

Redirecting demand towards credible COVID-19 vaccine sources

Nishan Gantayat (1), Saransh Sharma (1), Jeff Mulhausen (2), Tim Sweeney (2)

¹ Final Mile Consulting, USA,

² Upstream Thinking, USA

Abstract. The menace of misinformation online is among the critical challenges facing societies and governments in the 21st century. The spread of false narratives and (mis)information around vaccinations in the context of the COVID-19 pandemic poses a serious public health concern. In many ways, the behavior of the unvaccinated population globally remains inconsistent and unexpected. Most of these people trust their doctors and have taken immunizations before. However, they are motivated uniquely by the context of COVID-19 and a desire for safety for themselves and family to reach their conclusions rather than rely on public health professionals. This study focuses on the African American population living in the southern states of the United States of America (USA). We used a sequential mixed method approach to analyze the nature of demand for alternative sources and counternarratives, and inform the need to design demand-side interventions to tackle the spread of misinformation on video-based platforms such as YouTube. Five core needs of users that underlie the demand-side of misinformation were identified based on the qualitative study. Building on the same, three behaviorally informed solution concepts to channel demand towards vaccine-confident content were developed in the generative phase and later evaluated using a quantitative study. In conclusion, the study adds to the existing literature on misinformation from a demand-side perspective and offers foundational insights for behaviorally-informed designs to address the demand for counter-narratives.

Keywords: Demand, Misinformation, Nudge

1 Introduction

Over the past decade, the world has witnessed unprecedented growth in access to information with a deepening of internet penetration (World Bank, 2021). With the internet, information is available at any time and everywhere. However, along with information proliferation, an ‘infodemic’, i.e., the menace of misinformation has emerged in the internet era. In simple terms, ‘misinformation’ can be defined as false or inaccurate information, and the same when deliberately spread with malicious content is referred to as ‘disinformation’ (UNHRC, January 2021). Quick access to information and removal of the gatekeepers of traditional media who make editorial decisions (Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012) has led to democratization of the publication space; however, it has also made the space extremely vulnerable to misinformation as traditional editorial filters for quality, veracity etc are no longer operational.

Misinformation cannot be characterized as merely the lack of relevant knowledge; instead, it is a belief in factually inaccurate information. The growth of the internet has caused a change in the nature and magnitude of the misinformation crisis even though they have existed for years in communication channels. Content on the internet is characterized by an increased ambiguity due to diverse perspectives and a diverse base of content creators, blurring the lines of expertise (Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012). By design, social media networks cause selective exposure and the formation of ‘echo chambers’ that facilitate and allows people to interpret and be exposed to information consistent with their existing beliefs and not the ones in conflict with their viewpoint (Vicario, Bessi, & Zollo, December 4, 2015).

The issue of misinformation does not merely influence individual decisions; instead, it has serious social repercussions. Factually incorrect information, when believed by a large group, can become the basis for political and societal decisions that are against the general interest of society and have been predicted to cause “digital wildfires. (Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012, World Economic Forum, 2013). A large-scale analysis of images on Facebook revealed that 23% of political images scanned had misinformation. It revealed the constant sexism and targeting of female politicians in these posts (Yang, Davis, & Hindman, 28 February 2023). The gravity of the harm caused by misinformation is further escalated in the case of matters related to health. A recent study analyzed the 50 most popular social media articles on breast cancer, prostate cancer, colorectal cancer, and lung cancer. A striking finding revealed that, of the 200 articles, 32.5% contained misinformation, and 30.5% had harmful misinformation (Johnson, et al., July 2022). A similar study in the context of Zika virus epidemic found that rumors had much higher popularity in terms of shares in comparison to verified news (Sommariva, Vamos, Mantzaralis, Dao, & Tyson, June 2018). Due to misinformation’s impact on people’s health and well-being, the need and search for effective ways of countering the ‘infodemic’ has become critical.

In recent years, governments, companies, and other multilateral organizations have taken proactive measures to provide corrective information and control the supply of misinformation online. One of the primary challenges countries face while legislating on controlling the supply of misinformation is striking a balance with freedom of expression (Smith, August 2020). Legislation in Germany to tackle misinformation has been criticized for its effect on free speech (Guerrini,2020). Studies have also shown detailed counters with context had more statistically significant efficacy against merely tagging misinformation on social media (Kreps & Kriner, February 2022). The tension between fair moderation and censorship highlights two related problems that arise in flagging online content as fake or legitimate: firstly, what kind of content counts as a problem such that it should be flagged, and secondly, what criteria must be used to label instances of such content in an unbiased manner?

