental gene therapy” that may cause chronic, long-term inflammation that will disable and kill large swaths of the population. The video garnered over 95,700 views.

Mercola and other PMIs also commonly use manipulated statistics and hard-to-verify data as the backbone of their mis- and disinformation. Throughout the pandemic, PMIs in particular have commonly shared studies and clinical trial data that are hard for users to interpret or verify, while presenting misleading claims about the findings. For instance, in a May 25, 2021, Facebook post, Mercola decontextualized findings from a study from Washington University School of Medicine in St. Louis showing that “mild Covid-19 induces lasting antibody protection.” Mercola’s post added, “This is good news for everyone except the vaccine manufacturers and health authorities who have said otherwise.”

Simone Gold, a licensed emergency room physician, was the second most prominent PMI across Virality Project’s tickets. Gold is the leader of America’s Frontline Doctors, a right-leaning group opposed to pandemic mitigation measures. Gold has been influential since the summer of 2020, when the White Coat Summit, an event
broadcast online in which members of America’s Frontline Doctors spoke on the steps of the Supreme Court. The White Coat Summit promoted hydroxychloroquine both as a preventative measure and as a cure for COVID-19; one speaker at the event, Dr. Stella Immanuel, had previously been best known for claiming that gynecological issues are caused by having sex with witches and demons. Since then, Gold consistently made news and appeared in almost half of VP’s weekly briefings for spreading a variety of false and misleading claims about the COVID-19 vaccine—ranging from encouraging her followers to boycott companies for their vaccine protocols to organizing a cross-country tour to fight back against “censorship, chaos, and the undeniable slide towards communism that lurks beneath the tyrannical lockdowns for governmental ‘public health’ policy.”

Along with repeatedly posting false and misleading information about the COVID-19 vaccine, Gold was an influential political voice in 2020 with crossover into the right-leaning political influencer group. She persistently amplified false information about the 2020 election and voter fraud and was also arrested for participating in the January 6, 2021, storming of the US Capitol. Gold remains active on Twitter and TikTok. On her Twitter account, she consistently links to the American Frontline Doctors website, where a range of anti-vaccine films, blog posts, and petitions are promoted.
4.1.4 Conspiracy Theory Influencers

Conspiracy Theory Influencers (CTIs) focus their content on false or misleading stories about COVID-19 vaccines that frame the vaccines as part of a larger nefarious agenda by an overarching enemy—usually some form of global elite, such as the pharmaceutical industry, government, or billionaires like Bill Gates. Many influencers in this group have a history of spreading other conspiracy theories and often inflected COVID-19 vaccine conspiracy theories in such a way as to make them fit the belief structures of other movements, such as followers of the right-wing conspiracy theory QAnon.51

Though only 20% of total incidents had a conspiracy theory focus, CTIs generated the largest share of engagement—37%—during the Virality Project’s observation period. Alt-platforms (e.g., Rumble, Gab, BitChute, and Righteon) and encrypted messaging apps (e.g., Telegram) are popular with CTIs, as many of the adherents have been, or anticipate being, deplatformed by mainstream social media. The audience size of alternative social media is often smaller, but may be more committed and passionate about the topic. Despite the smaller size of their initial potential audience, alt-platform conspiracy theory influencer ideas do still make their way to mainstream social sites through devoted fans sharing across platforms the links and screenshots that they believe to be the truth.

One recurring actor within this category was Pfizer’s former vice-president, Michael Yeadon.52 Yeadon lost his position at Pfizer in 2011, but continued to work in the pharmaceutical industry as founder of the drug company Ziarco, which he sold in 2017 to Novartis.53 In late 2020, Yeadon began to spread multiple false claims about COVID-19 and the vaccines, aided by reports that incorrectly attributed his claims to “Pfizer’s head of research.”54

Yeadon has made various false claims about the harms of the vaccines,55 including that they cause infertility, but has folded these claims into the larger conspiracy that government-led vaccine rollouts are a ploy to initiate a “massive-scale depopulation” event.56 In March 2021, America’s Frontline Doctors featured Yeadon in an article making a similar claim.57 Coverage of the interview received roughly 10,000 Facebook interactions, primarily within COVID-19 misinformation groups.58 In June 2021, Yeadon spoke on War Room: Pandemic with Steve Bannon, where he claimed that the vaccine is 50 times more likely to kill young people than the virus is, citing the alleged danger of the “spike protein.”59 He has also appeared on conspiracy theory and conservative outlets like SecuringAmericaTV and The Full English Show.60

Videos appeared in 41% of conspiracy theory tickets. CTIs leverage an elaborate network of alt-channels, documentaries, podcasts, and media to spread their claims among their community. By relying on other CTIs to interview and host each other (see Figure 4.3 on the facing page), “experts” can claim special insider knowledge
about a vast range of topics related to COVID-19 vaccines or their rollout, and can advance these claims on outlets not constrained by journalistic principles. The CTIs do not necessarily have high-profile media accounts at a national- or mainstream-audience level; instead, they act as mini-celebrities, or local influencers, within the conspiracy theory community. While nationally prominent influencers may have the ability to reach many, or to reinforce the preexisting beliefs of their audience, local online influencers like CTIs can work to shape opinion within these local networks to greater effect because they are trusted members of the community. Further, the quantity of interviews and discussions hosted on these accounts gives the impression that conspiracy theorists draw from a diverse and robust pool of experts who each provide a piece to the overall puzzle of the conspiracy theory narrative.

Figure 4.3: The insular and self-reinforcing world of conspiracy theories. This includes other CTIs present in VP’s data and briefings such as Christiane Northrup, Reiner Fuellmich, and Stew Peters.

### 4.1.5 Right-Leaning Political Influencers

Right-leaning Political Influencers (RLPIs) include right-leaning media figures, politicians, and publications. These figures produce some of the most engaged-with daily stories on social media, far beyond the topic of vaccines. RLPIs can easily harness a well-oiled machine of media publications and influencers. By using a highly organized, multimedia, cross-platform approach—involving cable television, online blogs, and social media (mainly Facebook and alternative platforms
like Gab, Parler, and Telegram)—this group can maximize engagement among their followers and targeted audiences—who also tend to fall on a spectrum to the right of the American political center—to boost virality of their messages, which are often framed sensationaly. Yochai Benkler, co-director of the Berkman Klein Center for Internet and Society, has joined other scholars in analyzing and documenting these dynamics, describing how the right-leaning media ecosystem differs from the rest of the media environment in ways that make it especially vulnerable to the spread of mis- and disinformation.64 According to Benkler, false stories find their birthplace on explicitly non-journalistic and extreme websites, such as InfoWars, and are then amplified by more established media outlets that boast extremely large, established audiences.65 Recurring actor dynamics among RLPIs span an array of topics determined to be politically expedient; prior to the COVID-19 pandemic, researchers from the Election Integrity Partnership observed several of the same prominent actors spreading false and misleading claims about voter fraud.66 From the start of the COVID-19 pandemic, RLPIs incorporated uncorroborated claims about COVID-19 and vaccines.

Fox News has played a particularly pivotal role in spreading vaccine misinformation and anti-vaccine beliefs during the COVID-19 pandemic. Media Matters observed that in a two-week period between June 28 and July 11, 2021, Fox News ran 129 segments about the COVID-19 vaccine on its cable broadcast; more than half of those segments included unverified claims that undermined vaccination efforts.67 A November 2021 study from the Kaiser Family Foundation found that people who trust Fox News and other conservative media outlets are more likely to believe falsehoods about COVID-19 and the vaccines than people who primarily consume other news.68 After Fox News pivoted to a more pro-vaccine stance in late July 2021, a Morning Consult poll found that vaccine hesitancy among their viewers dropped by ten percent, suggesting that Fox News may have outsize influence on its viewers’ vaccine beliefs and behavior.69

Fox News television host Tucker Carlson has been one of the most prominent and sensationalist spreaders of false or misleading information about vaccines throughout the COVID-19 pandemic. Despite claiming he is not anti-vaccine, Carlson has repeatedly featured vaccine-skeptical segments.70 He has encouraged viewers to question the COVID-19 vaccine, referred to the vaccine rollout as “eugenics” and “draconian,” and highlighted unconfirmed instances of adverse vaccine reactions (a practice he would continue as the vaccine rollout progressed).71 Carlson has also promoted skepticism by inviting guests onto his show who claim they experienced vaccine injuries. On July 1, 2021, Carlson interviewed a mother who claims that her 12-year-old daughter suffered severe harm after participating in the Pfizer vaccine trial. The interview received over 700,000 views on Facebook. The claim that the daughter was harmed by the vaccine has not been proven.72

Like Carlson, Candace Owens is a conservative author and media personality who
4.1. A Cast of Recurring Actors

has featured prominently in the COVID-19 vaccine conversation. Owens hosts a talk show on the leading conservative news website, The Daily Wire, and is known for co-headlining on Turning Point USA’s Campus Clash speaking tour.\textsuperscript{73} She has consistently used her expansive platform to spread COVID-19 mis- and disinformation by making false claims about COVID-19 statistics, the CDC, and the COVID-19 vaccines.\textsuperscript{74}

In one example, Owens claimed that the CDC proposed putting those at high risk of coronavirus infection in camps.\textsuperscript{75} In another, Owen became a key amplifier of an accidental false claim from Britain’s Chief Scientific Adviser, Sir Patrick Vallance.\textsuperscript{76} On July 19, 2021, Vallance accidentally overstated the rate of vaccination in the percentage of patients hospitalized with COVID-19 infections. Though Vallance’s claim was quickly corrected,\textsuperscript{77} that same day Owens posted about the news story, claiming that the “Biden regime’s censorship team” would attempt to keep information from the public.\textsuperscript{78} Owens’ politicization of the misstated statistic was meant to expose the Left’s “hypocrisy” with respect to censorship, fueling the notion that any platform removal of RLPI posts (due to mis- and disinformation moderation policies) is part of a malicious attempt by the Left to silence right-leaning voices. Even after Vallance’s correction, Owens continued to sow doubt in the safety and efficacy of the vaccines by updating her original post with language that was critical of Vallance’s correction. This post amassed over 50,000 reactions.

Media personalities and politicians have a thriving symbiosis. Tucker Carlson’s claims have repeatedly found purchase with right-wing politicians, who play an important role in the vaccine conversation and the spread of false and misleading claims. Senator Ron Johnson (R-WI) appeared on Carlson’s show to discuss why he was not getting vaccinated in a clip rife with misinformation about vaccine deaths.\textsuperscript{79} This clip received over 100,000 views and 2,000 shares on Facebook. Sen. Johnson also promoted hard-to-verify claims of vaccine injury when he tweeted quotes from Carlson’s interview with the mother of the child in the Pfizer vaccine trial. In June 2021, he held an event for individuals to testify about adverse reactions to COVID-19 vaccines. The event was subject to criticism from medical professionals.\textsuperscript{80} However, when pressed about it, Sen. Johnson stated that he was not anti-vaccine and was just “asking questions”—a tactic commonly used by spreaders of misinformation to deflect culpability.\textsuperscript{81}

Perhaps the most well-known politician amplifying vaccine mis- and disinformation is Representative Marjorie Taylor Greene (R-GA).\textsuperscript{82} Rep. Greene has been characterized as a far-right conspiracy theorist\textsuperscript{83} with an extensive track record of promoting COVID-19 conspiracy theories and politicizing measures to curb the spread of the virus.\textsuperscript{84} She has made inflammatory remarks about COVID-19 mitigation polices (such as comparing COVID-19 safety measures to the persecution of Jews during the Holocaust) and has actively defied congressional COVID-19 regulations, receiving an estimated $48,000 in fines for refusing to wear her mask on the House floor.\textsuperscript{85}
4. Actors & Tactics

On March 29, 2021, Rep. Greene livestreamed a video in which she framed vaccine passports as Biden’s “Mark of the Beast” (see Chapter 3: Narratives). Rep. Greene’s comments received widespread engagement, with over 500,000 interactions across Facebook, Twitter, Instagram, and Reddit, elevating this narrative in the public eye and bringing it to audiences beyond communities focused on conspiracy theories and religious content.

4.1.6 Medical Freedom Influencers

The “health choice” or “medical freedom” movement espouses an aversion to government interference in individual or family health choices. It emerged as a distinct force in the mid 2010s, as a growing spate of preventable disease outbreaks (such as the 2014–2015 Disneyland measles outbreak⁸⁶) led state legislatures to crack down on the abuse of childhood vaccine exemptions for school attendance. Anti-vaccine parents who objected to the strengthening of public health requirements found that pseudoscience-based arguments about vaccine safety and efficacy were largely ineffective in combating those legislative processes, and so they pivoted to arguing that the state was overreaching, usurping parental authority. They partnered with local Tea Party organizations and others sympathetic to concerns about government overreach more generally to expand their base.⁸⁷ During the late 2010s, the health freedom movement expanded through newly established political action committees in Texas, Oklahoma, Oregon, Michigan, and Ohio.⁸⁸ Organizations soon formed in almost every state.

The newest iteration of medical freedom, adapted for COVID-19, challenges the legitimacy of government or corporate vaccine mandates and public health interventions specific to COVID-19, including vaccine passport systems and masking requirements. As COVID-19 emerged, the same health freedom groups previously dedicated to opposing childhood immunization requirements expanded to protesting social distancing, masks, and other prevention measures,⁸⁹ and gained large followings on social media by networking with right-leaning and libertarian actors. While these groups explicitly advocate for freedom of “vaccine choice,” they additionally often propagate vaccine doubt by merging arguments for choice with misleading claims about efficacy, risk, or adverse medical consequences. For example, local organizations rallying around vaccine choice or medical freedom often simultaneously promote vaccine mis- and disinformation and boost posts from vaccine conspiracy theory and pseudomedical communities.⁹⁰

Medical freedom influencers (MFIs) active in the anti-COVID-19-vaccine movement were fairly distinct from other categories of influencer in that rather than hinging on a handful of key (and often celebrity-status) individuals, they spread their narratives via a franchise model across all 50 states. One variety of these entities, the self-described “Freedom Keepers,” put out content specific to a local geographic area, but leveraged Instagram-friendly yellow branding and motivational messaging
4.1. A Cast of Recurring Actors

common to the national movement—such as using language describing themselves as “inclusive global community,” “raising awareness,” or “empowering individuals” (see Figures 4.4 and 4.5 on the following page).91

![Figure 4.4: Instagram icons for local Freedom Keepers groups.](image)

No single “vaccine choice” organization has amassed a following of a size comparable to the prominent influencers discussed in other categories here. However, the networked amplification of specific messages, linking local experiences to a national movement, and the use of state and local chapters to further boost the content of anti-vaccine influencers, suggests the aggregate engagement across these localized organizations is worth additional study.

Many of the MFI incidents focused intensely on debates about vaccine mandates and passports. As medical freedom activists have fought requirements imposed by states, cities, or private employers, they have learned from each others’ successes and failures—at a messaging and an organizing level—and have brought those lessons to their local communities. What one state does, another state will often echo. For instance, on July 17, 2021, hundreds of people gathered outside several Henry Ford Health System facilities in Detroit to protest the vaccine mandate issued for its employees.92 Posts discussing the protests garnered at least 60,000 interactions. Later that month, California Health Choice Advocacy released a letter to Governor Gavin Newsom demanding he not implement vaccine requirements for government and healthcare employees.93 The latter incident did not garner high engagement online, but together, the responses showcase the consistent frequency and geographic range of these groups.

