



TONY BLAIR
INSTITUTE
FOR GLOBAL
CHANGE

Combatting Vaccine Misinformation in the Age of the Internet

BENJAMIN TSENG

Contents

The Problem
The Opportunity
The Roots of Anti-Vax Sentiment
The Role of Online Platforms in Spreading Vaccine Misinformation
The Structural Impediments to Online Platforms Successfully Tackling Vaccine Misinformation
How to Make Progress
The Need for Global Collaboration

Published at <https://institute.global/policy/combating-vaccine-misinformation-age-internet> on December 17 2020

The Problem

Despite broad scientific consensus around the safety, efficacy and overwhelming public-health benefits of vaccination, unfounded concerns around vaccine safety (often called “vaccine hesitancy” or “anti-vax sentiment”) have achieved an alarming rate of acceptance. While fringe viewpoints have always existed, vaccine hesitancy has achieved mainstream attention in modern times due to [a discredited 1982 TV “documentary” called Vaccine Roulette](#) and [Andrew Wakefield’s since-retracted-and-debunked 1998 Lancet paper](#) alleging a link between MMR (measles, mumps and rubella) vaccines and autism in children. While anti-vax proponents remain in the minority, it’s clear that this is a large enough issue that global leaders need to take action.

This surge in anti-vax sentiment is widely believed to have resulted in a resurgence of measles in the developed world, including the United States ([over 1,200 cases of measles were reported in 2019](#)), despite the country’s decades-long effort to eradicate the disease through vaccination. Experts warn that continued anti-vax sentiment may continue to erode the “herd protection” needed by those who have compromised immune responses and people who cannot be or have yet to be vaccinated. This problem is compounded by the fact that [some published models](#) predict that, with current patterns of antivax-sentiment spread, anti-vaccination views will become the dominant view over the next decade. With the world still struggling to contain the Covid-19 pandemic, the issue of anti-vax sentiment has achieved special significance with surveys suggesting that [35 per cent of Americans would choose not get a free, FDA-approved vaccine against the disease](#). Such a substantial level of vaccine scepticism would significantly blunt the potential positive health and economic impact of a Covid-19 vaccination campaign.

While there are myriad forces at work, at its heart, the popularity of anti-vax sentiment represents an inability of authorities and experts to address vaccine misinformation. Not surprisingly, studies have shown that those who subscribe to anti-vax views also appear to be [more likely to subscribe to unfounded conspiracy theories](#), with some reports [linking those beliefs to opposition towards mask-wearing during the Covid-19 pandemic](#). As a result, vaccine hesitancy is a key front in the important battle between institutions and a rising tide of misinformation.

The Opportunity

While the challenge of grappling with anti-vax misinformation is significant, the nature of the problem makes the current moment an interesting sandbox for policymakers seeking to understand the most effective ways of combatting misinformation. The similarity in the mechanisms of spread and reinforcement as well as the strong association with other forms of anti-science and conspiracy-theory viewpoints suggest that the successful handling of anti-vax misinformation will have additional positive effects in dealing with misinformation more broadly.

Beyond the public-health imperative, anti-vax misinformation is also valuable to tackle as it represents a lower-risk misinformation category with which to start. Unlike other categories of misinformation, the issue of anti-vax misinformation has well-defined boundaries, lowering the risk of freedom of speech or political bias concerns that analogous efforts in other areas may engender. With clear support in the form of medical and scientific consensus and the obvious public-health implications, policymakers should also see less resistance (and likely support) from technology and media circles.