This focus on intervening on supply-side of misinformation can be limiting. Our view is that till there is a widespread demand for alternative sources and counternarratives to the mainstream narrative, new supply sources will emerge to

cater to it, even if all the existing ones are eliminated. And even sincere efforts to counter misinformation on the supply side are complicated by concerns about civil liberties and political bias. Understanding the needs of individuals who are inclined and partial towards fringe media and counternarratives and fulfilling those needs with user-centric design of mainstream narratives and communication campaigns, can be a more sustainable approach to countering the spread of misinformation. Exploring misinformation through the lens of demand can engage academics and policy makers to the drive that fuels the perpetual supply of misinformation and thereby support the development of effective strategies that social platforms and even regulators can deploy to augment their supply-side interventions. Therefore, we propose that a major part of the long-term solution to the issue of misinformation lies on addressing the widespread public demand for alternative sources and counternarratives.

This study, done in the context of COVID-19, analyses how a desire for safety for self and family and an aim to arrive at one's conclusion drives people's demand for information. The demand-side of misinformation is characterized by the demand for alternative, often fringe, sources and counter-narratives to the mainstream ones. The study addresses the gap in the current literature regarding the demand for misinformation and proposes demand-side interventions against supply-side controls to arrest the consumption of misinformation.

2 Method

The study leveraged YouTube as a social media platform proxy for understanding and intervening in the consumption of vaccine-hesitant content among individuals who are not fundamentally vaccine hesitant. In this exploratory approach, the identified the needs that drive online content consumption behavior, as well as search & judgment levers that address both these needs and demonstrate the potential for nudge-based online solutions.

The study followed a sequential mixed-method approach which involved three phases- qualitative, generative, and then quantitative.

2.1 Qualitative Phase

A purposive sample of 18 participants was drawn to conduct six interviews each consisting of a triad. The participants were recruited using a recruitment screener that was administered using a survey. Participants who met the points of representation for the sample were selected into the sample. The sample represented Americans – aged 18+ years old, belonging to African American ethnicity, residing in a Southern US State, residing in a small town/rural geography (populations not residing in cities or nearby suburbs with populations approx.. 100,000 or more), had a prior history of vaccination for themselves or for their family, were not COVID-19 vaccinated, and

desired safety from COVID-19. Here the southern states refer to Texas, Louisiana, South Carolina, Georgia, Mississippi, Alabama, Tennessee, Arkansas and Missouri. At the time of this study the percentage of population fully vaccinated were approximately 52% and this number was lower for the southern states.

To gain a comprehensive understanding of the decision-making that supports the information consumption behavior of the participants on video-based social media platforms, we conducted a 60–90 minute semi-structured triad interview with the groups consisting of open-ended, sequenced questions in the presence of a research moderator. The interviews took place in October 2021. Pattern identification and theme analysis were used to analyze the qualitative data and organize the content from each triad into information origins, user needs, and information processing heuristics.

2.2 Generative Phase

The qualitative data along with secondary literature was reviewed and analyzed by a group of human-centered design (HCD) experts and behavior science researchers with a focus on developing nudge concepts intended to fulfill needs identified in the prior qualitative study for search and judgment separately. The ideation goals included addressing at least two core needs per intervention idea. Three behaviorally informed design concepts to channel the demand toward vaccine-confident content were developed.

1. When a user searches for any COVID-19-related content, at the top of the search results appears a bar showing a playlist containing user-sourced helpful videos with a message "other people like you have found these videos most useful in deciding about COVID-19 vaccines".
2. When a user views COVID-19 content, he/she is presented with an informational cue at the top of the comment section that presents an emotional sentiment analysis of the comment section.
3. A prompt, containing an alternate set of videos determined by a diversity-based algorithm, is presented to the user on his/her profile feed as well as during their specific content viewing session, alerting the user that that he/she may be missing out on diversity aspect due to algorithmic bias.

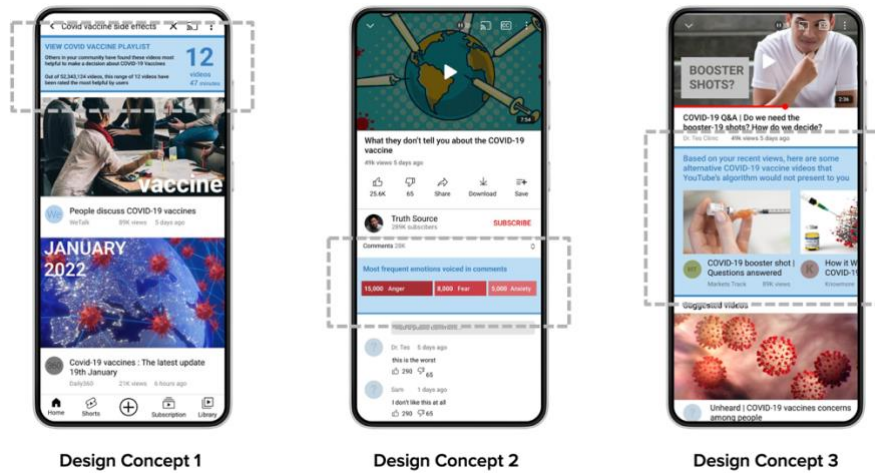


Fig.1. Design concepts from the generative phase of the research

2.3 Quantitative Phase

After administering a quantitative survey to 1842 individuals, 484 participants were recruited upon satisfying the recruitment criterion - African American individuals aged 18+ years old, residing in rural counties in Southern USA, not vaccinated for COVID-19, harbor need to protect themselves or their loved ones from COVID-19, having some history with past vaccination for themselves or with children, and have safety concerns for COVID-19 vaccines.