As observed in Detroit, MFIs have been successful in taking advantage of their local
community focus to gather in-person momentum, driving real-world turnout at protests. Decentralized medical freedom movements have sprung up around the world, advocating against mandatory vaccination, lockdowns, and mask mandates. During the VP observation period, the largest of these in-person events was the Worldwide Rally for Freedom, which held three mass protests in March, May, and July 2021.94 These rallies were held in as many as 100 cities worldwide, with demonstrations taking place in some US states such as Hawaii, Minnesota, California, and Texas.95 Many posts featured flyers advertising the demonstrations and used the hashtag #WeWillALLBeThere, which has received a total of 427,645 interactions on Facebook since February 2021.

4.1.7 Defies Classification

Some nontraditional influencers popped up as recurring figures and drivers of the conversations on social media about COVID-19 vaccines. Although they occasionally espouse some of the same beliefs as influencers in the categories previously described in this chapter (notably, medical freedom), are amplified by those other influencers, and have gained a significant degree of trust across several communities, they remain difficult to classify beyond, perhaps, “anti-establishment.”

Dr. Naomi Wolf, once a feminist icon who became known for an array of conspiratorial claims as early as 2014,96 expanded her pseudomedical claims to COVID-19
4.1. A Cast of Recurring Actors

vaccines. However, Wolf does not squarely fit into clearly delineated conspiracy theory groups, nor does she attribute COVID-19-related events to larger cabal-like plots. Still, her COVID-19 vaccine misinformation has been far-reaching. On one occasion she tweeted that unvaccinated women are reporting reproductive cycle changes after being near vaccinated people, suggesting it needs more investigation (a common misinformation rhetorical strategy). Her conspiratorial messages centered in large part around vaccine passports; in spring 2021 these claims achieved high traction within conspiracy theory communities. At the time, she appeared on Fox News, where she called a passport rollout “the end of human liberty in the West,” pushing government overreach narratives popular in right-leaning communities and communities rallying around medical freedom. Wolf was deplatformed by Twitter in June 2021 for spreading vaccine misinformation.

Alex Berenson is perhaps the most significant influencer who defies categorization. A former New York Times reporter and a bestselling novelist with no specific anti-vaccine background (vaccine-related tweets such as the one shown in Figure 4.6 suggested he supported vaccine requirements for routine childhood vaccines), Berenson was however hesitant about potential COVID-19 vaccines, however, and over time evolved into a key player in repeatedly spreading false and misleading information about the COVID-19 pandemic and vaccines. He underplayed the danger of the virus and challenged the efficacy of vaccines and masks, even as evidence supported their value as life-saving public health measures.

Berenson’s popular posts on Twitter notably claimed to be “digging up” or “uncovering” information that was hidden from the public about vaccine safety or effectiveness. In one incident in July 2021, Berenson amplified a conspiracy theory
4. Actors & Tactics

...from a statement filed with a lawsuit from America’s Frontline Doctors stating that the government was covering up more than 45,000 vaccine-related deaths. Berenson’s 17-tweet thread, which received over 16,000 interactions on July 21, 2021, claimed that the CDC had “quietly more than DOUBLED” the number of deaths reported in VAERS, suggesting the CDC had misled the public. In other viral incidents, Berenson used screenshots of scientific articles to attempt to prove mRNA vaccines are a form of gene therapy, as well as screenshots from a data leak from the European Medicines Agency’s 2020 cyberattack (involving email correspondence about the approval process for the Pfizer vaccine) that suggest the FDA felt “extreme pressure” to allow the “fast sale” of the vaccine without “substantive” evidence that the vaccine is safe and effective. Twitter permanently deplatformed Berenson in August 2021 for repeated violations of Twitter’s COVID-19 falsehoods policy. At the time he lost his account, he had more than 200,000 followers. He has since become a prominent user of the subscriber-based blog-hosting site Substack, where he puts out a newsletter, Unreported Truths, that continues to spread a variety of false and misleading claims about COVID-19 vaccines. Berenson additionally notes that Substack has been a lucrative endeavor for him, as he is now able to charge $60 for annual subscriptions, and claims to have amassed hundreds of thousands of subscribers.1

4.2 Spanish and Chinese Influencers

Recognizing that mis- and disinformation targeting non-English-speaking communities often goes understudied, Virality Project researchers prioritized having visibility into content in Spanish and Chinese in the US. The challenges of addressing mis- and disinformation in languages other than English on social media are well documented; this holds for vaccine mis- and disinformation as well. During the course of the Virality Project, three analysts (all with native language proficiency) monitored Spanish- and Chinese-language content circulating in US communities. It should be noted that the volume of non-English-language content far surpassed the Virality Project’s ability to capture and analyze all of it. Furthermore, several non-English-speaking communities share news on social platforms such as WhatsApp and China-based WeChat, which are far less visible to analysts than more open platforms such as Facebook and Twitter.

The anti-vaccine narratives spreading within non-English-speaking communities on mainstream platforms in the US were similar to English-language content; for example, they frequently centered around safety concerns. English-language content, particularly videos from popular activists and influencers, was often translated and

---

1Substack only publicizes the order of magnitude, not specific numbers, of subscribers.
2Specifically, VP monitored textual content in Simplified and Traditional Chinese, as well as video content in spoken Mandarin.
4.2. Spanish and Chinese Influencers

subtitled, word-for-word, and shared to Spanish and Chinese channels. Yet English-language content did not need to be translated word-for-word for the narrative to take hold and spread. Channels often shared a link to an English-language article or video alongside an in-language summary, excerpted quote, or misinformative analysis. This allowed non-English-language channels to keep up with the English-language narrative zeitgeist and share content more quickly than word-for-word translations of videos and articles. Of all the videos analysts identified in Spanish and Chinese, 47% originally appeared in English and were translated via subtitles or paraphrased in text accompanying the post.

English-speaking actors whose posts and videos were used in Spanish- and Chinese-language communities spanned the cast of recurring characters:

• Long-standing anti-vaccine activists, like RFK Jr (see Figure 4.7).113
• Pseudomedical influencers, including Dr. Judy Mikovits,114 Dr. Sheri Tepenny,115 and Dr. Dolores Cahill.116
• Conspiracy theory influencers, such as Dr. Christiane Northrup (a doctor, but also a QAnon adherent),117 and Stew Peters (see Figure 4.8 on the next page).118
• Right-wing media personality Tucker Carlson.119
• Religious influencers; these actors achieved some reach in Spanish-language channels, but were not popular in Chinese-language ones.

Figure 4.7: Screenshot of a subtitled video of Children’s Health Defense founder RFK Jr., uploaded by Spanish-language media entity BLes Mundo.
Spanish and Chinese channels also featured figures who created original content in-language. The actors discussed in this section are native speakers involved with anti-vaccine activism, while foreign state actors and state media are discussed in Section 4.3 on page 96. In both Spanish and Chinese, influencers spread misinformation about the safety and efficacy of vaccines, as well as vaccine conspiracy theories. Pseudomedical influencers were particularly popular in both Spanish- and Chinese-language online communities. Notably, content from wellness influencers appeared to be less popular in both Chinese- and Spanish-language communities. Analysts did not surface significant content from specific wellness influencers within those communities, nor did they find significant content from English-speaking wellness influencers that had been translated and reposted to non-English-language channels.

### 4.2.1 Spanish: Doctors for Truth and Natalia Prego Cancelo

In Spanish, many Telegram channels featured individuals whose medical credentials conferred undue authority on topics like vaccines. Some of these influencers were in fact medical doctors; however, they spread information about vaccines not accepted
by the medical community at large. VP analysts identified five Spanish-language Telegram channels affiliated with a loose organization of pandemic-skeptical, Spanish-speaking “médicos por la verdad” (“doctors for truth”), who often spread content from pseudomedical influencers. Apart from a general “Médicos y Ciudadanos por la Verdad” (“Doctors and Citizens for Truth”) channel, other channels purported to represent doctors and healthcare workers from Spain, Argentina, and Mexico (see Figure 4.9). As of September 1, 2021, those five Telegram channels had a combined membership of more than 105,000 users. Various “Médicos por la Verdad” channels appeared to be loosely formed, and did not include information about who runs them or proof of their credentials.

Figure 4.9: A post on Telegram channel Médicos por la Verdad Argentina, which says “Cheers to those remaining humans who have kept their DNA intact.”
4. Actors & Tactics

One of the most prominent, recurring anti-vaccine influencers in Spanish successfully spread vaccine misinformation through a separate (though identically named) organization, Médicos por la Verdad. The organization’s co-founder, Dr. Natalia Prego Cancelo, is a COVID-19 denialist and anti-vaccine influencer based in Spain. It does not appear that any of the Telegram channels discussed above are formally affiliated with the organization run by Prego. Prego is a family doctor who was fired by the private clinic she worked for because of false statements she made about the COVID-19 pandemic. However, Prego still uses her credentials as a doctor to lend credence to the variety of false claims she has spread—including that COVID-19 vaccines are “contagious” and that the vaccine produces dangerous spike proteins. She has also promoted misinformation from other pseudomedical influencers like Robert Malone. VP analysts observed Prego and others creating and spreading anti-vaccine content in Spanish.

Despite her rise in popularity since the early days of the pandemic, Prego’s personal reach today remains small in comparison to English-speaking influencers. Prego is most active on her personal Telegram accounts, where, as of January 2022, she had more than 28,000 subscribers. She uploads sporadically to YouTube and frequently to alternative video sites like Rumble and Odysee (see Figure 4.10 on the facing page); she then promotes these videos through her Telegram channel. The videos have less reach than the content of many of her English-speaking counterparts—usually receiving low-four-digit engagements. Prego uses her channels to spread conspiracy theories primarily about the pandemic and COVID-19 vaccines, and to organize events and promote in-person protests. Though Prego’s content was unique to Spanish-language channels, she spread many of the same conspiratorial narratives about vaccines familiar in English-speaking channels and promoted English-speaking anti-vaccine influencers to her followers.

4.2.2 Chinese: Guo Wengui and Yan Li-Meng

Guo Wengui is an exiled Chinese businessman and critic of the Chinese Communist Party (CCP) who has been referred to as the “center of a vast network of interrelated media entities which have disseminated online disinformation.” Guo and his media entities have spread mis- and disinformation related to several topics, including vaccines and the COVID-19 pandemic. From the beginning of VP’s observational period in February 2021, Guo and his network created and spread content promoting false narratives about COVID-19 vaccines on Telegram, Twitter, YouTube, and Guo’s own websites. VP analysts did not surface content created by Guo’s network on WeChat—perhaps unsurprising, given WeChat’s strict moderation policies and Guo’s criticism of the CCP.

Among the conspiratorial claims that Guo and his network of media properties spread were allegations that the CCP created the virus as a bioweapon, and that only the CCP truly understands how to stop the virus (using their own antidotes). One
4.2. Spanish and Chinese Influencers

![Image](image_url)

Figure 4.10: Dr. Natalia Prego Cancelo claims in a video uploaded to Rumble that COVID-19 vaccines are contagious.

such meme read, “Antidotes are possessed and controlled by the CCP, but there is no vaccine!” Chinese virologist Yan Li-Meng—who came to prominence in 2020 for asserting that COVID-19 was purposefully created in a lab as a CCP bioweapon—made similar claims. Both Guo and his colleague, American conservative political operative Steve Bannon, have been instrumental in promoting Yan’s claims. According to reporting by the Washington Post, Yan claims affiliations with the Rule of Law Society and Rule of Law Foundation, both founded by Guo. Guo has also used his various media properties to promote Yan and her claims in both Chinese and English. Bannon has further bolstered Yan’s claims on his podcast, War Room: Pandemic, which has repeatedly spread mis- and disinformation about COVID-19 vaccines.

Much like other pseudomedical influencers, Yan’s credentials lend her claims credence. As the Washington Post reported, Yan has a medical degree and a PhD in ophthalmology, and was previously a University of Hong Kong postdoctoral fellow. Her credentials enable others, like Guo and Bannon, to elevate her as an expert despite the misinformative nature of her claims. Guo’s influencer status among people critical of the CCP, along with the reach of his media properties in multiple languages, were an ideal combination to push these claims. Likewise, Bannon’s status in the US helped him spread these claims to English-speaking American audiences.
4.3 Foreign Actors

The borderless nature of the coronavirus and the global effort by governments to vaccinate their citizens created a complicated dynamic between vaccination efforts and broader diplomatic efforts. This relationship sometimes took the form of “vaccine diplomacy,” wherein wealthy nations donate vaccines to other nations in order to strengthen diplomatic and economic relationships while bolstering their own reputations. Other times it manifested as leaders touting the efficacy of their national effort or of the vaccines their country had created. While all nations have engaged in some form of diplomacy or messaging related to vaccines, the Virality Project focused on observing state media and social media networks linked to Russia, China, and Iran. These three nations not only promoted their own efforts, but also attempted to undermine their rivals.

Researchers and journalists documented social media manipulation efforts targeting the US COVID-19 vaccine conversation with false or misleading claims even prior to the vaccine rollout. The Virality Project created a researcher pod that focused on understanding state media narratives from these actors as the vaccine rollout progressed.

Russia, China, and Iran used a full-spectrum propaganda approach, with capabilities spanning both media and social media in attempts to influence vaccine conversations in the US and around the world. “Full-spectrum propaganda” refers to nation-state actors’ use of both overt and covert channels across media types in

Figure 4.11: GNews (one of Guo’s properties) graphic shared by a Guo supporter on Twitter. The text claims that the CCP made a viral biological weapon for which only they have the antidote.
order to influence public opinion from a variety of angles.\textsuperscript{143}

These include:

- Traditional diplomatic channels
- State-run or -affiliated media outlets
- Overt state-affiliated social media accounts
- Covert social media accounts
- Grey media properties\textsuperscript{144}
- Cyber activity

These approaches are often used simultaneously, with some degree of coordination, to achieve strategic objectives. Although the covert operations analyzed below typically used negative, outwardly adversarial and divisive messaging to denigrate Western vaccines and institutions, the corresponding overt media outlets clearly attributable to the state tended to use comparatively neutral language, selective topical coverage, and self-promotion in order to promote their own governments’ vaccines and vaccination effort relative to those of their counterparts in the West. Virality Project researchers observed that the narratives used by overt and covert foreign-state actors to influence US conversations around the pandemic and COVID-19 vaccines are very similar. However, the actual mechanisms used by overt and covert campaigns tend to differ in meaningful ways. Our analysis focuses specifically on these full-spectrum efforts.

When state-actor tactics include unambiguous mistruths or inauthentic accounts, platforms are more willing to label or take down content or to suspend accounts; the decidedly opinionated messages of overt state media propaganda rarely fall cleanly into these categories. Recognizing this challenge, social media platforms began to label state media properties and state-linked accounts early in the COVID-19 pandemic to assist their audiences in understanding the source—and, thus, the possible agenda or motivation—of the content or speaker.\textsuperscript{145}

### 4.3.1 Russia

The first operation linked to Russian actors involved a small network of fake personas that targeted right-leaning audiences on alternative social media platforms Gab and Parler, as well as the far-right discussion forum patriots[,],win, as early as November 2020. The accounts in the network posted a series of memes, articles, and messages that appeared aimed to exacerbate existing social and political tensions in the United States, including around the Biden administration’s response to COVID-19 and the vaccine rollout. Many of these posts featured highly inflammatory political cartoons, which may have been created by the actors themselves, and leveraged a
variety of divisive partisan themes, such as suggestions that Biden was benefiting from Trump’s hard work to combat the pandemic while ignoring other issues such as rising US inflation. The accounts also attempted to amplify anti-vaccine conspiracy theories, spread doubts about vaccine efficacy, and depict the vaccine rollout as state-imposed oppression designed to enrich pharmaceutical companies and erode American freedoms. The actors behind this operation have been linked to the same group responsible for the “Newsroom for American and European Based Citizens” (NAEBC), a fake right-wing news outlet that targeted US audiences ahead of the 2020 election and was connected to Russia’s Internet Research Agency.

