



Evaluating COVID-19 Vaccination-Related Speech Policies on Social Media platforms

The tables below correspond to the social media platforms' policies as they currently stand as of the publication of this blog post on February 11, 2021. The tables provide analysis on vaccine advertisement policy, the categories of vaccine narratives and their corresponding ratings, and platform moderation strategy.

To be considered comprehensive for each of the different categories:

Category	In order to be comprehensive, the policy must have:
Safety	The policy clearly states that it will address false or misleading claims about <i>both</i> the contents of the vaccine and if there are any side effects or adverse health risks. The policy provides examples of what this type of content would look like on their platform.
Efficacy and Necessity	The policy clearly states that it will address <i>both</i> false or misleading claims about whether or not the vaccine is effective and whether or not it is necessary to receive in order to prevent against COVID-19.
Vaccine Development and Government Distribution	The policy clearly states that it will address false or misleading claims about <i>both</i> the process of making the vaccine and how the vaccine will be distributed. For the distribution of the vaccine, platforms should also address false or misleading claims about government mandates.
Conspiracy	The policy clearly states that it will address known conspiracy theories and it provides transparency into what those conspiracies are.



Facebook

Facebook first tackled vaccine-related misinformation on its platform in [March 2019](#). Facebook geared the focus of these policies towards reducing vaccine misinformation, especially through the discoverability and content from Pages and Groups that spread false information, and elevating authoritative sources and introducing educational pop-ups on the app. Facebook also removed vaccine controversies as a targeting option for ads, however a spokesperson stated that the policy does not ban advertising expressing opposition to vaccines when BuzzFeed News [found anti-vaccine ads](#) running on its platform. Facebook has since updated its advertisement policy and [does not allow](#) ads to discourage people from vaccination or advocate against vaccines. In [December 2020](#), Facebook published its plan to remove false claims COVID-19 vaccines and promote public health messaging. The company [sees this policy](#) as another way that we are applying our policy to remove misinformation about the virus that could lead to imminent physical harm. More recently, on [February 8, 2021](#), the platform expanded its policies to remove a wider list of claims related to vaccine misinformation.

Table Key: Red text reflects Facebook’s policy changes made February 8, 2021.

Category	Policy Language	Corresponding Rating
Safety	<ul style="list-style-type: none"> ● This could include false claims about the safety, efficacy, ingredients or side effects of the vaccines ● Claims that COVID-19 vaccines are the cause of or infect people with COVID-19 ● Claims that COVID-19 vaccines kill or seriously harm people ● Claims that COVID-19 vaccines cause autism ● Claims that building immunity by getting COVID-19 is safer than getting the vaccine ● Claims that COVID-19 vaccines are unsafe for a specific group of people, if that group is identified based on protected characteristics or other identifiers not directly related to their personal health, age, or disabilities (e.g. social status, religion, or political views) ● Claims that the COVID-19 vaccine changes people’s DNA 	Comprehensive. Facebook’s policy clearly states that it will address different claims commonly associated with vaccine safety, including vaccine ingredients and side effects.



	<ul style="list-style-type: none"> ● Claims that COVID-19 vaccines cause infertility ● Claims that COVID-19 vaccines contain toxic, prohibited, or harmful ingredients, microchips, animal products, or anything not on the vaccine ingredient list ● Claims that COVID-19 vaccines contain the mark of the beast <p>Ex:</p> <ul style="list-style-type: none"> ● “The COVID vaccine causes autism!” ● “The COVID vaccine will kill you” ● “The COVID-19 vaccine causes COVID-19!” [Feb. 8] 	
<p>Efficacy and Necessity</p>	<ul style="list-style-type: none"> ● This could include false claims about the safety, efficacy, ingredients or side effects of the vaccines ● Claims that COVID-19 vaccines do not provide any immunity to people ● Claims that COVID-19 vaccines are not effective in preventing COVID-19 <p>Ex: “The COVID vaccine provides no immunity”</p>	<p>Non-Comprehensive. While Facebook’s policy clearly addresses the efficacy of vaccine-related speech, it does not address content that falls under claims of whether the vaccine is necessary in order to prevent against COVID-19.</p>
<p>Vaccine Development and Government Distribution</p>	<ul style="list-style-type: none"> ● Content coordinating interference with the administration of the COVID-19 vaccine. ● Content calling to action, advocating, or promoting that others not get the COVID-19 vaccine. ● Claims that COVID-19 vaccines do not exist or have not been approved ● Claims that COVID-19 vaccines are untested ● Claims COVID-19 vaccines are not tested against a placebo during clinical trials ● Claims that people died as a result of the COVID-19 Pfizer/BioNTech vaccine during clinical trials (Note - We allow claims that people died during the COVID-19 Pfizer/BioNTech clinical trials) 	<p>Non-Comprehensive. While Facebook’s policy addresses coordinating interference with the distribution of the COVID-19 vaccine, it does not address <i>false claims</i> about government distribution of the vaccine.</p>