The Roots of Anti-Vax Sentiment

It is vital to understand the worldviews which form the underlying basis for anti-vax sentiment. While anti-vax sentiments capture an incredibly broad range of viewpoints – from anxieties around vaccination schedules to disputes about the severity of diseases and conspiracy theories about the use of vaccines to implement mind control – at their heart there are three core beliefs (see Table 1 below):

- Vaccine safety is suspect
- Scientific authority is not absolute
- Big business and government are untrustworthy

Table 1 – The ideas that underpin anti-vaccine sentiment

| Categorisation of Ideas | Vaccine Safety Is Suspect | Science Doesn't Know Everything | Big Business and Government Are Untrustworthy |
|-------------------------|---|--|--|
| Mainstream | May lack scientific knowledge, but trusts the scientific/medical field's judgement on vaccine safety and efficacy | May have confusion about rapid change in scientific understanding about what counts as healthy diet and behaviour, but generally trusts scientific authority | Government regulations, medical providers and pharmaceutical companies are working in concert to deliver vaccines and treatments that help alleviate disease |
| | Concerned about vaccine side effects and may confuse accelerated funding and safety reviews with rushed development of vaccines | Believes that alternative medicine (homeopathy, naturopathy, traditional ethnic medicine, etc.) has positive health value | Big Pharma's profit motive may be suspect but regulations and well-trained physicians make sure medicine and vaccines are beneficial |
| | Believes vaccines cause serious illness (e.g., autism) and provide little health benefit | Believes conventional medicine represents a dangerous orthodoxy and poses a danger to people's health | Profit motive has corrupted regulatory oversight and physician incentives, casting doubt on all medical recommendations of drugs and vaccines |
| Extreme | Corporations, governments and the medical establishment are conspiring against the well-being of the public for their own gain and use vaccines to perpetuate the interests of the elite above those of the public. This may extend into anti-Semitic globalism conspiracies, as well as outright Covid-19 denialism. This can result in the belief that vaccines have greater risk than reward and are intrusive to liberty, or in conspiracy theories around vaccines as a means of delivering tracking microchips or mind control, or as form of sterilisation or unethical experimentation. | | |

These beliefs can manifest independently of one another – for example, concern about the development of a vaccine does not necessarily entail suspicion of pharmaceutical companies – and they tend to exist along a spectrum. Often the starting points are entirely understandable: Confusion

around why science seems to change its position so frequently on what constitutes healthy foods and behaviours, concern about the magnitude and frequency of vaccine side effects, scepticism about motives of large pharmaceutical companies, broader mistrust of the health-care establishment linked to concerns around racism and colonialism (from ethnic minorities and those outside of the West), and a broader cultural shift towards valuing more “natural” lifestyles are all, in isolation, not imminently harmful.

The problem emerges when, absent of intervention, individuals with these starting points become “radicalised” by bad actors who push them towards more extreme viewpoints. Research on people who “convert” to anti-vax sentiment on [Twitter](#) and conspiracy theories more broadly on [Reddit](#) shows that “converts” tend to exhibit characteristics which predispose individuals to conspiratorial thinking (negative sentiment, greater anxiety and paranoia, etc.), but a key factor in their conversion are the external factors (greater engagement with misinformation, active conversations with conspiracy theory proponents, etc.) which push them down the spectrum towards more extreme views.

Once these extreme viewpoints take hold, they are extremely difficult to dislodge as they are generally rooted in unfalsifiable logic and emotional appeals which play to an individual’s pride (for example, the idea that “only those who have opened their eyes can see the truth of what vaccines are doing to our children”), altruism (“we have to prevent children from suffering from vaccine-linked autism”) and scepticism (“the pharma companies are just trying to push something you don’t need to make money off of you!”). [One randomised study on combatting vaccine misperceptions published in the medical journal *Pediatrics*](#) suggests that directly confronting individuals holding anti-vax views with refutations is oftentimes ineffective and, in some cases, backfires by providing an opportunity to amplify those viewpoints. Holders of these viewpoints may also begin to radicalise those they are connected to as well. This, coupled with [findings that holders of anti-vax sentiment appear to begin with less certainty](#), suggests that effective intervention needs to happen earlier on the path to “radicalisation”.