Figure 4.12: Memes originally posted by Russian account NAEBC on patriots[,]win implying that Kamala Harris (left) and Joe Biden (right) are forcibly vaccinating individuals; the Biden meme also implies that vaccines contain Microsoft microchips.

A second operation was attributed to a marketing firm operating from Russia called Fazze. While it is not clear who Fazze was working for, the company appears to have engaged in a concerted effort to covertly disseminate narratives online denigrating Western-made COVID-19 vaccines, including by attempting to hire YouTube influencers to deliver unattributed messages. In late 2020, for example, the group spread memes on Facebook and Instagram suggesting the AstraZeneca vaccine was dangerous because it was developed from a chimpanzee adenovirus—claims echoed by some Russian state media outlets that have attacked AstraZeneca for producing a “monkey vaccine.” Months later, the same actors were caught attempting to seed misleading information online about the safety of the Pfizer vaccine, this time by disseminating an alleged internal AstraZeneca report on vaccine mortality rates. Fake personas posted copies of the report online alongside a propaganda article, which claimed that the report was hacked or leaked. Fake social media accounts then posted these articles to groups concerned with COVID-19 information and vaccine safety. Fazze also contacted prominent social media influencers, offering to pay them to post videos amplifying the claims to YouTube, Instagram, and TikTok. The effort was quickly exposed, in part because the targeted influencers posted instead about the manipulative outreach, but some reports have noted that at least two influencers posted videos that appear to match the instructions that Fazze distributed.
4.3. Foreign Actors

The narrative themes present in the efforts undertaken by the covert actors, particularly attempts to exacerbate existing social and political cleavages, were mirrored by Russian state media and representatives. For example, RT amplified claims suggesting vaccine passports are akin to government overreach and population segregation, proclaimed the emergence of a global “#Covid apartheid,” and suggested that harms associated with mRNA vaccines were downplayed by the media in collusion with US business elites. Russian state media also amplified research of questionable veracity that suggested Western vaccines are ineffective or harmful. However, in contrast to the covert campaigns, which were not intended to be obviously pro-Russian, the overt state media approach additionally contrasted the supposed failures of Western vaccines with the success of the Sputnik V vaccine. When the rollout of the Johnson & Johnson vaccine was paused to review a possible risk of blood clots, Russian state media paired neutral coverage of the specific story with articles highlighting the lack of blood clots associated with the Sputnik V vaccine. Russian state media additionally featured interviews with anti-vaccine influencers who had been deplatformed by mainstream social media platforms—implying or suggesting that the US was censoring those critical of potentially dangerous vaccines.

Figure 4.13: Copy of the instruction document allegedly sent to French influencers by Fazze and published by French outlet factandfurious.com.

---

99
4.3.2 China

Several covert influence operations linked to the CCP were taken down in 2020–2021, and included content related to COVID-19 or the COVID-19 vaccines. Because it gained international prominence as it spread through Wuhan, messaging about the virus was a topic of high priority for the well-established CCP overt propaganda apparatus; leading officials from the propaganda department were assigned to the small committee responsible for COVID-19 messaging. And, as with Russia, this messaging appeared in both overt attributable state media, which focused mostly on telling positive stories of the CCP’s handling of the virus and its vaccine development and rollout, as well as covert networks, which primarily denigrated Western vaccine efforts, on Western social media platforms.

One long-standing China-aligned network, identified and dubbed “Spamouflage” by Graphika, consists of pro-Chinese accounts dispersed across different platforms that post political spam and continuously regenerate new accounts after existing accounts are taken down. As described in Graphika’s reports on the network, the participating accounts present on-platform with a wide range of personality types. The vast majority of Spamouflage accounts have Western or other non-Chinese names and accompanying profile photos, and post near-daily videos on a range of geopolitical topics with narratives aligned with those of the Chinese government. However, most posts achieve low levels of engagement. In 2019, accounts in this network were linked to accounts targeting the Hong Kong protests, which Facebook and Twitter attributed to PRC state-controlled actors.

In early 2020, Spamouflage turned its attention to COVID-19, spreading content that praised Chinese efforts against the virus and highlighted the shortcomings of Western countries’ handling of the pandemic, especially events in the US. In February 2021, a Graphika report on the network assessed that China’s COVID-19 response had become a key theme stressed by accounts in the network on Twitter, YouTube, and Facebook, and that this cross-platform activity had started to reach audiences outside of its own network. The report also showed Chinese diplomats overseas using their social media accounts on these platforms to promote Spamouflage video productions. When it was reported that Sinopharm and Sinovac vaccines had low efficacy, the Spamouflage network changed focus away from China’s international “vaccine diplomacy” and started casting doubt on the effectiveness of Pfizer, Moderna, and AstraZeneca vaccines. Recently, security firm FireEye also reported that China’s inauthentic networks have spread to a variety of smaller platforms and expanded to languages outside of English and Chinese. Notably, one goal of this network, albeit an unsuccessful one, was to mobilize in-person protests in the US in response to COVID-19.

Around the same time, in early 2020, that the cross-platform inauthentic campaign argued that China’s COVID-19 response was more effective than the West’s, several Chinese state media outlets began publishing articles questioning the validity of
foreign research on the virus and the virus’s alleged origins in the United States.\textsuperscript{168} In January 2021, the \textit{Global Times}, CCTV, Xinhua News, and \textit{China Daily} reported on a number of older adults who died in Norway after receiving the Pfizer COVID-19 vaccine.\textsuperscript{169} While health officials in Norway clarified that this was largely due to non-vaccine-related reasons, such as illness and frailty, Chinese state outlets published articles calling the safety of the Pfizer vaccine into question and suggesting countries “broaden” their vaccine supply to include Chinese vaccines (see timeline below).\textsuperscript{170}

\begin{itemize}
\item **January 15, 2021:**
Norway Health Officials
Say vaccine may be too risky for those with severe frailty or the terminally ill.
Norwegian Medicines Agency (NMA) says: “We are not alarmed by this”

\item **Global Times**
Article and Facebook post says Chinese health experts call for the suspension of Pfizer after elderly deaths in Norway.

\item **CCTV & Xinhua & China Daily**
Identical Facebook posts reporting that “an adverse reaction… may be responsible” according to the NMA.

\item **January 16, 2021:**
Xinhua News
“Norwegian health authorities say deaths among elderly likely related to vaccine”

\item **January 19, 2021:**
Global Times Article, January 19
Chinese health experts advise Australia to halt approval for Pfizer vaccine following Norwegian deaths

\item **January 19, 2021:**
Global Times
“Chinese health experts advise Australia to halt approval for Pfizer vaccine following Norwegian deaths”
\end{itemize}

Figure 4.15: Timeline of the Norway story and its reporting in various Chinese state-media outlets.
4. Actors & Tactics

4.3.3 Iran

Iran had also been an early and active participant in the global conversation about the virus, using it to bolster the country’s political positions and to denigrate the US in a variety of ways. Graphika was able to attribute one covert campaign that attempted to influence information about COVID-19 to Iranian actors. The campaign occurred from approximately February 2020 to April 2020 and was conducted by the International Union of Virtual Media (IUVM), which has been directly linked to the Iranian government by the US government. An October 2020 press release from the US Department of the Treasury stated that IUVM is “owned or controlled” by the Islamic Revolutionary Guard Corps (IRGC), and the US placed sanctions on the IRGC and related entities for attempting to influence the 2020 US election.

The covert IUVM campaign used social media accounts and independent-seeming websites to launder cartoons, memes, and articles with pandemic-related narratives. Specific narratives promoted included ones suggesting that the US had purposely caused the pandemic; that the US was waging a biological attack via COVID-19 to beleaguer China; that Trump was mishandling the pandemic, and the US was suffering extensively because of him; and that Western media outlets were lying about the COVID-19 situation within Iran.

The social media assets that the IUVM operation appeared to have created included Facebook News Pages that seem to have been targeting different countries, as well as multiple Instagram accounts.

One Instagram account exclusively posted content from multiple IUVM channels. The operators of the account attempted to reach out to high-profile users, such as Ivanka Trump and Usain Bolt, but those efforts were unsuccessful, and it ultimately amassed only 485 followers.

Some of the key narratives in the covert IUVM operation—that the US was to blame for COVID-19 and its role in the pandemic was not to be trusted—were mirrored by overt official statements and actions by Iran. Some of the overt messaging was even used in the covert IUVM operation: for example, in March 2020, after Supreme Leader Ali Khamenei alleged that COVID-19 could be a biological attack, IUVM Press, one of the IUVM websites, quoted him in an article. Many Iranian officials, including Khamenei, Major General Hossein Salami, and IRGC Brigadier General Gholamreza Jalali, claimed it was possible that COVID-19 was a biological attack or had inorganic state-sponsored origins, explicitly or implicitly pointing to the US as the likely culprit. Former officials like former president Mahmoud Ahmadinejad also echoed these claims. Government actions, such as Khamenei’s ban of the import of American and British vaccines into Iran and claim that they were “completely untrustworthy,” likely served to further hit home the narrative that the West’s role in the pandemic was somehow nefarious.

By repeating similar narratives across a variety of channels, both explicitly at-
4.3. Foreign Actors

Figure 4.16: Top, a February 24, 2020, article from the IUVM Press website. Bottom, images posted on February 25, 2020, by the IUVM Pixel website.174
4.4 Conclusion

The figures and categories described in this chapter do not comprise a comprehensive list of all actors or archetypes who spread COVID-19 vaccine mis- or disinformation in 2021, but instead represent the actors and tactics that the Virality Project observed as highly prominent in the anti-vaccine rhetoric online. There have been other studies, such as the Center for Countering Hate’s “Disinformation Dozen” report, which used a different methodology to highlight the actors they focused on. Rather, this chapter highlights important groupings of actors that were active and frequent participants in spreading COVID-19 vaccine misinformation. We chose to highlight specific figures within their categories based on their prevalence in high-engagement incidents that had significant reach, or prominence that justified their inclusion in the Virality Project weekly briefings. These are figures who have strong parallels to long-standing participants in the discourse around vaccines, and whose online presence is critical for stakeholders working across government, research, public health, and social media to be aware of and understand—not only as archetypes or even as individuals, but also as nodes in a broader network committed to undermining COVID-19 vaccine uptake.
Notes

1. T. Smith, “Vaccine Rejection and Hesitancy: A Review and Call to Action.”
2. T. Smith.


23. See, e.g., Candace Owens (@RealCandaceO), “Nuremberg codes stipulate that consent must be voluntary. Telling a child they cannot go to school, firing people from work, or preventing their travel is NOT voluntary consent. Threatening someone’s livelihood unless they take an experimental vaccine is NOT voluntary consent,” Twitter, July 7, 2021, http://web.archive.org/web/20210710105514/https://twitter.com/RealCandaceO/status/1412880295368478723; Earthley (@earthley), “Thank you, Candace!

25. Anesthesiologist Dr. Christopher Rake was escorted out of his workplace at UCLA for refusing to be vaccinated. UCLA Health (@UCLAHealth), “These comments do not represent the views of UCLA Health. Unvaccinated people are more likely to contract COVID-19 & we encourage employees to be vaccinated. We adhere to the state public health order requiring health care workers to be vaccinated or undergo regular testing,” Twitter, August 21, 2021, https://twitter.com/UCLAHealth/status/1432754622695817218; CBSLA Staff, “Willing To Lose Everything: Unvaccinated Anesthesiologist Escorted From UCLA Medical Facility, Placed On Unpaid Leave,” CBS Los Angeles KCAL, October 7, 2021, https://losangeles.cbslocal.com/2021/10/07/willing-to-lose-everything-unvaccinated-anesthesiologist-escorted-ucla-medical-facility-placed-on-unpaid-leave/.


41. D. J. Mercola, “This good news....”


47. Dr. Simone Gold (@drsimonegold), “The founder of Shake Shack says his company will require proof of vaccination for both employees AND customers. I encourage everyone to boycott this business. No corporation or government has a right to demand to see your private health documents,” Twitter, August 1, 2021, https://twitter.com/drsimonegold/status/1421750765866029058; Dr. Simone Gold (@drsimonegold), “WOW: @RoyalCaribbean cruises is now requiring you to get an experimental vaccine
to book a cruise with them. I will be boycotting the company instead. Who else is with me? #BoycottRC," Twitter, March 20, 2021, https://twitter.com/drsimonegold/status/1373361332062412810.


53. Stecklow and Macaskill.


58. Sones.

59. Stecklow and Macaskill, “The Ex-Pfizer Scientist Who Became an Anti-Vax Hero.”


75. Holmes Lybrand, “Fact check: Candace Owens falsely claims CDC proposed putting high-risk people in camps during the pandemic,” Facts First, August 11, 2021, https:

77. Sir Patrick Vallance (@uksciencechief), “Correcting a statistic I gave at the press conference today, 19 July. About 60% of hospitalisations from covid are not from double vaccinated people, rather 60% of hospitalisations from covid are currently from unvaccinated people,” Twitter, July 19, 2021, https://twitter.com/uksciencechief/status/1417204235356213252.

78. Candace Owens, “This is shocking! 60% of people being admitted to the hospital with #COVID19 in England have had two doses of a coronavirus vaccine, according to the government’s chief scientific adviser,” Facebook, July 19, 2021, https://www.facebook.com/159351817405271/posts/5768837649854055.


103. Alex Berenson (@AlexBerenson), “Fact check: True. Our daughter got MMR and other shots today. (Which is not to say I’m going to be running to the head of the line for a rushed COVID vaccine. You go first! Have at it.),” Twitter, June 15, 2020, https://web.archive.org/web/20210822043632/https://twitter.com/AlexBerenson/status/1272592613460381698.


106. Alex Berenson (@AlexBerenson), “As the kids say, lolcatz. Though Moderna avoids calling what it does gene therapy (in part because regulators historically look harder at gene therapy than other biologics), the scientific literature is less discreet. Katalin Kariko invented the idea back in 1990; she would know,” Twitter, March 14, 2021, http://web.archive.org/web/20210822182205/https://twitter.com/AlexBerenson/status/1371249209777721345.


108. Thompson, “The Pandemic’s Worst Man.”


110. Kertscher.


127. By comparison, RFK Jr.’s anti-vaccine organization Children’s Health Defense boasts 173,000 followers on its Facebook Page alone, as of September 2021.


141. Timberg, “Scientists said claims about China creating the coronavirus were misleading. They went viral anyway.”


144. Propaganda has long been classified on a white-to-black spectrum. Grey media are state-backed fronts, where ownership is questionable. This media is not clearly or demonstrably attributable to the state, but does not completely obscure or misdirect its source or intent.


149. “July 2021 Coordinated Inauthentic Behavior Report.”

150. “July 2021 Coordinated Inauthentic Behavior Report.”

151. “July 2021 Coordinated Inauthentic Behavior Report.”


165. Nimmo, Hubert, and Chen.


172. Nimmo et al., “Iran’s IUVM Turns To Coronavirus.”

174. Nimmo et al., “Iran’s IUVM Turns To Coronavirus.”

175. Nimmo et al.


179. CCDH, “The Disinformation Dozen.”
Chapter 5

Platform Policies for Anti-Vaccine Content

For years, researchers and journalists have raised flags about anti-vaccine misinformation on social media platforms and its potentially deadly offline consequences.\(^1\) Yet building effective platform policies to combat misleading health information is difficult. Over a year into the COVID-19 pandemic, when reporters asked President Biden if he had a message for social media platforms, he responded that Facebook and other platforms were “killing people” by enabling the spread of COVID-19 misinformation.\(^2\) The president later walked back his remarks, but the initial accusation revived previous discussions about what responsibility social media platforms bear in addressing vaccine-related misinformation. More recently, in October 2021, Facebook whistleblower Frances Haugen gave Senate testimony on the harms of misinformation,\(^3\) including vaccine misinformation, stating that “Facebook consistently resolved those conflicts in favor of its own profits.”\(^4\)

This chapter provides an overview of the content moderation approaches that various social media platforms have taken to address false and misleading information about vaccination, as well as a timeline of policies specific to the COVID-19 vaccines. It highlights potential gaps that may enable anti-vaccine content to continue to spread. It addresses the challenges of moderating health misinformation: the dynamics of evolving scientific consensus, limited facts, the subjectivity of personal experiences, and the difficult balance between minimizing harmful misinformation with maximizing freedom of expression.