This phase of the research involved user-testing of the nudge concepts prototypes developed in the generative phase. The data collection was carried out using a structured online survey questionnaire as the research instrument. The survey questions, consisting of the prototyped nudge concepts (design concepts), were meant to - test the concept's effectiveness, helpfulness, and relevance in meeting the core needs, framing, and to rank each solution concept against each other on their inclusivity, upholding freedom to choose and need satisfaction characteristics.

3 Results and Findings

3.1 Qualitative Phase

We found that individuals began their COVID-19 vaccine consideration journey with a firm desire to make a well-informed decision to protect themselves and their loved ones. This drive to come to own conclusion stems from a historical lack of trust in the mainstream media, commercial health/pharma entities and regulatory bodies. During the pandemic, this distrust was amplified by instructive rather than persuasive style of communication, dismissing the concerns of the unvaccinated, media portrayal of

individuals who are not vaccinated and lack of any reliability markers for COVID-19 related information. The goal of making own decisions drives their information search on social media.

We identified 5 core needs regarding COVID-19 vaccine-related content that individuals try to fulfill during information search, evaluation and sharing on social media:

- (1) **Balance & Diversity:** The demand for counter-narratives is based on a desire to balance out a perceived mainstream bias.
- (2) **Freedom to Choose (Decision Autonomy):** Individuals are driven to make informed decisions for themselves and, therefore seek out content that informs rather than instructs.
- (3) **Trust:** Individuals prefer first-hand experiences of people due to an existing distrust towards institutions.
- (4) **Transparency:** Individuals collect all available information to judge incentives and inclinations of various sources supplying them with the necessary information.
- (5) **Certainty about the value of time spent:** Individuals avoid spending time on content that does not fulfill their needs.

From the qualitative study, some of the heuristics that were seen being applied for searching and judging video content were:

- Video links or mention of the video on a secondary source such as Yahoo News or other news websites serve as pathways for users to view the content on YouTube.
- Content is selected based on video length, title, familiarity with the channel, a Google search, number of subscribers, and recent comments.
- Individuals judge a video favorably if it contains authentic experiences and provides a diversity of views.
- The decision to continue is based on how little conflict is created with existing views, novelty, and diversity of the video.
- The decision to stop viewing the video is usually based on the feeling of fatigue, confusion, and repetition.

3.2 Quantitative Phase

A comparative result of the quantitative study conducted to test the effectiveness, helpfulness, and core needs fulfilled by the behavioral nudge has been presented in Table 1.0. A Likert-scale of 1-10 was used to collect participant's response. 1 being less effective and 10 being most effective.

Table 1.0 (Note: *% refers to the percentage of respondents in the quantitative study)

Design Concept No:	Effectiveness Rating	% rating the concept between moderate to extremely helpful	% rating the concept moderate to most likely to satisfy needs	% rating the concept as moderate to most inclusive and neutral	% rating the concept as moderate to most supporting autonomy
1	69% rated 6 or more.(Mean:6.57)	82%	75%	76%	78%
2	71% rated 6 or more.(Mean:6.52)	85%	67%	64%	61%
3	70% rated 6 or above (Mean 6.67)	80%	59%	59%	62%

According to the research findings, design concept 1, which proposed presenting users with a definite playlist of videos, was rated ahead of the other two concepts in terms of three of the four indicators, i.e., satisfying needs, being inclusive and neutral, and supporting autonomy. Design concept 2 proposed presenting users with information prompts highlighting the sentiments in the comment. It had the highest percentage of respondents that found the concept extremely helpful. Design concept 3 proposed presenting suggestive content for users that would otherwise be missed due to YouTube's algorithm. In terms of effectiveness, design concept 3 had the highest mean score. However, this concept had comparatively lower acceptance when compared to the other two design concepts on all four indicators.

In addition, some critical findings regarding the design of interventions include:

Design Concept 1:

- Videos rated helpful by other users are “more likely” to motivate users to click the playlist as compared to information on views and likes.
- Respondents preferred the framing of the message as ‘others similar to you’ over ‘other YouTube users’ signalling a preference for the wisdom of tribe than the wisdom of the crowd.