The Role of Online Platforms in Spreading Vaccine Misinformation

While fringe viewpoints have always existed, the outsized role that online platforms now play in how people access news has been a key catalyst in their spread. Published studies have indeed confirmed the role of online platforms as a conduit for vaccine misinformation. [One study published in May 2020 in the journal *Nature*](#) on the “online ecology” of the vaccination debate found that anti-vaccination Facebook pages with around 4 million members appeared to be far more active in engaging with “undecideds” than members of communities which were pro-vaccination. [A study published in *The Permanente Journal* in 2018](#) on a random set of tweets bearing vaccine-related hashtags showed that anti-vax tweets tended to experience higher engagement in the form of retweets, meaning these messages are being distributed to a wide audience. Parents of young children, in particular, appear to be targeted with this content. [A survey conducted by the Royal Society for Public Health](#) found that, in the UK, around 50 per cent of parents of children under the age of 5 were exposed to negative messages about vaccines on social media. This has resulted in an alarming normalisation of anti-vax sentiment in some online circles with likely spill-over effects for offline conversations as well.

The easy spread of anti-vax misinformation reflects the underlying product architecture and design decisions behind the major online platforms. While not intended to spread misinformation, the platforms’ emphasis on engagement pushes features and practices with wide-ranging consequences for the spread of misinformation. First, they can inadvertently prioritise the spread of emotionally appealing misinformation over real information or rebuttals. [A study published in *Science* in 2018](#) looked at over 120,000 “rumour cascades” on Twitter and found that falsehoods were much more likely to spread farther (in terms of the number of people reached and the number of successive re-shares) than truths (as judged by a consensus of fact-checking organisations).

Second, the existence of algorithmic feeds and the fact users select which pages and profiles to follow can lead to “filter bubbles”. While the impact of this on aggregate political polarisation is a subject of academic debate, studies point to this functionality as creating “echo chambers” where misinformation can self-amplify. [One “algorithmic audit” of YouTube’s search and recommendation algorithms published in May 2020](#) found that watching videos containing misinformation skews YouTube’s future search results and recommendations towards surfacing further misinformation. Additional studies on different social-media services and on discussions about global warming on Twitter suggest the combination of recommendation algorithms and users choosing which groups they engage with can lead to people mainly engaging with like-minded individuals, shielding misinformation from rebuttal and acting as an amplification mechanism.

The impact of online-platform feature design on adoption of non-mainstream conspiratorial viewpoints is borne out in empirical data across France, Germany, Britain and the US. Responses from the [TBI Globalism Study](#) – a global polling project conducted by YouGov between July and August 2020 in collaboration with the Tony Blair Institute for Global Change, the *Guardian* and researchers at the University of Cambridge – showed that individuals who use social media to discuss or read about current affairs frequently, compared to those who do not, are more likely to believe a wide range of conspiracy theories. While there is cultural variation in the results, the pattern largely holds across a wide variety of non-mainstream views, including several key health-and-science-oriented conspiracy theories.

[As we explain in a companion piece](#), this data should be treated carefully and used only as grounds for further research, rather than to draw definitive conclusions. It’s clear, however, that while anti-vax sentiment is a minority view, it’s sufficiently widespread that it could threaten vaccine-enabled herd immunity and the underlying conspiracy thinking could further pollute public discourse. It’s therefore essential that leaders act now.

Figure 1 – Belief in select conspiracy theories, cut by use of social media for reading or discussing current affairs, by country

Q Would you say the following statement is true or false? (charts 1, 2, 4, 6). Thinking about coronavirus...As far as you know, are each of the following statements true or false? (charts 3, 5). (Please select from one option on each row: Definitely true/Probably true/Probably false/Definitely false/Don't know.)

The Structural Impediments to Online Platforms Successfully Tackling Vaccine Misinformation

Despite broad consensus that action needs to be taken to curb misinformation, most online platforms have shown hesitancy to act. This is driven by three key factors.