5.1 Platform Moderation

Platform policies have evolved, responding first to general anti-vaccine content and then to health and vaccine misinformation surrounding the COVID-19 pandemic.
5.1. Platform Moderation

5.1.1 Pre-COVID-19 Vaccine-Related Policies

In 2019, several major social media platforms created vaccine-related policies in response to rising misinformation about the safety of routine childhood immunizations such as the measles, mumps, and rubella (MMR) vaccine.\(^1\) The policies followed major measles outbreaks in New York City, American Samoa, and elsewhere. During and after the outbreaks, scientists and congressional leaders sought accountability from the platforms, inquiring about the extent to which vaccine hesitancy among impacted communities had been exacerbated by misinformation on their products.\(^3\)

The pre-COVID-19 vaccine-related policies developed by the platforms focused primarily on reducing the discoverability of anti-vaccine content and providing links to authoritative sources. They removed anti-vaccine groups from recommendations, downranked groups and prominent health misinformation domains in search, and labeled content with links to the CDC, WHO, or other public health institutions. Anti-vaccine activists, influencers, and organizations were prohibited from running ads to target specific types of users in their attempts to grow their followings on Facebook Pages and Groups.\(^6\)

While Facebook’s early policies reduced the discoverability of false and misleading health information, the platform struggled to strike a balance between reducing the prevalence of Groups that pushed demonstrably false health claims (“vaccines cause autism”) while still allowing Groups that declared themselves to be opposed to vaccines due to concerns about personal freedom and political views (“school vaccine requirements are unjust”). In response, some of the longtime pseudomedical-focused communities began to refocus their public messaging on liberty.\(^7\)

In the early period of the newly articulated policy, Facebook often did not action individual posts within Groups. Yet some of the claims shared in these Groups led to demonstrable direct harm. In February 2020, for example, one mother shared a post about her son’s flu diagnosis to “Stop Mandatory Vaccination,” a private Group with 178,000 members. Group members discouraged her from administering a doctor-prescribed medication and instead suggested she try homeopathic remedies. In this case, the mother’s four-year-old son died.\(^8\) Reducing Group and Page discoverability perhaps limited the growth of the anti-vaccine movement, but it appeared to do little to curb the spread of vaccine misinformation within existing Groups, or between adjacent Groups (such as wellness or QAnon communities). However, when the

\(^{1}\)This chapter refers to Facebook/Instagram, Twitter, YouTube, TikTok, and Pinterest as the major social media platforms, defined as such because of their prominence in US communities. These specific major social media platforms were chosen based on two criteria: first, there are tools to research and collect data from these platforms; second, these platforms are centers for social discussion and news. Platforms like LinkedIn, which is widely used but not necessarily at the forefront of medical discussion, are not included in this overview.

The policies were reviewed during the Virality Project’s analysis period and reflect how they stood during that period. Platforms may have changed their policies by the time this report is published.
COVID-19 pandemic began, the platforms quickly took steps to broaden these 2019 policies in an effort to minimize the impact on individual and public health.

### 5.1.2 COVID-19-Related Policies

The timeline in Figure 5.1 on the next page demonstrates how COVID-19 vaccine policies evolved over time, and became more granular, as platforms became far more proactive in addressing health misinformation during the pandemic. For example, when Pfizer and Moderna’s COVID-19 vaccines received emergency use authorization in late 2020, platforms quickly introduced policies to address false and misleading information specific to those vaccines. On December 3, 2020, Facebook announced it would “start removing false claims about these vaccines that have been debunked by public health experts.”

Twitter, Google, TikTok, and Nextdoor all updated their policies in quick succession to explicitly address misleading information about the COVID-19 vaccine.

However, there is often a gap between the existence of platform policies—even newly enacted ones—and the enforcement of the policies.

#### 5.1.3 Policy Enforcement: Remove, Reduce, Inform

The major social media platforms generally use a “Remove, Reduce, and Inform” framework for content moderation. While that specific terminology was publicly introduced by Facebook in 2016, discussing its facets is helpful for understanding the types of policy enforcement that platforms use more broadly. In the tables below, we use the framework to assess the policies of six platforms.ii,iii

| Yes. The platform explicitly states the policy publicly in either its community standards or its blog posts. |
| Partial / Not Explicit. The platform partially fulfills the policy, but the policy does not apply in all cases, or may apply but is not explicitly stated. |
| No. The platform does not have the policy stated publicly in either its community standards or its blog posts. |

---

iiIt is important to note that check marks in these charts do not represent the Virality Project’s endorsements of a policy. Some platforms do not need certain policies, while some policies—such as removing content that runs counter to public health experts—are not completely reliable. For example, experts have at times reversed their health recommendations, leading them to make statements that would potentially be actioned upon by platform policies based on prior health recommendations. Tim Elfrink, Ben Guarino, and Chris Mooney, “CDC reverses itself and says guidelines it posted on coronavirus airborne transmission were wrong,” Washington Post, September 21, 2020, https://www.washingtonpost.com/nation/2020/09/21/cdc-covid-aerosols-airborne-guidelines/.

iiiThe exact policy language and criteria for the tables can be found in Appendix D: Table Sources.
5.1. Platform Moderation

Figure 5.1: Timeline of policies that address content related to medical or vaccine misinformation. See Appendix C: Timeline Sources for a breakdown of methodology and individual sources.
5. Platform Policies for Anti-Vaccine Content

**Remove**

The strictest platform moderation intervention is to *remove* content and accounts. To determine whether or not to remove content for violating a policy related to false and misleading claims about COVID-19 vaccines, platforms chose public health expert consensus as the basis for evaluating the material. For example, Facebook’s Community Standards state that the company “remove[s] misinformation when public health authorities conclude that the information is false and likely to contribute to imminent violence or physical harm.” In addition to removing content, platforms may also suspend accounts that consistently share false claims.

Table 5.2: Platform “Remove” policies.

<table>
<thead>
<tr>
<th>Policy Description</th>
<th>Facebook</th>
<th>Twitter</th>
<th>YouTube</th>
<th>TikTok</th>
<th>Pinterest</th>
<th>Nextdoor</th>
<th>Reddit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Removes content that could lead to physical harm</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Removes content that runs counter to public health expert consensus</td>
<td>✔️</td>
<td>✗️</td>
<td>✔️</td>
<td>✗️</td>
<td>✔️</td>
<td>✔️</td>
<td>✗️</td>
</tr>
<tr>
<td>Clear strike policy for removing recurring actors</td>
<td>✗️</td>
<td>✔️</td>
<td>✔️</td>
<td>✗️</td>
<td>✗️</td>
<td>✗️</td>
<td>✗️</td>
</tr>
</tbody>
</table>
### 5.1. Platform Moderation

**Reduce**

Platforms can choose to reduce the spread of content by making it less discoverable for users. Examples might include reducing the distribution of a piece of content in a News Feed, video queue, or timeline to give time for it to be fact-checked, or removing the option to retweet or share something.\(^\text{11}\)

Demotion, which reduces the prevalence of content across the platform, is a step short of complete removal of content. Platforms apply this policy in different ways. For example:

- On Facebook, if fact-checked content is determined to be false, altered, or partly false, it will appear lower in News Feed; on Instagram, it will not be featured on the Explore page, where people can discover posts from users they don’t follow.\(^\text{12}\)

- On Reddit, subreddit communities can be subject to the “quarantining” policy. Introduced in 2015, this policy prevents people from stumbling upon problematic subreddits. A subreddit can be quarantined if the average Redditor finds the content “highly offensive or upsetting.”\(^\text{13}\) In September 2021, the policy was applied to 54 subreddits associated with COVID-19 denial, but only after a large-scale protest on the site led by popular subreddits.\(^\text{14}\)

<table>
<thead>
<tr>
<th>Search results for vaccine-related terms point to authoritative sources</th>
<th>Facebook</th>
<th>Twitter</th>
<th>YouTube</th>
<th>TikTok</th>
<th>Pinterest</th>
<th>Nextdoor</th>
<th>Reddit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Restricts certain forms of vaccine ads</th>
<th>Facebook</th>
<th>Twitter</th>
<th>YouTube</th>
<th>TikTok</th>
<th>Pinterest</th>
<th>Nextdoor</th>
<th>Reddit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reduces groups, pages, and content that promote vaccine misinformation</th>
<th>Facebook</th>
<th>Twitter</th>
<th>YouTube</th>
<th>TikTok</th>
<th>Pinterest</th>
<th>Nextdoor</th>
<th>Reddit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>
5. Platform Policies for Anti-Vaccine Content

**Inform**

Platforms can inform users by adding context and reliable information to posts about the COVID-19 vaccine. These policies elevate “trusted sources” by, for example, designating specific pages and tabs on their platform for sharing fact-checked information with the latest updates and guidance about COVID-19 vaccines, or partnering with public health or nonprofit organizations to amplify their content.¹⁵ They also provide additional context to vaccine-related posts through labels that point to trusted sources, or, in the case of posts that are misleading or false, to fact-checkers.

Table 5.4: Platform “Inform” policies. An important difference between the second and third policies in this table is what content triggers the label. For “applies banners & labels,” the content does not need to be about COVID-19 but must be considered misleading or false. For “universal label policy,” the content simply needs to mention COVID-19 to receive the label; the platform makes no judgment about the truth of the content.

<table>
<thead>
<tr>
<th>Provides COVID-19 information hub on platform</th>
<th>Facebook</th>
<th>Twitter</th>
<th>YouTube</th>
<th>TikTok</th>
<th>Pinterest</th>
<th>Nextdoor</th>
<th>Reddit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applies banners &amp; labels to misleading COVID-19-related post content</td>
<td>✔️</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Universal label policy for all content mentioning COVID-19</td>
<td>✔️</td>
<td>✗</td>
<td>✔️</td>
<td>✔️</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
</tbody>
</table>

5.2 **Nuances and Gaps in Vaccine-Related Policies**

There has been widespread public debate about the extent to which platforms should moderate COVID-19 vaccine misinformation; as with election-related misinformation, the policies are described as overreach by some, and inadequate by others. Trying to moderate vaccine-related misinformation has posed unique challenges for platforms. Three recurring types of posts in particular illustrate the challenge
of deciding whether, and how, to moderate: (1) personal anecdotes, (2) medical freedom claims, and (3) decontextualized headlines and studies.

### 5.2.1 Personal Stories

As discussed in Chapter 3: Narratives, content that consists of personal stories and experiences is difficult to verify and fact-check. Much of the anti-vaccine content that VP analysts observed was linked to hard-to-verify personal stories, often about someone who allegedly died from or was injured by the vaccine. These types of stories have an established history within anti-vaccine communities on social media as emotionally evocative content with high virality potential; for example, anti-vaccine organizers have long leveraged stories of Sudden Infant Death Syndrome to encourage people to join their anti-vaccine groups and learn “the truth.” These types of stories place platforms in a challenging position: a post describing an individual’s purported adverse reaction to the vaccine might not be verifiable. The authoritative sources that platforms usually turn to, such as the CDC or WHO, do not weigh in on every personal account of alleged vaccine injury.

Platform policies on personal anecdotes vary. Facebook, Twitter, and, most recently, YouTube state clearly in their policies and blog posts that they will not take action on personal stories. However, Facebook recently added a policy to remove accounts that are dedicated to sharing “personal anecdotes or news events of severe adverse events, either in hyperbolic terms or without context.” TikTok and Reddit do not explicitly address how they handle personal anecdotal content on their platform.

While these policies allow for free expression of people’s experiences, there are consequences: If personal stories do not fall under the umbrella of moderatable content, they are not subject to the circuit breakers or friction that comes with fact-checking or labeling. Platforms could potentially mitigate this by applying labels more liberally and leveraging the universal labeling policy to add context more broadly. For example, they could label any post that mentions VAERS with information about what VAERS data shows—and what it doesn’t show.

### 5.2.2 Medical Freedom Claims

Major social media platforms have also struggled with how they wish to moderate medical freedom claims, particularly because there are degrees of disagreement among the public about which vaccines should be required. Prior to the COVID-19 pandemic, the medical freedom movement pushed legislation across the country to fight school vaccination requirements. Users took to social media platforms to promote or fight against vaccine-related bills in New York, New Jersey, Ohio, and California. Political activism and legislation related to vaccines, and narratives about medical freedom, are tricky for platforms to moderate because much of the commentary involves statements of political opinion. In 2019, platforms’ policy
choices demonstrated that they consciously chose to avoid moderating the political debate about vaccines: medical freedom pages were left up, while falsifiable health-related claims (such as “vaccines cause autism” or “vaccines cause SIDS” were moderated.\textsuperscript{21}

While platforms have used “reduce” enforcement mechanisms on demonstrably false anti-vaccine health claims, content espousing medical freedom principles remains a vector for drawing people to health-related anti-vaccine movement content. For example, a search on Facebook for anti-vaccine Groups will not autocomplete or suggest related Pages and Groups or search terms. However, a search for medical freedom Groups, such as the state-level activist group “Freedom Keepers,” will. Although these groups use freedom of choice as their primary framing and branding, much of their content aims to reduce vaccine uptake (see Figure 5.2).

Figure 5.2: Left, a search for “anti vaccine” on Facebook does not auto-complete. Center and right, searches of common medical freedom terms on Facebook will auto-suggest related terms. Search performed January 25, 2021.

### 5.2.3 Sensationalized or Decontextualized Content

Sensationalized or decontextualized headlines and misleading presentation of facts and figures also pose a challenge for platform policies. These posts may contain true information, but when stripped of context, that information can be weaponized to support a misleading or outright false narrative. While such headlines are often linked to an article containing more information, multiple studies suggest that most Americans do not read past a headline before sharing a link on social media.\textsuperscript{22} Headlines thus play an important role in what information is spread. This vector for misinformation is often compounded when users share screenshots, which usually contain only a title and a few lines, but rarely link to the original source.

In other instances, decontextualized facts can be manipulative depending on their recontextualization by the poster (see Figure 5.3 on the next page). It is not sufficient
for platforms to scan for the shared memes; they must also monitor the post’s framing, which is difficult to do consistently at scale.

Figure 5.3: Top, a user shares a meme and puts the number of breakthrough cases in context to help decrease the sensationalization of the meme. Bottom, a user shares the same meme but frames it as an example of how the vaccine is harmful and experimental.

Hyperbolic, decontextualized, or misinterpreted facts are not explicitly acknowledged in most platform policies. Facebook is the exception. After an update of its vaccine misinformation policies on February 8, 2021, the company has since added a policy that addresses sensationalist or alarmist content: it will reduce the distribution and limit visibility of “content that otherwise does not violate our COVID-19 or vaccine policies above but that suggests that vaccines are ineffective, sacrilegious, unsafe, or irrelevant, in exaggerated, conspiratorial, or sensational terms.”