Design Concept 2:

- Respondents preferred framing of information in terms of ‘%’ over absolute numbers to evaluate sentiments of comments.
- Respondents preferred information that provides overall sentiment contained in comments over information that provided specific emotions contained in comments.

Design Concept 3:

- Flagging playlists as something that has been overlooked due to YouTube algorithm than individual judgment was preferred by users.
- Fear of missing out was not seen to be playing a role indicating that people exercise agency in search and choice and avoid feeling a sense of loss of time invested in past consumption behavior.

4 Discussion and Conclusion

The COVID-19 content landscape for those seeking to form their own opinion is heavily populated by highly accessible alternative narratives, that may be rooted in partial truths or blatant falsehoods or may be deliberately designed to spur actions away from vaccination. However, in the context of the pandemic, content that promotes vaccine confidence and uptake is not always accessible or appealing and falls short in addressing the 5 needs, whereas the anti-vaccine content specifically targeted these needs. Based on the results of the concept testing, we recommend 4 strategic features and concepts for social media platforms to adopt for meeting user needs and rechanneling demand towards credible sources and content:

1. Categorizing content ratings through the lens of “tribe” drives trust: Categorize content reactions and ratings by user specific attributes such as proximity, age, gender, race, etc.
2. Algorithm awareness can improve diversity, transparency autonomy, and thus drive trust in platform: Users know their content choices are skewed by algorithms when brought to their attention. Leverage this awareness to drive people to alternative content pathways.
3. Highest order aggregate data summaries enhance value of time: Instead of raw numbers or specific language, use summaries that put the least cognitive burden on users. Establishing content limits (ie specific number of videos) prevents confusion & frustration.
4. Decision- or domain-specific ratings drive trust: Frame content reactions and ratings to specific user decisions or goals (ie. most helpful for making a decision, provided most certainty).

The demand-side focused approach demonstrated in this study respectfully engages with the netizens who are partial towards fringe information sources and counternarratives to understand their context and needs that are being currently fulfilled by fringe sources and counternarratives. Public messaging, communication campaigns, media reporting and social media features need to respond to these user needs and become more user centric. Instead of restricting content and sources that contribute to misinformation, this approach aims to channel the demand tendencies towards credible information sources and narratives, and nudge users to restrict their consumption of misinformation using online behavioral interventions. The demand can be channelled away from misinformation and socially undesirable content, and towards socially

desirable content, by intervening in the processes of search, judging, and sharing of content on social media platforms. As their current audience moves away, creation and supply of misinformation and counternarratives will be disincentivized. We believe this is the only sustainable strategy to address the rampant spread of misinformation in the long-run.

References

- World Bank . (2021). *Individuals Using the Internet (% of population)* . Retrieved from World Bank : <https://data.worldbank.org/indicator/IT.NET.USER.ZS>
- UNHRC. (January 2021). *Using Social Media in Community Based Protection: A Guide* . UNHRC.
- Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and Its Correction: Continued Influence and Successful Debiasing. *Psychological Science in Public Interest* , 106-131.
- Vicario, M. D., Bessi, A., & Zollo, F. (December 4, 2015). The spreading of misinformation online. *PNAS* .
- Yang, Y., Davis, T., & Hindman, M. (28 February, 2023). Visual Misinformation on Facebook . *Journal of Communication* .
- World Economic Forum . (2013). *Global Risks 2013*. World Economic Forum .
- Johnson, S. B., Parsons, M., Dorff, T., Moran, M. S., Ward, J. H., Cohen, S. A., . . . Fagerlin, A. (July,2022). Cancer Misinformation and Harmful Information on Facebook and Other Social Media: A Brief Report. *Journal of the National Cancer Institute*, 1036-1039.
- Sommariva, S., Vamos, C., Mantzaralis, A., Dao, L. U.-L., & Tyson, D. M. (June,2018). Spreading the (Fake) News: Exploring Health Messages on Social Media and Implications for Health Professionals using a Case Study . *American Journal of Health Education* , 246-255.
- Smith, Z. M. (August, 2020). Part II: How Are Other Liberal Democracies Addressing False Information? In Z. M. Smith, *Burnt by the Digital Sun: HOW THE INFORMATION ENVIRONMENT IS TESTING THE METTLE OF LIBERAL DEMOCRACIES* (pp. 9-16). Center for Strategic and International Studies (CSIS) .
- Guerrini, F. (2020, March 3). *The Problems With Germany's New Social Media Hate Speech Bill*. Retrieved from Forbes : <https://www.forbes.com/sites/federicoguerrini/2020/03/03/the-problems-with-germanys-new-social-media-hate-speech-bill/?sh=39f75748592a>
- Kreps, S. E., & Kriner, D. L. (February, 2022). The Covid-19 Infodemic and Efficacy of Interventions Intended to reduce Misinformation . *Public Opinion Quarterly* , 162-175.