First, platforms are caught between business models that monetise user engagement and content-moderation issues that are uniquely difficult to tackle. Although misinformation is an unintended by-product, platforms' financial success has largely been built on cultures and processes which remove friction or limitations on highly engaging content. Set against examples of borderline content that are genuinely hard to create policies for, companies can be stuck between a rock and a hard place even if action is made a priority by senior management. This internal tension has resulted in shifting policies that are not always coherent. For instance, in March 2019, Facebook announced that it would ban ads that promoted anti-vax misinformation. Later, it was revealed that this policy was focused on ads which specifically contained anti-vax misinformation, but not ads which promoted less explicit anti-vax sentiment. Facebook then amended its policy in October 2020, banning all ads that discourage vaccination, but not Facebook pages or groups that actively promoted antivax content. Finally, it further amended that decision in December 2020 to remove all false claims about Covid-19 vaccines. This public back and forth (and reports of internal struggles within Facebook) reveals the underlying tensions online-platform owners face.

Second, the platforms are wary of the cost and feasibility of mass-scale misinformation regulation. This reflects the difficulty in determining, even for human fact-checkers, what counts as misinformation. With regards to breaking news, the truth can often take a significant amount of time and fact-checking resources to ascertain. Individual articles or posts can also be a mix of truthful and untruthful points, which complicates determining what sufficiently counts as misinformation. Reliance on external parties to make these judgements is also fraught with risks of political bias or capture – as even usually authoritative sources can make mistakes (for example, the WHO's and CDC's initially more lax position on mask-wearing was later revised).

Given these challenges, it is no surprise that technological solutions for regulating misinformation at scale are limited. While artificial intelligence and natural language processing have become increasingly adept at understanding the *meaning* of content, they have not yet proved their understanding of the *validity* of content. Some of the more promising efforts are based on matching claims in a post or article with previously established falsehoods. As a result, regulation of misinformation requires platforms to implement human moderation, which brings significant operational complexities with it. Not only is moderation costly, but it also creates headaches around how to set content policies and handle appeals. There have also been concerns raised in the past about the exploitation of moderators.

Third, the platforms generally do not want to be viewed as censors. This is as much a philosophical stance as a practical one. Many founding teams of social-media platforms view their role as providers of neutral tools to democratise and decentralise communication, in stark contrast to a legacy media environment where editors had control but also acted as gatekeepers against new ideas and voices. As a result, many of their philosophical starting points around misinformation are rooted in their intentions to be neutral and the faith that a free flow of ideas is better for users than centralised control or censorship.

There are also numerous practical concerns that the platforms have about being viewed as censors. One issue relates to the difficulty of regulating private online conversations. The platforms have broad reach and are able to support a large number of users within a single messaging group (WhatsApp supports up to 256, Twitter up to 50, Facebook up to 150), which makes it easy for any content (including misinformation) to be rapidly shared by sending a single message. As a result, these social platforms have now become quasi-public channels which some researchers have identified as a key conduit for misinformation. But, user expectations around privacy (sometimes enforced technologically through end-to-end encryption) make regulating and monitoring this challenging for the platforms.

Another practical issue is centred around legal concerns. While many countries have legislation and directives in place protecting the right of online platforms to engage in “good Samaritan” moderation (Section 230 of the Communications Decency Act in the US, the Electronic Commerce Directive 2000 and subsequent guidelines from the European Commission in the EU), the political scrutiny and calls for regulation they routinely meet as a result of any content policies they take on have understandably made many hesitant to act aggressively.

How to Make Progress

The central challenge with the spread of anti-vax misinformation (and misinformation more generally) is one of trust. Lack of trust in those in positions of authority and the poor availability of trusted information in the right forums have contributed to a dangerous rise in anti-vax sentiment.

Figure 2 – Trust in government as an accurate source of information about coronavirus

Q How much, if at all, do you trust **The Government** in [country name] in general to be an accurate source of information about coronavirus? (Please select one option on each row: A great deal/A fair amount/Not much/Not at all/Don't Know/Prefer not to say.)

To combat this rising tide, policymakers need to adapt their approaches to not only account for the key role that a rapidly changing online environment plays in their spread, but also to leverage that environment to build and restore trust in institutions and policy. A comprehensive approach would embrace three pillars.