However, determining what is “sensational” is somewhat subjective and, therefore, not always predictable. For most platforms, these posts may fall under the overarching umbrella of “misleading” content, which could lead to the post’s removal (if the figure or headline in the post is a clear misrepresentation) or to adding a label to the post (if the post’s veracity is ambiguous). However, platform policy does not always articulate what type of content justifies moderation or what kinds of labels should be used.

In practice, it can prove difficult for platforms to add context to a decontextualized or misleading interpretation of a fact or figure. In one example of this dynamic, a June 2021 study of pregnant women and COVID-19 vaccines was used to promote the
claim that vaccines cause miscarriages. Posts on various social media platforms highlighted the study’s findings about the rate of miscarriage among vaccinated pregnant women, sharing figures that did not accurately represent the study’s results. Posts circulated without the necessary context for users to interpret the findings correctly. Yet adding additional context would have made the post more confusing: longer, denser prose is less appealing than an easy graphic and challenging to condense into a tweet, making it easier for users to misrepresent the study than for others to explain why the misrepresentations are incorrect.

One of Twitter’s experimental design-based interventions offers a potential solution to this issue. In June 2020, Twitter began testing a policy that adds more friction to reshares of outside links: a dialog box pops up, asking the user if they want to read an article before sharing it. This intervention is speech agnostic—the platform doesn’t weigh in on what the text actually says—but it helps the user be more involved, rather than blindly sharing misinformation. A few months after launching the product change, Twitter released information on the intervention’s performance. The platform found that after seeing the prompt, users were 33% more likely to open the article before retweeting it—potentially making for more informed tweeting along with decreased spreading of sensationalized headings.

5.3 Authoritative Sources and the Deplatforming Debate

Surfacing authoritative content on social media platforms is a complicated task. To address it, platforms rely primarily on two approaches: first, raising up authoritative voices by elevating them in search results or on dedicated information carousels or tabs, and second, deplatforming those who repeatedly spread false and misleading information.

5.3.1 Elevating Authoritative Sources

Platforms have implemented a number of product changes to elevate authoritative information. For example, Pinterest, one of the earliest platforms to tackle vaccine misinformation, only surfaces content from leading public health institutions, such as the CDC and WHO (see Figure 5.4 on the facing page). When users search for vaccine-related content, only credible institutional sources come up. Similarly, both Twitter and Facebook provide links to authoritative sources to users searching for key vaccine-related words.

However, as COVID-19 showed, institutional sources did not always produce particularly timely content for platforms to surface, particularly in areas in which consensus took time to develop. Sometimes, in fact, institutions such as the CDC
5.3. Authoritative Sources and the Deplatforming Debate

![Figure 5.4: A search for medical freedom on Pinterest pulls up only credible vaccine-related information from trustworthy sources.](image)

and WHO were reticent to take early positions on pressing COVID-19-related questions that the public wanted answers to, creating a data void; the early debate about masks was one such issue. Recognizing that there was not always content from authoritative institutions to point to, and that it was often not the most compelling content to share or retweet, some platforms attempted to additionally elevate the authoritative voices of individual accounts. Twitter helped elevate these voices by getting experts from public health organizations and frontline doctors verified with blue check marks. However, while this diversified away from solely institutional authoritative voices, it was not a panacea; as researchers witnessed during the 2020 election, a number of accounts that spread false and misleading information came from large accounts that were verified. For example, prominent anti-vaccine activist Robert F. Kennedy Jr. is verified on Twitter, as was Dr. Naomi Wolf before she was deplatformed on June 5, 2021.

5.3.2 Deplatforming Harmful Sources

Deplatforming is the most severe action that a platform can take against an individual account. This approach also seems to have gained more prominence after the 2020 election, when, along with 70,000 other accounts on Twitter, President Trump was deplatformed. Against this backdrop, VP analysts observed the dynamics around when and why social media companies elected to deplatform anti-vaccine accounts, and the impacts, expected and unexpected, of deplatforming.
5. Platform Policies for Anti-Vaccine Content

Facebook announced updates to its vaccine-related policies in February 2021, as the vaccine rollout opened up to more groups in the US. Included in this update was the policy that “Pages, Groups, profiles, and Instagram accounts that repeatedly post misinformation or coordinate harm...related to COVID-19, vaccines, and health may face restrictions, including (but not limited to) reduced distribution, removal from recommendations, or removal from our site.” A few days after the updated policy was announced, Robert F. Kennedy Jr. and Del Bigtree’s High Wire Talk Instagram accounts were removed (799,000 and 162,000 followers, respectively).

Deplatforming on its own may not achieve reduced spread of false and misleading information by specific actors. First, platforms do not exist in a vacuum. While a platform can make decisions about an account within the sphere it controls, these actors may still have influence across the larger information ecosystem, especially if they are not removed from all other major platforms. Content shared by a prominent account on one platform (e.g., YouTube) can hop over to another platform (e.g., Facebook) whether or not that account has been banned from Facebook. As alt-platforms with laxer moderation policies, such as Rumble and Parler, proliferate, this issue will grow.

Second, anti-vaccine accounts may protect themselves against the effects of deplatforming by migrating to alternative platforms. Deplatforming would logically cause an account to lose the community the person has built on large platforms, but by preemptively using their mainstream platform accounts to highlight their accounts on alternative platforms, such as Telegram or MeWe, anti-vaccine influencers diminish the power of deplatforming. In the aftermath of Robert F. Kennedy Jr.’s and Del Bigtree’s Instagram deplatforming, Erin Elizabeth of Health Nut News took to Instagram Live to discuss the major deplatforming and encourage her 121,552 followers at the time to migrate to Telegram. Virality Project analysts traced 49 Facebook and Instagram accounts that promoted links to their Telegram channels (see examples in Figure 5.5 on the next page). Although some large accounts had been deplatformed, they were able to move a portion of their Instagram and Facebook followers to these alternative platforms, and accounts that remained active on Instagram and Facebook were able to continue to build their followings elsewhere. This is one cost associated with delaying deplatforming of recurring actors. Additionally, this migration to alternative platforms may lead to “cross-pollination” with other extreme views: anti-vaccine actors and content may be more likely to mix with far-right and conspiratorial fringes. Right-wing extremists and white supremacist groups have consistently discussed and spread COVID-19 conspiracy theories on alt-platforms.
5.4 Conclusion

Moderating medical misinformation comes with pitfalls and controversy, and thus far the anti-vaccine movement has proven both popular and savvy enough to persist on major social media platforms despite attempts to minimize its impact. While progress has been made since platforms first developed vaccine-related policies in 2019, clear gaps in platform policy exist with respect to moderating vaccine-related content, including posts that employ personalized stories, medical freedom claims, and misleading headlines and statistics. In addition, policies about the actions platforms can take to suppress content, promote trusted voices, and deplatform accounts vary widely from platform to platform and are still not enforced consistently, both within and across platforms.

It should be noted that understanding the impact of platform policy is limited by what information is publicly available. It is crucial that platforms provide more transparency on each moderation approach and allow external researchers the ability to independently verify the success and impacts of these interventions.
Notes


7. This strategy of adjusting messaging to focus on outwardly communicating themes of liberty instead of the discredited health claims had been previously observed on Twitter. DiResta and Locan, “Anti-Vaxxers Are Using Twitter to Manipulate a Vaccine Bill.”

8. Zadrozny, “On Facebook, anti-vaxxers urged a mom not to give her son Tamiflu. He later died.”


26. Twitter Comms [@TwitterComms], “We shouldn’t have to say this, but you should read an article before you Tweet it. ... So, we’ve been prompting some people to do exactly that. Here’s what we’ve learned so far,” September 24, 2020, https://twitter.com/TwitterComms/status/1309178715717369856.


31. Connett, “Naomi Wolf Banned from Twitter for Spreading Vaccine Myths.”


33. “COVID-19 and Vaccine Policy Updates & Protections.”

34. Center for Countering Digital Hate [@CCDHate], “Del Bigtree is still broadcasting dangerous misinformation about vaccines and Covid to half a million followers on @Facebook, @Instagram and @Twitter. It’s time they followed @YouTube and removed his accounts,” July 30, 2020, https://twitter.com/CCDHate/status/1288769298304061440/photo/2.


Chapter 6

Recommendations for Stakeholders

No single entity has the necessary visibility, capability, and reach to mitigate viral mis- and disinformation on its own. Instead, and in recognition of the fact that mis- and disinformation will never be “stopped,” a whole-of-society effort is needed in which stakeholders build robust and persistent partnerships to ensure that significant high-harm claims can be addressed as they arise. The Virality Project sought to do just that by bringing together four types of stakeholders: (1) research institutions, (2) public health partners, (3) government partners, and (4) platforms. Our recommendations recognize the collective responsibility that all stakeholders have in mitigating the spread of mis- and disinformation and empowering citizens to make informed decisions—about vaccination specifically, but also more generally given the structure and dynamics of the information ecosystem.

6.1 Recommendations for Research Partners

The Virality Project offers an early template for structuring interaction between research institutions and nonacademic stakeholders (including government entities, health practitioners, and private companies). In early conversations with the Virality Project, government and civil society entities tasked with mitigating false and misleading claims expressed dissatisfaction with the often one-sided relationship they had with researchers. The Virality Project tried to avoid this experience for its broad array of stakeholders by prioritizing the communication of tailored and actionable information through weekly briefings and regular check-in sessions with key public health partners.

This sentiment goes both ways: Virality Project research institutions felt that they struggled at times with sporadic communication from civil society organizations that had better and far more nuanced visibility into their communities than could be achieved through list-monitoring. This created limitations and potential for bias.
An area that required ingenuity was creating a framework for facilitating the *intake* of tips from civil society and government partners. Structuring tip intake remains a hurdle, as many civil society organizations are understaffed or unfamiliar with workflows like task management or ticketing software. However, their tips are often highly valuable, so overcoming this challenge is a priority for future efforts.

**Define roles and responsibilities based on competencies.**

Collaboration across research institutions can enable researchers to share their comparative skills and advantages to build a powerful research team. Partners should lean into their own strengths and seek out partners that complement those strengths. For example, the Virality Project’s stakeholders often asked the partnership for assistance with counter-messaging. Because the Virality Project partner organizations did not possess the necessary health expertise to generate such messaging, VP briefings instead drew on material from outside resources, such as the Public Good Project’s Stronger campaign,\(^1\) which fights mis- and disinformation in the scientific, medical, and vaccine spheres. Future partnerships should consider incorporating partners focused on counter-messaging specifically.

**Streamline tip line processes for civil society and government partners.**

Set up an efficient channel for intaking external tips. Many external partners are operating with limited resources, so it is important to minimize the barriers for entry on an external tip line. The Virality Project often had to leverage informal exchanges, such as Zoom meetings or calls with our partners, to receive the tips verbally or encourage additional reporting. In future projects, external reporting channels should be strengthened via an easier means of reporting and increased access to the reporting channels, especially for partners on the ground (such as health practitioners or government health officials).

**Establish a constant feedback loop with partners.**

Set up a collaborative feedback process with external stakeholders to understand the value they derive from the research and to constantly improve the work’s impact. For example, several VP stakeholder groups indicated that they had a need for a deeper understanding of mis- and disinformation targeting Spanish- and Chinese-language communities. The VP team was then able to prioritize that need by assigning analysts to focus on activity in these languages.

### 6.2 Recommendations for Public Health Partners

One goal of the Virality Project was to provide situational awareness of key vaccine-related narratives to public health partners. Frontline public health stakeholders are uniquely positioned to counter or respond to misinformation because of their trusted position and their interest in increasing vaccine uptake. However, they often
have minimal or no access to relevant social media data, and need help identifying and understanding what narratives are going viral, and where. Various scholars and organizations have already outlined best practices for communication around public health issues affected by mis- and disinformation. Based on this project’s experience, the Virality Project offers some additional insights for public health communicators.

Focus on themes, not individual incidents on social media.

Individual anecdotes are often unverifiable, and attempting to debunk or counter them one at a time may prove counterproductive, or be perceived as dismissing a person’s individual experience. Consider addressing overarching themes instead.

Be transparent about known unknowns.

Be transparent about the degree to which information about a health topic is early or speculative. Scientific consensus appropriately evolves as more information becomes available; if consensus changes, transparency remains the best policy. In today’s social media information environment, those who seek to undermine confidence in institutions can easily obtain past statements, tweets, or video of official stances and use the material to create a false perception that the institution is attempting to cover something up or is flip-flopping. Public health communicators might consider takeaways from Wikipedia’s approach to contest issues, which offers a publicly visible edit history that shows shifts in consensus over time and has a stated process for deciding how to handle interpretation disputes.

Mix personal stories about the vaccine’s benefits with support from data.

Vaccine-related messaging is more effective when qualitative messaging and on-the-ground stories are coupled with easy-to-understand quantitative evidence. During a July 15, 2021, panel with the Virality Project, US Surgeon General Vivek Murthy discussed the importance of vaccination by sharing his own story about COVID-19 (including the death of a close relative) alongside data around the effectiveness of the vaccines. Where data is available, it is critical to provide context and information about the source and meaning of that data. The Virality Project repeatedly saw data misrepresented or outright manipulated by anti-vaccine activists and influencers to support their narratives.

6.3 Recommendations for Government Partners

COVID-19 continues to impact every state and county across the US. There is an ongoing need for federal, state, and local leaders to continue to build out their multistakeholder partnerships and messaging to ensure consistency and effectiveness across communities. While the federal government (through DHHS, the CDC, and the Surgeon General) has ramped up its engagement and communications, more can
be done moving forward. There are several areas where government officials can focus to improve their ongoing response to mis- and disinformation surrounding the COVID-19 vaccines:

**Maintain clear channels of communication across all levels of government.**

Federal, state, and local government officials should coordinate real-time response to emerging mis- and disinformation. Narratives related to the COVID-19 vaccines cannot be moderated away; instead, they must be addressed directly, and accurate information must be transparently and frequently provided to the public via reputable, trusted figures at all levels of government. For example, as voting-related mis- and disinformation arose in the 2020 presidential election, the Election Infrastructure Information Sharing and Analysis Center (EI-ISAC) served a critical role in sharing information with the Election Integrity Partnership and pushing its rapid response analysis back out to election stakeholders across all states. There was no equivalent coordinating hub or center to support public health organizations nationwide, or to connect them to other stakeholders. Moving forward, the government should support the establishment of such an information-sharing mechanism.

**Address misinformation through collaborations with trusted voices.**

Federal, state, and local health officials should prioritize engagement with underserved communities, particularly non-English-speaking and/or marginalized communities, by focusing on empowering trusted communicators within those communities. In the context of vaccine misinformation specifically, some examples of engagement best practices can be found in the Virality Project’s July 15, 2021, hosted discussion with Surgeon General Vivek Murthy, which highlighted the partnership formed between hair salons, barbershops, and the government to combat vaccine hesitancy within the Black community. During the panel, hair salon owner Katrina Randolph shared the impact she was able to make within her community by communicating facts and encouraging friends, colleagues, and clients to get information from health providers. Randolph’s recognition of her role as a community leader, modeling of beneficial behavior (she is vaccinated), and messaging of that behavior to her customers and community is a strong approach that might be replicated across communities.

**Implement a Misinformation and Disinformation Center of Excellence (CoE) housed within the federal government.**

In recognition of the fact that structural factors in the information environment play a role in mis- and disinformation independent of topic, the CoE would centralize expertise on mis- and disinformation within the federal government at the Cybersecurity & Infrastructure Security Agency (CISA) with its existing mis- and disinformation team. Its role would be threefold:

- **Bring together subject-matter experts conducting prospective research into emerging areas of information distortion that threaten national**
6.3. Recommendations for Government Partners

**security.** This research would be supported by close collaboration with academic researchers, nonprofit organizations, the intelligence community, and private forensic firms.

