

Summary Project Report: Increasing COVID-19 Vaccine Confidence

2024

Project Period

January 15, 2021 - June 30, 2024

Project Team

Rishi Goyal, Associate Professor, Columbia University Department of Emergency Medicine and the Institute of Comparative Literature and Society

Dennis Yi Tenen, Associate Professor, Columbia University Department of English and Comparative Literature, Data Science Institute

Arden Hegele, Lecturer in the Discipline of English and Comparative Literature

Emily Madison, Lecturer in the Discipline of English and Comparative Literature

Melissa Stockwell, Chief of the Division of Child and Adolescent Health and an Associate Professor of Pediatrics, Vagelos College of Physicians and Surgeons, and Population and Family Health, Mailman School of Public Health

Jeremiah Trinidad-Christensen, Columbia University Libraries

Moacir P. de Sa Pereira, Columbia University Libraries

Partners

Vaccine Confidence Project, London School of Hygiene & Tropical Medicine

Maine Center for Disease Control & Prevention (Maine CDC)

Department of Health & Mental Health of Ulster County, New York

Acknowledgments

The **Increasing COVID-19 Vaccine Confidence** project was supported and funded by Columbia World Projects, an initiative under Columbia Global, which brings together major global initiatives from across the university to advance knowledge and foster global engagement.

Columbia World Projects (CWP) is a university-wide initiative established in 2017 to bridge scholarly knowledge and real-world action. With the goal of achieving the greatest possible impact on pressing challenges of our time, CWP mobilizes Columbia University's scholars, researchers, practitioners, and students to identify and implement interdisciplinary solutions to complex societal challenges in partnership with targeted change agents, such as policymakers, government agencies, non-governmental organizations, and corporations.

For more information, contact:

Dennis Tenen (dt2406@columbia.edu) and **Rishi Goyal** (rkg6@columbia.edu)

Increasing COVID-19 Vaccine Confidence

Introduction

Vaccine hesitancy has risen sharply over the last two decades and is a major public health concern worldwide. In 2019, the World Health Organization (WHO) cited vaccine hesitancy [1] as one of the top 10 threats to global health. The ***Increasing COVID-19 Vaccine Confidence*** project was born out of the belief that to effectively combat vaccine hesitancy, it is critical to understand its emotional, ideological, and rhetorical base. This CWP project spanned 3.5 years (January 2021-June 2024) and harnessed the power of data science and artificial intelligence (AI), in combination with tools of literary scholarship, to better understand the complexities of vaccine hesitancy and to inform more effective messaging in support of vaccination.

Background

Over the past twenty years [2], an increasing proportion of the U.S. population has come to view vaccines with skepticism and, in many cases, has opted out of getting themselves or their children vaccinated. The COVID-19 pandemic and vaccine rollout exacerbated this trend and heightened debates about the benefits of vaccines. While the procedural, behavioral, and access barriers to vaccine uptake have been studied extensively, the emotional, ideological, and rhetorical bases for vaccine hesitancy are poorly understood. This limited understanding undermines efforts to achieve high levels of vaccination coverage for COVID-19 and other diseases.

“We [Rishi Goyal and I] met a few years ago, before COVID. We were looking at hesitancy around the measles vaccine. At the time, the United States was in danger of losing its measles elimination status according to the WHO. We were also thinking about the various ways in which language shapes our perception of illness and health. Then we started to look at the way people express hesitancy online, and in literature, and in the news. As we were thinking about this, the COVID-19 pandemic happened, and the issue of vaccine hesitancy became more pressing.”
– **Dennis Tenen**, Project Lead

Recognizing that online forums have become the primary platforms for discussing and disseminating vaccine-related concerns, the project team set out to harness their unique backgrounds—emergency medicine, data science, and comparative literature—to unlock the meaning behind the language used by various vaccine-hesitant groups and use the analysis to reach these communities with tailored, positive vaccine messages. The project also engaged public health officials from the Maine Center for Disease Control and Prevention and the Department of Health and Mental Health in Ulster County, New York (NY) to inform the development of targeted, online pro-vaccine ad campaigns.

Objectives

The CWP *Increasing COVID-19 Vaccine Confidence* project, which brought together a unique, interdisciplinary team of literary scholars, medical professionals, data scientists, political scientists, community leaders, and public health officials, sought to use state-of-the-art data science methods combined with tools of literary scholarship to:

1. Collect a significant database of anti-vaccine rhetoric, found in online forums, discussion groups, and social media.
2. Analyze vaccine hesitancy as a cultural, linguistic phenomenon so as to better understand its causes and concerns.
3. Propose and implement new ways of presenting vaccines to the public that increase acceptance and participation, in partnership with public health practitioners and the private sector.