First, it is necessary to embrace a modern regulatory framework that balances harm reduction with freedom-of-speech concerns and the practical need to enlist platform owners to participate. The objective of policy should be to incentivise platforms to take action against misinformation, recognising that the most effective procedures for this may vary across platforms. This requires balancing the desire to reduce harmful misinformation with the need to avoid curbing legitimate speech (e.g., scientific discussions about vaccine safety and efficacy, vaccination policy disagreement, etc.). In practice, policymakers should therefore encourage proactive monitoring on platforms to curtail misinformation while providing strong “good Samaritan” legal safeguards and other inducements.

However, for this twin-track approach to work, strong transparency and procedural review and audit systems must be in place to understand how platforms are meeting their obligations. The rapidly evolving nature of misinformation and online platforms requires a regulatory regime that can adapt to changing technologies and situations. Prioritising procedural transparency and review can enable an ecosystem of auditors that can find issues, make recommendations, encourage platform providers to adapt their rules continuously rather than simply wait and see what policymakers explicitly call for, and reduce broader information asymmetries between policymakers and platforms. It can also foster trust with platform users and stakeholders around the intentions and effects of moves by the platforms to curb misinformation, even when they involve new rules and monitoring that did not previously exist.

Second, governments should make it easier for online platforms to curb anti-vax misinformation. Where possible, public-health agencies should make available high-quality, layperson-friendly, de-politicised public-health information that platforms can use. As the pandemic has shown, institutions themselves are not infallible, so this will require far more humble political leadership and openness about scientific evidence. More readily available objective information may also make it easier for platforms to prioritise and amplify posts from reliable, trustworthy practitioners, who could also benefit from basic social-media and science-communications training delivered by platforms and public bodies to make their rebuttals and points more salient in the debate. Improved official communication could also help reach people who are sceptical about vaccinations but are not social-media users.

Governments should also collaborate with platforms to equip their citizenry with basic training and information on how to detect and resist influence and misinformation campaigns. While this is ultimately a long-term strategy with a long-term payoff, this is necessary to deal with the ever-present challenge of widespread misinformation that cannot be tackled with technical or content-policy solutions alone.

Governments can also exercise their regulatory authority and convening power to encourage platform operators to share data they have gathered on anti-vax misinformation – such as the strategies and accounts of bad actors, information to pinpoint malicious foreign actors or networks of bots, the narratives or talking points which are gaining traction, and other best practices – to better coordinate responses. Regulators should also consider approaches to encourage platform operators to experiment with different tactics (e.g., the impact of removing a sharing feature or adding an element of friction, like Twitter has done by adding a prompt to encourage users to read an embedded link before retweeting it) and make the findings available for others.

Lastly, public-health officials should leverage online platforms to develop social-mobilisation campaigns around vaccination. Studies on the online vaccination debate suggest that pro-vaccine advocates have unfortunately ceded significant ground by not engaging on the same emotional level as anti-vax proponents or participating in the conversations where “undecideds” are seeking information. This vacuum of ambiguity is then filled with misinformation and conspiracies, rather than reassuring or explanatory content. With the impending arrival of well-validated Covid-19 vaccines, this represents a key public-health information gap, but one which is also an opportunity for public-health officials to take advantage of, by using social-mobilisation methods which focus on empowering communities at all levels to promote public-health messaging.

To support broader social-mobilisation efforts on vaccination, public-health agencies should take advantage of online platforms to proactively monitor anti-vax conversations and adapt their messaging in response to new talking points and themes that emerge. They should also employ these platforms to identify and reach out to key community leaders and influencers to coordinate on messaging. With these platforms, experts and public-health officials can also participate directly in groups and conversations to help break down “filter bubbles” and “echo chambers” and productively intervene with “undecided” individuals or “softer sceptics” who might otherwise be further radicalised. One example of this having been successfully deployed is in Taiwan, where the government is a pioneer in digital civic engagement. They famously used memes and internet tools to convey public-health guidelines and respond to questions and concerns from the public around Covid-19 in a refreshingly digital-native fashion.