- **Coordinate counter-messaging efforts between relevant agencies.** The Center would help coordinate agencies, provide them with support in monitoring mis- and disinformation narratives (including a rumor-control mechanism), and engage relevant stakeholders.

- **Build resilience to mis- and disinformation, both within and outside government.** The CoE should spearhead initiatives to educate the American public on issues of information integrity and help start a national conversation on how we consume information.

**Establish a rumor-control mechanism to debunk nationally trending narratives.**

A Rumor Control page could serve as a central authority with information to clarify emerging narratives and get ahead of predictable tropes and narratives (see Figure 6.1). The page can also help coordinate public messaging, much like the CISA’s Rumor Control page did for the 2020 US election.10

![Figure 6.1: The cycle of rumor detection and response a Rumor Control page should follow to counteract mis- and disinformation.](image)

Operators of a Rumor Control page must exercise caution in deciding when to respond to emerging narratives, so as not to accidentally elevate narratives that would otherwise have little reach. First Draft offers five criteria to consider when determining whether to respond to an emerging anti-vaccine narrative.11 These
6. Recommendations for Stakeholders

criteria are in line with the decision-making techniques employed by the Virality Project when drafting weekly briefings:

- **Engagement**: Rumor Control pages should address misinformation that has received a high level of attention across multiple posts.
- **Audience**: While online anti-vaccine communities and accounts post vaccine misinformation incessantly, Rumor Control pages should only address a rumor found in these groups if it has spread broadly to the general public.
- **Multiple Social Media Platforms**: Rumor Control pages should consider whether misinformation has spread across multiple social media platforms.
- **Influencer/Verified Accounts**: Rumor Control pages should address misinformation that has been amplified by verified accounts and other influencers who may have a large and active following that is likely to amplify their message.
- **Large Media Outlets**: Rumor Control pages should address misinformation that has been amplified by large media outlets like cable news or online newspapers. Even if the media outlets are debunking and counteracting the misinformation, it will be important to include it on the Rumor Control page. Additionally, operators should be aware of how to present facts effectively. Using accessible language, rumor explanations should highlight a factual statement first, followed by a single sentence summarizing the rumor, and finally a deeper factual debunking of the rumor. Additional information about creating effective Rumor Control pages can be found on the Virality Project website.  

6.4 Recommendations to Platforms

At the outset of the pandemic, many platforms had insufficient policies to meet the increasing burden of COVID-19 misinformation shared across their channels. The Virality Project acknowledges the improvements made across platforms while offering the following recommendations to mitigate inevitable future misinformation spread:

**Consistently enforce policies against recurrent actors.**

Platform policies evolved during the pandemic. While many platforms have improved transparency around content moderation, there is still inconsistent enforcement of policies, notably in the case of recurring actors. More consistency and transparency is needed around enforcement practices, particularly when prominent or verified accounts are involved. While past policy environments have been slower to enforce policies against prominent accounts, these are the accounts with the greatest potential for impact. If anything, they may merit closer scrutiny.
6.5 Conclusion

The Virality Project focused its work on mis- and disinformation in the COVID-19 vaccine conversation, and established a nonpartisan, multistakeholder model consisting of health sector leaders, federal health agencies, state and local public health officials, social media platforms, and civil society organizations. Based on lessons learned from the Election Integrity Partnership, VP refined its intake and observational processes to create greater efficiency as well as greater capability to assess public narratives across platforms. This allowed for more detailed reporting and enhanced support to stakeholders.

Over the course of its seven months of work, the Virality Project saw narratives that questioned the safety, distribution, and effectiveness of the vaccines. Domestic actors leveraged already existing networks to undermine confidence in the vaccines. These actors were able to leverage bottom-up dynamics of social media to elevate content, often decontextualized or anecdotal, to amplify claims about the alleged danger or lack of effectiveness of the vaccines. While some claims emerged out of genuine concern, uncertainty, or attempts at collaborative sensemaking, others emerged from long-standing anti-vaccine activist communities who have sought to undermine even well-established and demonstrably safe vaccines. Foreign state actors were additionally present in the conversation, though at far lower levels. They pushed narratives both covertly and overtly through state media channels and social media accounts to disparage the safety and efficacy of Western vaccines while bolstering their own vaccines and national interests. Both foreign and domestic threats will continue to be concerns going forward, and we should be prepared for both.
6. Recommendations for Stakeholders

The Virality Project has demonstrated that a focused, well-coordinated collaboration between academia, nonprofits, and the private sector can provide nimble analysis and support to stakeholders who desperately need it. Moving forward, there is a need for a non-governmental independent research entity spearheading VP-style collaboration around emerging mis- and disinformation more generally. The primary focus of this entity should be to institutionalize the cross-platform observation and analysis capability that the Virality Project implemented in response to the ongoing narratives regarding the COVID-19 vaccine. The entity would maintain a permanent research infrastructure, consisting in a set of processes, tools, and roles, that would remain constant across enquiries. The team profile would evolve with the needs of the issue studied, including identifying emerging issues. The entity would package findings in a form most consumable for policymakers, government entities, subject-matter experts, and nonprofit leaders.

The Virality Project was formed out of a conviction that the challenge of high-harm misinformation is not going away and must be met with a collective, resilient, and dynamic response. Given the enormity of the challenge, the Project attempted to meet that call. The collective response must continue in the ongoing fight against misinformation.
Notes

1. PGP, “About Stronger.”


5. A recording of the event can be found here: Confronting Health Misinformation: A Discussion with the Surgeon General, uploaded by Stanford Cyber Policy Center July 26, 2021, https://www.youtube.com/watch?v=GUCdoXapOl0.


8. Masterson et al., “The Case for a Mis- and Disinformation Center of Excellence.”


Appendix A

Research Limitations

This rapid response initiative was not without its limitations. Because of the expansive nature of social media, and the changing complexities of the COVID-19 pandemic, Virality Project partner organizations strove to clearly and narrowly scope the research in such a way as to mitigate the challenge of “content glut,” while not overfiltering, and while working within personnel and data availability constraints. We drew on prior scholarship and our own prior work assessing election mis- and disinformation to ground our categories, actor types, and targeted communities, but there is a degree of subjectivity inherent in determining whether a given piece of content is in or out of scope.

Analysts encountered a myriad of challenges that had the potential to introduce bias into the dataset, such as researchers’ own value judgments or a more general lack of visibility into restricted areas on social media platforms. This report highlights four challenges of interest to social media researchers, particularly those studying mis- and disinformation.

Challenges of Defining Communities

Although there are some quantitative social science methodologies in which communities are identified by relationships or similarities in their networks, or in the content that accounts produce, these methodologies require a degree of data accessibility available on very few major social media platforms; they are most useful on Twitter. On a platform such as Facebook, which has been a core platform for social media misinformation, Virality Project researchers faced a challenge of subjectivity when defining a “community.” The boundaries between certain actor types are porous—for example, to which CrowdTangle list should a prominent influencer who primarily posted wellness content but regularly veered into anti-vaccine pseudoscience, and occasionally parental rights themes, be added? The scoping of
A. Research Limitations

audiences was also a challenge—what Pages, media outlets, or influencers were most relevant to, or widely followed by, Latino or Black communities on Facebook? Analysts sought to make the best possible decisions to ensure that these challenges did not result in simply abandoning classification of a potentially impactful actor, or ignoring a potential targeted community, while recognizing that researcher bias would be inherent in the lists. There were very few methodological discussions of this challenge to inform Virality Project community determination; analysts used social media followings and prevalence of topics covered on the social media posts of an account or page over time to make the best possible assessments. Virality Project outreach coordinators additionally reached out to civil society organizations and community leaders to listen to feedback about our choices and improve our understanding; however, many of these organizations, too, are bandwidth constrained and do not necessarily feel equipped to serve as arbiters defining a community. Over time, Virality Project pod focus shifted toward understanding non-English content, a decision that produced unique insights into those dynamics specifically, though uncertainty remained about VP researchers’ capacity to adequately assess the dynamics of mis- and disinformation in targeted marginalized communities.

Other organizations have faced similar challenges. In December 2021, First Draft Media published a report on COVID-19 vaccine misinformation in Hispanic communities that effectively points out many of the nuances that make this type of research so difficult. As First Draft notes, label terminology can be reductive or inaccurate; it also may not be used by the people it attempts to describe. How to understand, First Draft asks, “the online conversations and flow of information in a nebulous community which largely resists the confines of the umbrella terms under which it has been placed?” This question is salient for other groups who have been lumped together in research that has historically homogenized groups that are in reality quite heterogeneous.¹

It is also challenging to define who content is specifically targeting and how it does so. Because of the lack of publicly available metadata and the nature of closed messaging apps, it can prove challenging to identify who is creating content, who it aims to target, and who and how many people ultimately consume it. Conversations in Hispanic communities, for instance, may be taking place in English, in Spanish, or in Indigenous languages. Shortcuts used by research institutions may result in long-lasting consequences, and, as First Draft comments, the “erasure of many people who are most vulnerable to vaccine misinformation.”²

The importance of understanding the impact of health mis- and disinformation on targeted and marginalized communities, and the challenges inherent in present methodologies, suggest a potential stand-alone research area for future efforts, and a place where interorganizational research communities might form.
Query-related challenges

Keyword-based searches to detect emerging mis- and disinformation about COVID-19 vaccines posed a different sort of challenge. First, they required regular pruning and/or expansion. Many of the broadly relevant vaccine-related keywords in broad early queries returned significant quantities of out-of-scope content (i.e., related to animal vaccinations), so analysts attempted to balance weeding out keywords that would fill their analysis space with irrelevant content, with creating inadvertently restrictive Boolean queries that would filter out relevant content.

Metrics-related challenges

There is no clear definition or engagement threshold that presently defines what constitutes a “viral” or “high-engagement” post. These metrics vary by platform and are influenced by the size of the user base, by community norms, and by platform affordances. Thresholds for what numerical value or rate of change constituted a “viral” or “high-engagement” post varied even within one platform: a community cluster might not have a very high number of followers in the group, but could be highly active in sharing behaviors, or it might have smaller follower counts but several prominent influencers capable of creating virality rapidly across communities. Conspiracy theory-oriented content, which appeared in fewer incidents overall, nonetheless had higher engagement per incident than other content, perhaps because of novelty, and had a propensity to seep into broader anti-vaccine rhetoric over time. Repeatedly, claims that originated from within this relatively small community were picked up by the conservative media ecosystem, which had extremely large follower counts.

Privacy constraints

While subsequent rapid-response partnerships will provide an opportunity to remedy some of these challenges, refine processes, and reduce sampling and other biases, our final and most critical research limitation is likely to persist: lack of visibility into restricted areas on social media platforms. Our work respects privacy policies and terms of service, and we did not gather or engage with data from any private or secret communities on any social platform. Our research and analysis drew exclusively from public social media content (posts on public Facebook Pages and Groups, tweets, Reddit posts, public Instagram accounts, and public Telegram channels, etc.). However, given the well-documented behavior of activist communities using private and secret groups for organizing purposes, this commitment and structure means that our visibility into the overall narrative space is limited.
Definitions

Anti-vaccine, anti-vaccine activist, and anti-vaccine influencer: In discussing the narratives shaping the public conversation about COVID-19, it is important to acknowledge the spectrum of perspectives about vaccines, particularly the distinction between “vaccine-hesitant” and “anti-vaccine” attitudes.³ In this report, we use the term “anti-vaccine” in very specific cases. Drawing on the literature to define the phenomenon, we considered a range of definitions while establishing ours: Someone who is anti-vaccine believes that vaccines do not work or are not safe⁴ and disagrees with their use in the general population. People who are anti-vaccine typically refuse vaccines for themselves and their children (if applicable).⁵ They may deny the existence or challenge the validity of the science supporting vaccination.⁶ They may also deny the existence or severity of the illness, which can play into narratives that the vaccine may be distributed by entities with ulterior motives. On the other hand, vaccine-hesitant people may simply be uncertain or undecided about vaccine use.

While a significant number of people have anti-vaccine beliefs, anti-vaccine activists are further distinguished by their strategic attempts to mobilize public sentiment and collective action in the interest of social or political change. As with all activism, such mobilization is usually seen in the context of the creation, growth, and maintenance of a movement. In this report, the term activist refers to those who actively involve themselves in anti-vaccine legislative and organizing efforts. These were long-term movement builders who organized events, lobbied legislatures, filed FOIAs, and engaged in legislation debates and efforts around school requirements for vaccination. Anti-vaccine influencer, however, refers simply to actors popular on social media who use their platforms to spread false and misleading claims about vaccines. Activists often overlapped with the long-standing anti-vaccine influencers discussed in Chapter 4: Actors & Tactics.

Disinformation is false or misleading information purposefully produced and/or
spread for an objective, such as a political or financial goal.\textsuperscript{7} Disinformation may mislead through its content, or may work by deceiving its audiences about its origins, purpose, or producers’ identity. It differs from other forms of misinformation in that it is \textit{intentionally} produced and/or spread. It may layer falsehoods or exaggerations around a plausible or true core,\textsuperscript{8} and typically functions as a campaign rather than a single piece of content. As a disinformation campaign progresses, unwitting participants may take part in its production and spread; therefore, not every entity that spreads disinformation does so with intent to deceive or knowledge that they are spreading false or misleading content.\textsuperscript{9}

\textbf{Incidents} are specific events or stories that support a narrative’s argument. In this report, they refer to a ticket associated with a particular actor.

\textbf{Misinformation} is information that is false, though not necessarily by intent.\textsuperscript{10} Misinformation is at times used as an umbrella category for disinformation, false rumors, and other types of false and misleading information.\textsuperscript{11}

\textbf{Narratives} are stories that connect a series of related events or experiences, providing compelling interpretations that can help people make sense of these events and experiences. In the context of this report, they are central arguments that further vaccine hesitancy or rejection.

\textbf{Pseudomedicine} is a complex term derived from conceptions and descriptions of \textbf{pseudoscience} (first described by Karl Popper).\textsuperscript{12} Here, we use a modern view that posits that science is distinguished by a set of social norms around its production, evaluation, and discussion. Naomi Oreskes argues that what makes science science is the critical scrutiny of claims, which ensures faulty claims are rejected. Michael Strevens argues that science is unique in that by proving one’s own argument, one gives the reader the tools (through data and methods) to prove the scientist wrong. Pseudoscience, on the other hand, adopts the language of science while violating those norms that make science trustworthy, including that critical scrutiny and engagement with the scientific community.\textsuperscript{13}

In this report, we acknowledge that the term \textbf{pseudomedical} can be problematic, but we use the above definitions to help characterize various actors as pseudomedical—those who speak with the language and authority of medical science while violating scientific norms, and who appeal to that alleged authority to amplify anti-vaccine narratives. Some may be or may once have been credentialed practitioners (e.g., doctors of medicine or osteopathy) or recognized as experts in the scientific or medical communities. Still others are so-called “epistemic trespassers”: experts who pass judgment on questions from fields in which they lack expertise. In all cases, however, pseudomedical influencers are those that capitalize on their actual or perceived status as experts, “doctors,” or alleged authoritative figures in science, public health, or medicine.

\textbf{Recurring actors} are those actors whose social media content regularly included
B. Definitions

vaccine mis- or disinformation, who had large social media followings, and who demonstrated recurring success making content go viral.
Appendix C

Policy Timeline (Sources)

These policies exist in the platforms’ community standards or in blog posts. Policies included in the timeline were created to specifically address vaccine-related content, though an important caveat is that some platforms have stated that preexisting policies unrelated to vaccines may still address vaccine-related content. For example, in September 2021, Reddit said vaccine misinformation is already covered by a “rule against posting content that ‘encourages’ physical harm.”  