“What we’ve seen with COVID-19 is a greater politicization and polarization of vaccine hesitancy, but also a growing variety of logic and rhetoric. Vaccine hesitancy is not one thing and cannot neatly be understood in simple demographic terms.” – **Rishi Goyal**, Project Lead

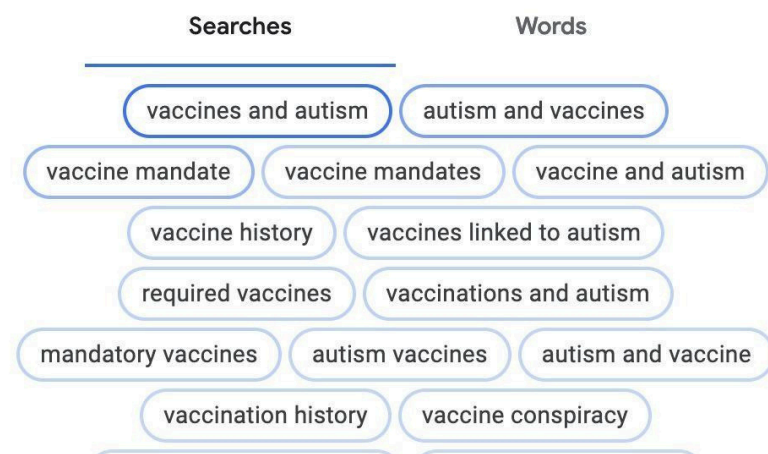
Methods

In the first phase of the project, coinciding with the rollout of the first COVID-19 vaccines in 2021, the team worked with Columbia Libraries to “scrape” and monitor social media forums and website conversations about vaccines. This resulted in the development of a significant computational public data set consisting of 2,090,054 sampled “vaccine” themed messages. Data sources included Facebook, YouTube, Reddit, Twitter, Parlor, and Telegram over the timeframe 2019-2022.

In the second phase, the team used Natural Language Processing (NLP), in combination with the tools of literary close reading [3], to analyze the language that people use around vaccination and vaccine hesitancy. The analysis revealed different categories, or “linguistic identities” of vaccine hesitancy, each characterized by clusters of recurrent topics and terms. The team sought input from public health partners at Columbia University’s Mailman School of Public Health (MSPH) and the Maine and Ulster County, NY health departments, including an analysis of vaccine hesitancy statistics, vaccination strategies and messaging, and observation at demonstrations and rallies in these jurisdictions.

In the third phase, the team developed four online, pro-vaccine campaigns tailored to the vaccine-hesitant identities defined in phase two. The team managed a six month (November 2023-April 2024) set of campaigns on Google Ads, which included tailored and untailored, 90-character textual snippets that met Google Ads requirements. Users who entered the search term “vaccine” plus any of the associated keywords (e.g., “autism,” “election”) were shown either the CWP project-developed, tailored ad or a generic, pro-vaccine ad developed by the U.S. Centers for Disease Control and Prevention (CDC). The team compared which ads got more user clicks.

Top searches and words within searches where people saw your ads



Most-shown search ads				
BEST_Vaccine Hesitancy Pilot (Parenting)_ad groups #2 Experiment				
<div> <div>Vaccine Autism Vaccine Fevers Vaccine Required</div> <div>[Ad] cu-hll.github.io</div> <div>Getting vaccinated helps protect your dear child's health! Mamas protect their dear children by keeping them caught up on vaccines.</div> </div>				
Ad	Status	Impressions	Clicks	CTR
Enabled	Not eligible	17,093	566	3.31%

Findings and Results

Analysis of the 2,090,054 sampled vaccine-themed messages resulted in a collection of general observations about the data which were further developed into distinct categories, or “linguistic identities” of vaccine hesitancy. **The vaccine hesitant “identities” were defined as: *political, race and ethnicity, natural healing, parenting, and religion*.** Characteristics of each of these identities informed the design of the four vaccine messaging campaigns on Google Ads (the campaign designed for the religion identity was not run, at the advice of the Maine and Ulster County health departments).

As a result of the campaigns, thousands of vaccine-hesitant searchers clicked through to accurate vaccine information. Users conducting a search for “vaccine” plus any number of the associated keywords within one of the identity categories were more likely to click on the ad tailored to that particular category versus the generic CDC ad. Overall, the campaign had **37,900 individual impressions** and **1,621 users** clicking through to the project-developed vaccine information. While the project did not measure vaccine uptake, when vaccine hesitant searchers clicked on the project-developed or CDC-developed ads, they were directed to accurate vaccine information. Rather than validating their vaccine hesitancy, the information offered reassurance about the safety and importance of vaccines.