保護自己也保護他人
跟總柴一起防疫新生活

肥皂勤洗手

出入實名制

維持社交距離

室內 1.5 公尺

室外 1 公尺

無法維持者請配戴口罩

2020.07.11

Taiwan’s Covid-19 public-health guidelines have made use of memes involving Zongchai, a real Shiba Inu dog belonging to an officer of the Ministry of Health and Welfare. *Translation: Headline: Protect yourself and others! Subtitle: Join Zongchai in embracing anti-pandemic habits; Upper-Left: Wash*

Your Hands with Soap; Upper-Right: Real name contact tracing; Middle Bottom: Maintain social distancing (1.5m indoors / 1m outdoors). Bottom: If you cannot socially distance, please wear a mask. Source: [Taiwan Ministry of Health and Welfare's Twitter account](#).

The Need for Global Collaboration

Anti-vax sentiment unfortunately plagues all inhabited areas of the world. While published studies have generally shown anti-vax sentiment to be more prevalent in Europe and North America, recent measles outbreaks in Asia and Brazil and widespread mistrust about vaccination in Africa reveal a troubling trend globally. The cultural and political factors involved as well as the identities and strategies of the actors responsible for spreading anti-vax sentiment will vary, such as strong links between anti-colonial and anti-vax movements outside the West. But, sadly, part of the problem does appear to be linked to the spread of misinformation from the West.

To tackle vaccine scepticism within their borders effectively, global policymakers will need to grapple with regional-specific sources of misinformation, bad actors who spread it, and attitudes and cognitive biases that underlie anti-vax sentiment. However, there is no doubting the need for greater international collaboration to tackle the pandemic of anti-vax sentiment. Only global cooperation can manage the challenges generated by the combination of global platforms, a global internet and a global pandemic.

In an interconnected world, the risks that pandemics present are profound and public-health must be treated as a global public good. No single country is immune to what happens elsewhere, and this applies as much to information about vaccinations as to the spread of the virus itself. The key online platforms that enable the spread of misinformation at scale have multinational footprints, which may leave smaller nations lacking the wherewithal to exert regulatory pressure in isolation. The transnational nature of some of the bad actors and cultural forces involved also suggests that, at the minimum, regional multinational collaboration, even if more informal, would pay dividends in controlling anti-vax misinformation.

In recent months, the institutions designed to deliver this type of coordination have often fallen short, but it is clear that the need has never been greater. Anti-vax views may currently be in the minority, but they are already sufficiently large to present a significant public health risk. However, it's also clear that there is an opportunity, too. The recommendations set out in this report show how a new compact between governments and social-media platforms – built on transparency, strong communications and social mobilisation – could help to stem the flow of anti-vax misinformation and secure public health for the long term.

Editor's Note:

Social Media Use Survey Question: *Generally speaking, how often, if at all, do you use social media to discuss or read about current affairs? (Several times a day or more / Once a day / Every few days / Once a week / Less than once a week / Not applicable - I do not use social media to discuss or read about current affairs.)*

Some data in charts and text may also vary slightly due to rounding. Participants for the survey were selected from an online panel, which should be taken into account in responses to questions about online activities, particularly in countries with low levels of internet access. More information about the research and results can be found here: <https://yougov.co.uk/topics/yougov-cambridge/globalism-project>

FOLLOW US

facebook.com/instituteglobal

twitter.com/instituteGC

instagram.com/institutegc

GENERAL ENQUIRIES

info@institute.global

FIND OUT MORE
INSTITUTE.GLOBAL

Copyright © December 2020 by the Tony Blair Institute for Global Change

All rights reserved. Citation, reproduction and or translation of this publication, in whole or in part, for educational or other non-commercial purposes is authorised provided the source is fully acknowledged. Tony Blair Institute, trading as Tony Blair Institute for Global Change, is a company limited by guarantee registered in England and Wales (registered company number: 10505963) whose registered office is One Bartholomew Close, London, EC1A 7BL.