Pre-COVID-19

- Pinterest

- Facebook/Instagram
C. Policy Timeline (Sources)

• YouTube

• Twitter

Early 2020

• TikTok

• Pinterest

• Nextdoor

• YouTube

• Google
December 2020

• Facebook/Instagram

• Google

• TikTok

• Twitter
  ► Twitter Safety, “COVID-19: Our approach to misleading vaccine information.”

2021

• Nextdoor

• Facebook/Instagram
C. Policy Timeline (Sources)


- Twitter


- YouTube


- TikTok


Platform Policy Tables (Sources)

“Remove” Policies for Vaccine Misinformation

The Virality Project’s approach to evaluating platform policies stems from previous work it published in February 2021. The work is focused only on publicly communicated policies. While platforms have internal policies, the Project considered only those that were published either to their community standards or in a blog post and that publicly state they are applied to vaccines.

Table Key:

**Yes**: The platform explicitly states the policy publicly in either its community standards or its blog posts.

**Partial/Not Explicit**: The platform partially fulfills the policy, but it does not apply in all cases, or the policy may apply, but it is not explicitly stated.

**No**: The platform does not have the policy stated publicly in either its community standards or its blog posts.

*Text in italics is additional commentary on the policy or justification for our policy assessment. The exact policy language is in quotation marks.*
### Table D.1

<table>
<thead>
<tr>
<th>“Remove” Policies</th>
<th>Removes content that could lead to physical harm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook / Instagram</td>
<td>Yes. “Under our Community Standards, we remove misinformation when public health authorities conclude that the information is false and likely to contribute to imminent violence or physical harm.”</td>
</tr>
<tr>
<td>Twitter</td>
<td>Yes. “We are focused on mitigating misleading information that presents the biggest potential harm to people’s health and wellbeing.”</td>
</tr>
<tr>
<td>YouTube</td>
<td>Yes. “YouTube doesn’t allow content about COVID-19 that poses a serious risk of egregious harm.”</td>
</tr>
<tr>
<td>TikTok</td>
<td>Yes. “Medical misinformation that can cause harm to an individual’s physical health.”</td>
</tr>
<tr>
<td>Pinterest</td>
<td>Yes. “Medically unsupported health claims that risk public health and safety, including the promotion of false cures, anti-vaccination advice, or misinformation about public health or safety emergencies.”</td>
</tr>
<tr>
<td>Nextdoor</td>
<td>Yes. “As a commitment to neighbor safety, we prohibit misinformation about vaccines that has the potential to cause harm or threatens public safety.”</td>
</tr>
<tr>
<td>Reddit</td>
<td>Yes. “We have long interpreted our rule against posting content that ‘encourages’ physical harm, in this help center article, as covering health misinformation, meaning falsifiable health information that encourages or poses a significant risk of physical harm to the reader.”</td>
</tr>
</tbody>
</table>
Sources

• Facebook/ Instagram

• Twitter
  ▶ Twitter Safety, “COVID-19: Our approach to misleading vaccine information.”

• YouTube

• TikTok

• Pinterest

• Nextdoor

• Reddit
### Table D.2

<table>
<thead>
<tr>
<th>&quot;Remove&quot; Policies</th>
<th>Removes content that runs counter to public health expert consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Facebook / Instagram</strong></td>
<td><strong>Yes.</strong> “Under our Community Standards, we remove misinformation when public health authorities conclude that the information is false and likely to contribute to imminent violence or physical harm.”</td>
</tr>
</tbody>
</table>
| **Twitter** | **Not Explicit.** “We are expanding the policy and may require people to remove Tweets which advance harmful false or misleading narratives about COVID-19 vaccinations, including: False claims which have been widely debunked about the adverse impacts or effects of receiving vaccinations.”<sup>a</sup>  
<sup>a</sup>We are interpreting “widely debunked” to mean a process that includes consultation of experts. However, the policy does not explicitly state it will take down content that runs counter to public health expert consensus. |
| **YouTube** | **Yes.** “YouTube doesn’t allow content that spreads medical misinformation that contradicts local health authorities’ or the World Health Organization’s (WHO) medical information about COVID-19.” |
| **TikTok** | **No.** While TikTok consults with experts and fact-checkers to develop its policies and assess the accuracy of the information, these policies do not explicitly state that TikTok will remove content that goes against public health expert consensus. |
## Table D.3

<table>
<thead>
<tr>
<th>&quot;Remove&quot; Policies</th>
<th>Removes content that runs counter to public health expert consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pinterest</strong></td>
<td><strong>Yes.</strong> &quot;Since 2017 it’s been against our community guidelines to promote anti-vaccination advice and other health misinformation, and we’ve never allowed it to be advertised. We’ve long used guidance from these health institutions to inform our policy and how we enforce it.”&lt;br&gt;&quot;We remove or limit distribution of false or misleading content that may harm Pinners’ or the public’s well-being, safety or trust, including: Medically unsupported health claims that risk public health and safety, including the promotion of false cures, anti-vaccination advice, or misinformation about public health or safety emergencies.”*&lt;br&gt;*We are interpreting &quot;medically unsupported&quot; to mean going against health expert consensus.</td>
</tr>
<tr>
<td><strong>Nextdoor</strong></td>
<td><strong>Yes.</strong> &quot;As a commitment to neighbor safety, we prohibit misinformation about vaccines that has the potential to cause harm or threatens public safety. We rely on trusted public health resources to evaluate content that may be misinformation, including the World Health Organization (WHO) and the Centers for Disease Control and Prevention (CDC).”&lt;br&gt;<strong>Partial.</strong> “Health Disinformation. Our rule against impersonation, as described in this help center article, extends to ‘manipulated content presented to mislead.’ We have interpreted this rule as covering health disinformation, meaning falsifiable health information that has been manipulated and presented to mislead. This includes falsified medical data and faked WHO/CDC advice.”*&lt;br&gt;*While Reddit’s policy includes removing Health Disinformation related to the CDC or WHO, the platform is not explicit on how it will handle other authentic attempts at spreading content that runs counter to the CDC or WHO.</td>
</tr>
<tr>
<td><strong>Reddit</strong></td>
<td><strong>Partial.</strong> “Health Disinformation. Our rule against impersonation, as described in this help center article, extends to ‘manipulated content presented to mislead.’ We have interpreted this rule as covering health disinformation, meaning falsifiable health information that has been manipulated and presented to mislead. This includes falsified medical data and faked WHO/CDC advice.”*&lt;br&gt;*While Reddit’s policy includes removing Health Disinformation related to the CDC or WHO, the platform is not explicit on how it will handle other authentic attempts at spreading content that runs counter to the CDC or WHO.</td>
</tr>
</tbody>
</table>
D. Platform Policy Tables (Sources)

**Sources**

- **Facebook/ Instagram**
  - “COVID-19 and Vaccine Policy Updates & Protections,” Facebook Help Center.

- **Twitter**
  - Twitter Safety, “COVID-19: Our approach to misleading vaccine information.”

- **YouTube**

- **TikTok**

- **Pinterest**
  - Ozoma, “Bringing authoritative vaccine results to Pinterest search.”

- **Nextdoor**
<table>
<thead>
<tr>
<th>“Remove” Policies</th>
<th>Clear strike policy for removing recurring actors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook / Instagram</td>
<td><strong>Not Explicit.</strong> While Facebook has a policy on “counting strikes” against content that violates Facebook or Instagram Community Standards, the strikes are not clearly communicated.</td>
</tr>
<tr>
<td>Twitter</td>
<td><strong>Yes.</strong> Five-strike system for removing accounts under COVID-19 misleading Information Policy. If your content violates this policy, we'll remove the content and send you an email to let you know. If this is your first time violating our Community Guidelines, you'll likely get a warning with no penalty to your channel. If it’s not, we may issue a strike against your channel. If you get 3 strikes within 90 days, your channel will be terminated.”</td>
</tr>
<tr>
<td>YouTube</td>
<td><strong>Yes.</strong></td>
</tr>
<tr>
<td>TikTok</td>
<td><strong>No.</strong> While TikTok’s community standards state it will “suspend or ban accounts and/or devices that are involved in severe or repeated violations,” it does not state the threshold of violations for which it will then take action.</td>
</tr>
<tr>
<td>Pinterest</td>
<td><strong>No.</strong> While Pinterest’s policies state that “accounts may be suspended due to single or repeat violations of our Community Guidelines,” it does not state the threshold of violations for which it will then take action.</td>
</tr>
<tr>
<td>Nextdoor</td>
<td><strong>No.</strong> While Nextdoor has a policy on account suspension, it is not clear how many violations an account must commit before it is removed.</td>
</tr>
<tr>
<td>Reddit</td>
<td><strong>No.</strong> While Reddit has a list of enforcement approaches that range from “Asking you nicely to knock it off” to “Banning of Reddit communities,” its policy does not explicitly state what the threshold of violations is before suspending accounts or communities.</td>
</tr>
</tbody>
</table>
D. Platform Policy Tables (Sources)

**Sources**

- **Facebook / Instagram**

- **Twitter**

- **TikTok**
  - “Community Guidelines,” TikTok.

- **YouTube**

- **Pinterest**

- **Nextdoor**

- **Reddit**
"Reduce" Policies for Vaccine Misinformation

Table D.5

<table>
<thead>
<tr>
<th>&quot;Reduce&quot; Policies</th>
<th>Search results for &quot;vaccine&quot;-related terms point to authoritative sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook / Instagram</td>
<td>Yes. “We will reduce the ranking of groups and Pages that spread misinformation about vaccinations in News Feed and Search. These groups and Pages will not be included in recommendations or in predictions when you type into Search.”</td>
</tr>
<tr>
<td>Twitter</td>
<td>Yes. “This new tool means that when someone searches for certain keywords associated with vaccines, a prompt will direct them to a credible Government public health resource that provides critical information about immunization and vaccination.”</td>
</tr>
<tr>
<td>YouTube</td>
<td>Yes. “When people now search for news or information, they get results optimized for quality, not for how sensational the content might be.” “We're updating our COVID-19 information panels to include links to COVID-19 vaccine info. The updated panels may* show in search results and on watch pages related to COVID-19 or COVID-19 vaccine info.” *While the policy says panels “may” show up, we interpret this policy as sufficient for providing authoritative sources for user’s search.</td>
</tr>
<tr>
<td>TikTok</td>
<td>Yes. “We are updating our information hub in-app so that when people search for vaccine information in-app, they will be directed to trusted information about the vaccine from respected experts.”</td>
</tr>
<tr>
<td>Pinterest</td>
<td>Yes. “We’re introducing a new experience for vaccine-related searches on Pinterest. Now, when you search for ‘measles,’ ‘vaccine safety’ and other related health terms, you can explore reliable results about immunizations from leading public health organizations.”</td>
</tr>
<tr>
<td>Nextdoor</td>
<td>Yes. “When you search for coronavirus (or other related terms) on Nextdoor, you’ll receive a pop-up announcement that directs you to information from public health officials.”</td>
</tr>
<tr>
<td>Reddit</td>
<td>Partial. While Reddit has a policy of adding “banners directing users to the CDC and r/Coronavirus” on its home page and in search, this banner seems to surface only when you search “COVID-19,” not just “vaccines.” For example, searching for “COVID-19 vaccine” will produce the information banner linking to the CDC’s website. But a search for “vaccines” will not produce this banner.</td>
</tr>
</tbody>
</table>
D. Platform Policy Tables (Sources)

Sources

• Facebook/Instagram
  ▶ “Combatting Vaccine Misinformation,” Facebook Newsroom.

• Twitter

• YouTube
  ▶ Neal Mohan, “Perspective: Tackling Misinformation on YouTube.”

• TikTok
  ▶ Kevin Morgan, “Taking action against COVID-19 vaccine misinformation.”

• Pinterest
  ▶ Ozoma, “Bringing authoritative vaccine results to Pinterest search.”
  ▶ “Introducing the Today tab for daily inspiration,” Pinterest Newsroom.

• Nextdoor
  ▶ Team Nextdoor, “Nextdoor’s COVID-19 resources & updates.”

• Reddit
Table D.6

<table>
<thead>
<tr>
<th>“Reduce” Policies</th>
<th>Restricts certain forms of vaccine ads</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook/ Instagram</td>
<td>Yes. “Today, we’re launching a new global policy that prohibits ads discouraging people from getting vaccinated. We don’t want these ads on our platform.”</td>
</tr>
<tr>
<td>Twitter</td>
<td>Yes. “We already ensure that advertising content does not contain misleading claims about the cure, treatment, diagnosis or prevention of certain diseases and conditions, including vaccines.”</td>
</tr>
<tr>
<td>YouTube</td>
<td>Yes. “Below are some examples of COVID-19-related content that don’t meet our policies...Medical Misinformation: Content that misinforms users about health matters related to COVID-19.”</td>
</tr>
<tr>
<td>TikTok</td>
<td>Yes. “We also do not allow paid advertising that advocates against vaccinations, though PSAs or calls to action related to COVID-19 vaccines are accepted on a case-by-case basis if they’re in the interest of public health and safety.”</td>
</tr>
<tr>
<td>Pinterest</td>
<td>Yes. “Prohibiting ads that claim to offer cures or treatments or that are looking to exploit the crisis.”</td>
</tr>
<tr>
<td>Nextdoor</td>
<td>Yes. “We prohibit...anti-vaccination (or vaccine hesitancy) messaging.”</td>
</tr>
<tr>
<td>Reddit</td>
<td>Not Explicit. While Reddit has a policy that prohibits “Deceptive, Untrue, or Misleading Advertising,” it is unclear how this policy covers vaccine-related content.</td>
</tr>
</tbody>
</table>
Sources

• Facebook/Instagram

• Twitter

• YouTube

• TikTok
  ► “COVID-19,” TikTok.

• Pinterest
  ► “Introducing the Today tab for daily inspiration,” Pinterest Newsroom.

• Nextdoor

• Reddit
Table D.7

<table>
<thead>
<tr>
<th>“Reduce” Policies</th>
<th>Reduce groups, pages, and content that promote vaccine misinformation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook/Instagram</td>
<td><strong>Yes.</strong> “We will reduce the ranking of groups and Pages that spread misinformation about vaccinations in News Feed and Search. These groups and Pages will not be included in recommendations or in predictions when you type into Search.”</td>
</tr>
<tr>
<td>Twitter</td>
<td><strong>Yes.</strong> If a tweet is labeled under the COVID-19 misleading information policy, Twitter will &quot;Reduce the visibility of the Tweet on Twitter and/or prevent it from being recommended.&quot;</td>
</tr>
<tr>
<td>YouTube</td>
<td><strong>Not Explicit.</strong> <em>YouTube has a policy for reducing the spread of &quot;borderline content,&quot; which it defines as &quot;Content that comes close to—but doesn't quite cross the line of—violating our Community Guidelines.&quot; However, it is unclear how this is applied to vaccine misinformation specifically.</em></td>
</tr>
<tr>
<td>TikTok</td>
<td><strong>Yes.</strong> “We use a mix of technology and human moderation to enforce these policies, including by removing content, banning accounts, and making it more difficult to find harmful content, like misinformation and conspiracy theories, in recommendations or search.”</td>
</tr>
<tr>
<td>Pinterest</td>
<td><strong>No.</strong> <em>Pinterest does not have a policy approach for misleading vaccine-related content other than removing that content.</em></td>
</tr>
<tr>
<td>Nextdoor</td>
<td><strong>No.</strong> <em>Nextdoor does not have a policy for reducing content through its algorithm.</em></td>
</tr>
<tr>
<td>Reddit</td>
<td><strong>Yes.</strong> For Reddit communities that have been “Quarantined,” these communities will “display a warning that requires users to explicitly opt-in to viewing the content. They generate no revenue, do not appear in non-subscription-based feeds (e.g. Popular), and are not included in search or recommendations.”</td>
</tr>
</tbody>
</table>
D. Platform Policy Tables (Sources)

Sources

- Facebook/Instagram
  - “Combatting Vaccine Misinformation,” Facebook Newsroom.