Metrics showed that the use of search terms triggering the tailored messaging, as well as user engagement with those ads, validated the team’s findings about the linguistic identities. This was particularly true in the case of the campaigns focused on the *parenting* and *political* identities of vaccine hesitancy. With the *parenting* campaign, the modified messages received a total of 566 user clicks, 173% more than the generic CDC campaign. The campaign targeted at *political* vaccine hesitant identities, had 458 clicks versus 63 for the control campaign (726% higher). Results from the other two campaigns targeted to the *natural healing* and *race and ethnicity* identities showed that the CDC ad received more clicks than the project-developed ad, although far fewer ads ran in these categories and further analysis is needed to determine if the results are statistically significant.

Based on the results of the ad campaigns, the project leads provided consultation to the health departments in Maine and Ulster County, NY as they developed their COVID-19 vaccine and booster campaigns. Findings were shared with stakeholders early on through one-page *Key Insights* documents. Additionally, the project has trained a cohort of undergraduate and graduate students in data analysis and original research using the data set.

By monitoring social media online, we can better understand the reasons for vaccine hesitancy and tailor public health messaging accordingly to increase vaccine uptake. Examples using common vaccine hesitant themes and identities defined by this project include:

- **“My body, my choice.”** Vaccine hesitant activists emphasize a perceived loss of agency, freedom, and consent. In response, emphasize bodily autonomy, choice, and control in the decision to take the vaccine.

- **“No vaccine passports.”** The public worries about medical privacy and the way their medical decisions may become subject to travel restrictions or employment mandates. Messaging should stress traditional mechanisms of protecting medical privacy.
- **Low level of trust in public institutions and “big-pharma.”** Messaging should highlight local, personal stories from public health officials, medical professionals, and local vaccine manufacturers. For example, the Pfizer vaccine manufacturing involves plants in Chesterfield, Missouri and Portage, Michigan.
- **Vaccine victimhood.** Anti-vaccine groups are effective in foregrounding, for example by holding vigils and memorials, emotional stories of those who were thought to suffer from the vaccine. In response, allow the community to grieve by commemorating the lives of COVID-19 victims.
- **Vaccine ingredients.** Health conscious constituencies, “natural” parenting advocates, and religious groups are extremely concerned about vaccine ingredients. Messaging should feature the transparency, safety, and sourcing of vaccine ingredients.

Key Outputs

- Public data set including 2,090,054 sampled “vaccine” themed messages from popular social media outlets, collected over three years (2019-2022).
- 4 tailored, pro-vaccine campaigns, run on Google Ads.
- Paper (in review) on vaccination in the context of folk belief.
- [LA Times Op-Ed](#): *How ‘my body, my choice’ came to define the vaccine skepticism movement*. Published on May 22, 2021.
- Three *Key Insights* documents: one on the [reasons behind vaccine hesitancy](#), another on the [style of a typical vaccine-hesitant post](#), and another on [online vaccine hesitant identities](#).
- An [award-winning poster](#), *The Poetics of Vaccine Hesitancy and COVID-19 Response in Muslim, Black, and Indigenous Communities*, by undergraduate research scholars Bethel Adiele, Mansi Garneni, and Adiba Hussain.
- [Your questions about the COVID-19 vaccine, answered](#). Detroit Free Press, Published September 14, 2021.
- [Vaccine Hesitancy: You can’t answer a feeling with a fact](#). KGW News (Portland, Oregon), Aired April 26, 2021.



Vaccine Hesitant Identities Online

POLITICAL

Common	Unique
People	Presidency/
Trust	presidential/pres
Administration	Senate
Election	Foxnews
Country	Starving
Plan	Trumpvirus
Work	Starving
Believe	Vacuna
Media	Donaldtrumpjr
	Disinfectant
	Gopchairwoman

PARENTING

Common	Unique
People	Hep
Children	Meningitis
Disease	Thimerosol
Flu	Teacozoy
Health	Wakefield
Years	Breastmilk/
Read	breastfed
School	Preschool
Measles	Formula
Risk	Cervical

RACE

Common	Unique
People	Supremacy
Black	Brown
White	Asian/Asians
Community	Latino/Hispanic
Gates	Futility
Government	Daughters
Trust	Privilege
Population	Epiphany
Country	Targeting
America	Historical
	Disparities

NATURAL

Common	Unique
Immune	Ecosystems
Body	Stecker
Vitamin	Inflammation
Healthy	Supplements
Flu	Sunlight
Herd	Minerals
Antibodies	Quercetin
Disease	Innate
Effects	Cycles
Fight	Organic
Risk	Memory

Why This Project Was Important

Vaccine hesitancy has risen sharply in the last two decades and is a major public health concern. Through this project and partnership with the Maine and Ulster County departments of health, a powerful case was made that vaccine-focused public health strategies must bridge the gap between the language of scientific communication and the language communities use to articulate their beliefs regarding medical knowledge to ultimately increase vaccine confidence and uptake.