- Twitter
  - “COVID-19 misleading information policy,” Twitter Help Center.

- TikTok

- Reddit
“Inform” Policies for Vaccine Misinformation

Table D.8

<table>
<thead>
<tr>
<th>“Inform” Policies</th>
<th>Provides COVID-19 information hub on platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook/Instagram</td>
<td>Yes. “COVID-19 Information Center.”</td>
</tr>
<tr>
<td>Twitter</td>
<td>Yes. “We have also ensured the Events feature contains credible information about Covid-19.”</td>
</tr>
<tr>
<td>YouTube</td>
<td>No. <em>YouTube's platform does not provide a page for YouTube-curated sources on COVID-19 or COVID-19 vaccine information.</em></td>
</tr>
<tr>
<td>TikTok</td>
<td>Yes. “In February, we continued our work by partnering with trusted sources to introduce an information hub in-app to provide the TikTok community with access to accurate information.”</td>
</tr>
<tr>
<td>Pinterest</td>
<td>Yes. “Stay safe. Stay inspired” Information Board.</td>
</tr>
<tr>
<td>Nextdoor</td>
<td>Yes. “We have aggregated posts from public health officials into ’carousels’ at the top of the feed so that the information from these trusted sources is easily discoverable. Here, we’ll include posts that contain accurate, real-time information from public agencies, including cities, counties, and local and state health departments.”</td>
</tr>
<tr>
<td>Reddit</td>
<td>Yes. “A dedicated AMA series connecting users with authoritative experts on coronavirus and vaccines.”</td>
</tr>
</tbody>
</table>

Sources

- Facebook/Instagram

- Twitter
  - Twitter Inc., “Coronavirus: Staying Safe and Informed on Twitter.”

- TikTok
  - Kevin Morgan, “Taking action against COVID-19 vaccine misinformation.”

- Pinterest

- Nextdoor
  - Team Nextdoor, “Nextdoor’s COVID-19 resources & updates.”
<table>
<thead>
<tr>
<th>Platform</th>
<th>“Inform” Policies</th>
<th>Applies banners and labels on misleading COVID-19-related content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook / Instagram</td>
<td>Yes. “We’re continuing to expand our efforts to address COVID-19 vaccine misinformation by adding labels to Facebook and Instagram posts that discuss the vaccines. These labels contain credible information about the safety of COVID-19 vaccines from the World Health Organization.”</td>
<td></td>
</tr>
<tr>
<td>Twitter</td>
<td>Yes. “Starting in early 2021, we may label or place a warning on Tweets that advance unsubstantiated rumors, disputed claims, as well as incomplete or out-of-context information about vaccines. Tweets that are labeled under this expanded guidance may link to authoritative public health information or the Twitter Rules to provide people with additional context and authoritative information about COVID-19.”</td>
<td></td>
</tr>
<tr>
<td>YouTube</td>
<td>Partial. “When you search or watch videos related to topics prone to misinformation, such as the moon landing, you may see an information panel at the top of your search results or under a video you’re watching.” “As a continuation of our efforts to combat COVID-19 related misinformation, we’re updating our COVID-19 information panels to include links to COVID-19 vaccine info. The updated panels may show in search results and on watch pages related to COVID-19 or COVID-19 vaccine info.”</td>
<td></td>
</tr>
<tr>
<td>TikTok</td>
<td>No. While TikTok’s policy states it will apply COVID-19 information labels to any content that is about COVID-19, it does not state that it will label misinformation content differently or explicitly.</td>
<td></td>
</tr>
<tr>
<td>Pinterest</td>
<td>No. Pinterest does not have a policy approach for misleading vaccine-related content other than removing that content.</td>
<td></td>
</tr>
<tr>
<td>Nextdoor</td>
<td>Partial. While Nextdoor has a policy that applies labels to “certain posts related to COVID-19 that direct to the CDC for trusted information,” it is unclear if this policy is applied to all COVID-19 information or posts that share COVID-19 misinformation.</td>
<td></td>
</tr>
<tr>
<td>Reddit</td>
<td>Partial. While Reddit’s Quarantined Subreddits have a warning label that “requires users to explicitly opt-in to viewing the content,” Reddit does not apply labels to post-level content, where vaccine misinformation can spread.</td>
<td></td>
</tr>
</tbody>
</table>
Sources

- Facebook / Instagram

- Twitter
  - Twitter Safety, “COVID-19: Our approach to misleading vaccine information.”

- YouTube

- NextDoor
  - Team Nextdoor, “Nextdoor’s COVID-19 resources & updates.”

- Reddit
  - “Quarantined Subreddits,” Reddit.
### Table D.10

<table>
<thead>
<tr>
<th>“Inform” Policies</th>
<th>Universal label policy for all content mentioning COVID-19</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook/Instagram</td>
<td>Yes. “We’re rolling out labels on all posts generally about COVID-19 vaccines that point people to the COVID-19 Information Center globally, and plan to add additional targeted labels about COVID-19 vaccine subtopics.”</td>
</tr>
<tr>
<td>Twitter</td>
<td>No. Twitter does not apply a universal label to posts that mention COVID-19 vaccines.</td>
</tr>
<tr>
<td>YouTube</td>
<td>Yes. “As a continuation of our efforts to combat COVID-19 related misinformation, we’re updating our COVID-19 information panels to include links to COVID-19 vaccine info. The updated panels may show in search results and on watch pages related to COVID-19 or COVID-19 vaccine info.”</td>
</tr>
<tr>
<td>TikTok</td>
<td>Yes. “[a] new vaccine tag to detect and tag all videos with words and hashtags related to the COVID-19 vaccine. We will attach a banner to these videos with the message ‘Learn more about COVID-19 vaccines.’”</td>
</tr>
<tr>
<td>Pinterest</td>
<td>No. Pinterest does not use labels on post-level content.</td>
</tr>
<tr>
<td>Nextdoor</td>
<td>No. Nextdoor does not apply a universal label to posts that mention COVID-19 vaccines.</td>
</tr>
<tr>
<td>Reddit</td>
<td>No. Reddit does not apply a universal label to posts that mention COVID-19 vaccines.</td>
</tr>
</tbody>
</table>
Sources

- Facebook/Instagram

- YouTube

- TikTok
  - Kevin Morgan, “Taking action against COVID-19 vaccine misinformation.”

- Pinterest
  - “Introducing the Today tab for daily inspiration,” Pinterest Newsroom.

Notes


2. Longoria et al.


5. Benoit and Mauldin.


D. Platform Policy Tables (Sources)


“Watch Alex Jones Predict Door to Door Vaccination Squads.” Alex Jones Predictions, 2021. https://freeworldnews.tv/watch?id=60e5caa839348447c5729075.


Annenberg Public Policy Center of the University of Pennsylvania. “Vaccine misinformation and social media: People exposed to vaccine content on social media more likely misinformed than those exposed to it on traditional media.” ScienceDaily, February 17, 2020. https://www.sciencedaily.com/releases/2020/02/200217163004.htm.


Berenson, Alex (@AlexBerenson). “And there it is: although I must warn I don’t have firsthand knowledge of this interview - can someone provide a direct link if possible? https://t.co/XaMl6szSyV.” Twitter, June 12, 2021. https://web.archive.org/web/20210619061316/https://twitter.com/alexberenson/status/140380772345745411.


Bibliography


Burages, Tami (@tburages). “A week ago today my brother’s 13-year-old son had his 2nd covid shot. Less than 3 days later he died. The initial autopsy results (done Friday) were that his heart was enlarged and there was some fluid surrounding it. He had no known health problems. Was on no medications.” Twitter, June 20, 2021. https://web.archive.org/web/20210710054935/https://twitter.com/tburages/status/1406796255573991424.
——— (@tburages). “I have deleted a tweet thread about the sudden death of my nephew following his 2nd covid shot. Fox News and other far-right disinformation networks are politicizing the death and using my tweet to do it. They are causing my family more pain. If you have any questions, DM me.” Twitter, June 22, 2021. https://web.archive.org/web/20210712233418/https://twitter.com/tburages/status/1407343080026611713.


Bibliography


Center for Countering Digital Hate [@CCDHate]. “Del Bigtree is still broadcasting dangerous misinformation about vaccines and Covid to half a million followers on @Facebook, @Instagram and @Twitter. It’s time they followed @YouTube and removed his accounts,” July 30, 2020. https://twitter.com/CCDHate/status/1288769298304061440/photo/2.


Bibliography


Chad, The (@theChadFactor). “Daily reminder I don’t care if half the world is vaccinated I’m not taking the vaccine. I don’t even take flu shots. A lot of you people need to research Tuskegee seriously.” Twitter, July 27, 2021. https://web.archive.org/web/20210727211429/https://twitter.com/TheChadFactor/status/1420130427709366288.


——— (@earthley). “This...is heartbreaking. This is ‘safe?’ Then what is DANGEROUS??” Facebook, July 11, 2021. https://www.facebook.com/749122528556667/posts/2339372269531677.


Freedom From Mental Slavery. “Tuskegee Study: Black people thought they were receiving free healthcare but instead were injected with Syphilis. This was conducted by the U.S. Public Health Service and happened over a period of 40 years between 1932 and 1972. Let us never forget.” Facebook, July 25, 2021. https://www.facebook.com/161399403976816/posts/4137971696319547.


Bibliography


——— (@drsimonegold). “The founder of Shake Shack says his company will require proof of vaccination for both employees AND customers. I encourage everyone to boycott this business. No corporation or government has a right to demand to see your private health documents.” Twitter, August 1, 2021. https://twitter.com/drsimonegold/status/1421750765866029058.

——— (@drsimonegold). “WOW: @RoyalCaribbean cruises is now requiring you to get an experimental vaccine to book a cruise with them. I will be boycotting the company instead. Who else is with me? #BoycottRC.” Twitter, March 20, 2021. https://twitter.com/drsimonegold/status/1373361332062412810.


Jr, Robert F. Kennedy (@RobertKennedyJr). “#Norway upped the number of deaths under investigation, from 23 last week to 33, while in #Germany, health officials said they are investigating 10 deaths that occurred among elderly patients who received the #COVID #vaccine. #TheDefender https://t.co/Y5FxxyIJcK.” Twitter, January 18, 2021. https://web.archive.org/web/20210826175814/https://twitter.com/RobertKennedyJr/status/1351277936607948800.


Bibliography


KY, KING (@DaddyKy_). “Yeaaa it’s time to get @POTUS on up out of the house.. You can’t force ppl to get a vaccine. We remember the Tuskegee Syphilis Study sir.. Why they hell should we trust what the government say when all y’all ever done was let us down?? Now you want us to have faith in y’all? 😂.” Twitter, July 27, 2021. https://web.archive.org/web/20210831050326/https://twitter.com/ThatGuyOGPrime/status/1420042590821527552.


Mayhew, Ishmael (@ishcontent). “What part did we misunderstand about it? The Tuskegee Experiment went on for 40 years. Define not that long ago? It sounds like gaslighting to me. This was REAL and it impacted generations. We not about to skim over that. You shouldnt need to do that to defend the vaccine.” Twitter, July 27, 2021. https://web.archive.org/web/20210727195548/https://twitter.com/ishcontent/status/1420100592131887115.


NFSC. 喜台農.III (@NfscSuperpower). “有515100人因注射中共病毒疫苗而死亡，但為何報導只會告訴你接種疫苗實際死亡率僅1%或更少？因為虛假媒体持續打壓『疫苗真相』傳播!” [There are 51,5100 people who died due to the Chinese Communist virus vaccine, but why the report only tells you that the actual mortality rate is only 1% or less? Because false media continue to suppress the spread of the “vaccine truth!”] Twitter, June 16, 2021, 4:01 p.m. https://web.archive.org/web/20210617031119/https://twitter.com/NfscSuperpower/status/1405193921831964679.


Owens, Candace (@RealCandaceO). “Nuremberg codes stipulate that consent must be voluntary. Telling a child they cannot go to school, firing people from work, or preventing their travel is NOT voluntary consent. Threatening someone's livelihood unless they take an experimental vaccine is NOT voluntary consent.” Twitter, July 7, 2021. http://web.archive.org/web/20210710105514/https://twitter.com/RealCandaceO/status/1412880295368478723.

———. “This is shocking! 60% of people being admitted to the hospital with #COVID19 in England have had two doses of a coronavirus vaccine, according to the government’s chief scientific adviser.” Facebook, July 19, 2021. https://www.facebook.com/1593518174052711/posts/5768837649854055/.


Real Immunity (@realimmunitymovie). “UPDATE- For anyone who has tried to navigate the government VAERS (Vaccine Adverse Event Reporting System) site and has thrown up their hands in frustration, here is an easy way to access an updated summary.” Facebook, July 9, 2021. https://www.facebook.com/695769770631109/posts/1679826732225403.


Revolution, The Next (@NestRevFNC). “@naomirwolf sounds the alarm on the push for ‘vaccine passports’: ‘I am not overstating this, I can’t say it forcefully enough, this is literally the end of human liberty in the West if this plan unfolds…’” Twitter, March 28, 2021. https://twitter.com/NextRevFNC/status/1376370803059658754.


Rise Up Canada (@riseup_canada). “👉🏻👉🏻👉🏻 VAERS captures ONLY AN ESTIMATED 1% of Vaccine injury, because most injuries are not reported. VAERS CV Reports Through April 30…” Instagram, May 7, 2021. https://www.instagram.com/p/COmTtFlhqqq/.


Soular, Ash (@ashsoularmusic). “People have always speculated how the mark of the beast would be so widely accepted and implemented. How so many would be deceived into taking it instead of rejecting and fighting against it. I never understood a scenario that made sense to me until I saw how this mask mandate has been implemented…” *Facebook*, July 18, 2020. [https://www.facebook.com/ashsoularmusic/posts/10157426038694212](https://www.facebook.com/ashsoularmusic/posts/10157426038694212).


ThisIsOurShot (@ThisIsOurShot). “There are a few things that have been shown to affect male fertility, but the COVID-19 vaccine is not one of them. Trust science: Wear boxers and get vaccinated. #ThisIsOurShot #VacunateYa #WeCanDoThis.” Twitter, July 1, 2021. https://twitter.com/ThisIsOurShot/status/1410584124314861568.


Twitter Comms [@TwitterComms]. “We shouldn’t have to say this, but you should read an article before you Tweet it … So, we’ve been prompting some people to do exactly that. Here’s what we’ve learned so far,” September 24, 2020. https://twitter.com/TwitterComms/status/1309178715717369856.


UCLA Health (@UCLAHealth). “These comments do not represent the views of UCLA Health. Unvaccinated people are more likely to contract COVID-19 & we encourage employees to be vaccinated. We adhere to the state public health order requiring health care workers to be vaccinated or undergo regular testing.” Twitter, August 21, 2021. https://twitter.com/UCLAHealth/status/1432754622695817218.


Vallance, Sir Patrick (@uksciencechief). “Correcting a statistic I gave at the press conference today, 19 July. About 60% of hospitalisations from covid are not from double vaccinated people, rather 60% of hospitalisations from covid are currently from unvaccinated people.” Twitter, July 19, 2021. https://twitter.com/uksciencechief/status/1417204235356213252.


Yi, Joseph (yojimd), ed. “Relationships can be #HAM these days.” Instagram, September 2021. https://www.instagram.com/p/CTH9qICg2IV/.