The project data set is unique in that it was collected in real time during the rollout of a vaccine for a novel virus and also because it sampled a wide range of social media outlets, including mainstream sites such as Facebook, and sites such as Parler and Telegram that tend to reflect a broader spectrum of political ideologies. Additionally, a subset of the data set predates the COVID-19 pandemic, allowing users to compare and contrast forms of hesitancy. More broadly, this project showed the value of studying how diverse groups of people articulate their beliefs about medical knowledge, and how an interdisciplinary approach drawing on tools from data science, AI, public health, medicine, and the humanities can be effective in doing so.

“To help change a deep-seated reluctance to become vaccinated, we must understand the language of this anxiety—how people think and communicate. If we treat language as data—similar to the way we use epidemiological data to describe the epidemic itself, we can develop precise public health messaging to support vaccination. What might this messaging look like? We will need precision-tailored messaging strategies. We can oppose anti-vaccine language by turning its own syntax, word choice and emotive strategies against it. In countering ‘medical freedom’ discourse, for example, we need to employ the language of rights: ‘It is your right to get a vaccine.’ In essence, stand up and don’t let anyone deprive you of this right. For individualists who might be swayed by echoes of the 2nd Amendment: ‘Only you can protect yourself and your family.’ For freethinkers who do their own research: ‘No one else convinced me; I convinced myself to get the vaccine.’”

– excerpt from the OpEd: *How ‘my body, my choice’ came to define the vaccine skepticism movement*, published in the LA Times on May 22, 2021

Lessons on Implementing Social Impact Projects at Columbia University

The CWP *Increasing COVID-19 Vaccine Confidence* project exemplifies Columbia’s mission to leverage scholarly knowledge to create societal and global impact, in partnership with organizations and practitioners outside of academia. The project team and the methodologies used were truly interdisciplinary, drawing from literary studies, data science, medicine, and public health. Partnerships with public health officials in Maine and Ulster County, NY provided “on the ground” insights during the rollout of the vaccine campaigns and project findings about vaccine hesitancy and messaging. The dynamic public health environment during the pandemic required the project team to be flexible and adaptable while building a new research platform. The ability of the project team to adjust the project design and partnerships mid-course led to the largest data set of its kind that can now be utilized by other researchers across disciplines.

“When you look at the real world, it doesn’t have neat boundaries. It’s metaphor. It’s policy. It’s data. It’s medicine. This is a project that shows that the multiplicity of expertise can better address real world problems.” – **Dennis Tenen**, Project Lead

CWP as Catalyst

CWP funding and support enabled the project team to fill a gap in understanding vaccine hesitancy and to use these findings to craft more effective, pro-vaccine public health messaging—all in real time as COVID-19 vaccines were being rolled out across the U.S. This project has led the way for the team to establish the Health Language Lab (HLL) at Columbia University, a working group in collaboration with the [Incite Institute](#), [ICLS](#), the Data Science Institute, the Department of English & Comparative Literature, and the medical school. This interdisciplinary lab will continue to expand work on health messaging, including with public health practitioners focused on vaccine hesitancy.

Next Steps

The work of the project continues through the HLL. The project team is writing up project findings for public dissemination and will continue to share findings with public health communications teams and policy makers to improve vaccine messaging. Additionally, the team is working on publishing an open-source platform to house the dataset of vaccine hesitant language with the aim to promote new findings and new ways of “reading” online language.

Notes

[1] Vaccine hesitancy refers to delay in acceptance or refusal of safe vaccines, despite availability of vaccination services. Source: World Health Organization (2015). Vaccine hesitancy: A growing challenge for immunization programs. Accessed [here](#) on November 7, 2024.

[2] The rise of vaccine hesitancy before COVID-19 seems to be linked to Andrew Wakefield’s discredited work on autism and the link between it and the measles, mumps, and rubella (MMR) vaccine, which has been clearly debunked. As referenced in: <https://news.columbia.edu/news/covid-vaccine-hesitancy-project>

[3] Close reading, the dominant mode of analysis used in literary studies, is a method of literary interpretation which focuses on the rhetorical, structural, and figurative details of a text in order to perceive its deeper significance.