<p>Conspiracy</p>	<ul style="list-style-type: none"> Facebook will also remove conspiracy theories about COVID-19 vaccines that we know today are false: like specific populations are being used without their consent to test the vaccine's safety Facebook will continue to update the claims they will be removing based on guidance from public health authorities Claims about the side effects of COVID-19 vaccines which are incredulous or irrational, such as taking the vaccine turns you into a monkey Claims that COVID-19 vaccines are designed to or were developed in order to control a population for non-public health purposes Claims that specific populations are being used or targeted in order to test the true safety or efficacy of a COVID-19 vaccine <p>Ex:</p> <ul style="list-style-type: none"> "The COVID vaccine contains a microchip" 	<p>Comprehensive. Facebook's policy clearly states it will address known conspiracy theories and provides examples of what those are. In addition, Facebook's new policy provides a list of widely debunked claims by WHO and the CDC</p>
<p>Platform intervention (content-based level)</p>	<p>Distribution/ Labels</p> <ul style="list-style-type: none"> Once a piece of content is rated false by fact-checkers, we reduce its distribution and show warning labels with more context <p>If users interacted with COVID-19 misinformation, they will:</p> <ul style="list-style-type: none"> Receive a notification that says we've [Facebook] removed a post they've interacted with for violating our policy against misinformation about COVID-19 that leads to imminent physical harm. Once they click on the notification, they will see a thumbnail of the post, and more information about where they saw it and how they engaged with it. They will also see why it was false and why we removed it (e.g. the post included the false claim that COVID-19 doesn't exist) People will then be able to see more facts about COVID-19 in our Coronavirus Information Center, and take other actions such as unfollowing the Page or Groups that shared this content. 	



<p>Platform intervention (Pages/Groups-based level)</p>	<p>(as of March 7, 2019)</p> <ul style="list-style-type: none"> • 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. • When we find ads that include misinformation about vaccinations, we will reject them. We also removed related targeting options, like vaccine controversies. For ad accounts that continue to violate our policies, we may take further action, such as disabling the ad account. • We won't show or recommend content that contains misinformation about vaccinations on Instagram Explore or hashtag pages. • We are exploring ways to share educational information about vaccines when people come across misinformation on this topic. <p><i>Update on April 26, 2019 at 10AM PT:</i> We may also remove access to our fundraising tools for Pages that spread misinformation about vaccinations on Facebook.</p> <p><i>Update on February 8, 2020</i></p> <ul style="list-style-type: none"> • Pages, Groups, profiles, and Instagram accounts that repeatedly post misinformation 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.
<p>Vaccine Policy related to Ads</p>	<p>Facebook will not allow:</p> <ul style="list-style-type: none"> • Ads must not discourage people from vaccination or advocate against vaccines. This includes the claims introduced in Facebook's policy on February 8th, 2021. <ul style="list-style-type: none"> ◦ Ads which discuss vaccine legislation are allowed. • Ads that exploit the pandemic for commercial gain • Ads or organic posts that promote the sale of a COVID-19 vaccine, such as attempts to sell COVID-19 vaccine kits or expedited access to the vaccine • Ads that claim the vaccine is a cure for the virus <p>As of December 18, 2020, Facebook will allow ads that highlight the ability of a COVID-19 vaccine to prevent someone</p>



	from contracting the virus, as well as ads promoting ways to safely access a COVID-19 vaccine
General vaccine policy created in collaboration with global health organizations	<ul style="list-style-type: none">● Vaccines cause autism<ul style="list-style-type: none">○ Ex: Increased vaccinations are why so many kids have autism these days.● Vaccines cause Sudden Infant Death Syndrome<ul style="list-style-type: none">○ Ex: Don't you know vaccines cause SIDS?● Vaccines cause the disease against which they are meant to protect, or cause the person to be more likely to get the disease<ul style="list-style-type: none">○ Ex: Taking a vaccine actually makes you more likely to get the disease since there's a strain of the disease inside. Beware!● Vaccines or their ingredients are toxic, poisonous, harmful, or dangerous<ul style="list-style-type: none">○ Ex: Sure, you can take vaccines, if you don't mind putting poison in your body.● Natural immunity is safer than vaccine acquired immunity<ul style="list-style-type: none">○ Ex: It's safest to just get the disease rather than the vaccine.● It is dangerous to get several vaccines in a short period of time, even if that timing is medically recommended<ul style="list-style-type: none">○ Ex: Never take more than one vaccine at the same time, that is dangerous - I don't care what your doctor tells you!● Vaccines are not effective to prevent the disease against which they purport to protect<ul style="list-style-type: none">○ Ex: Vaccines actually don't do anything to stop you from getting the disease.

Twitter



Like Facebook. Twitter also addressed the quality of vaccine-related information on its platform around the [middle of 2019](#). The thrust of these policies was providing a search prompt that linked to credible public health resources when people search certain terms related to vaccines. On [December 16, 2020](#), Twitter announced its approach to misleading COVID-19 vaccine information. The announcement is an expansion of the [platform's ongoing policies](#) against COVID-19 misinformation. In order for Twitter to take action on the COVID-19 related content, the tweet must:

- advance a claim of fact, expressed in definitive terms;
- be demonstrably false or misleading, based on widely available, authoritative sources; and
- be likely to impact public safety or cause serious harm.

Twitter stated in another [blog post](#) two days later that it already didn't allow advertisements that contain misleading claims about the cure, treatment, diagnosis or prevention of certain diseases and conditions, including vaccines.

Category	Policy Language	Corresponding Rating
Safety	<ul style="list-style-type: none"> • The safety or efficacy of treatments or preventative measures that are not approved by health authorities, or that are approved by health authorities but not safe to administer from home • False claims which have been widely debunked about the adverse impacts or effects of receiving vaccinations • How vaccines are developed, tested, and approved by official health agencies as well as information about government recommendations. 	Non-Comprehensive. While Twitter's policy addresses vaccine safety as a whole, and specifically mentions side effects, its policy does not specify other key claims related to vaccine safety including vaccine ingredients. Twitter's policy for claims about how vaccines are developed may address vaccine ingredients. However, false claims about vaccine ingredients is an important trope in vaccine misinformation and therefore should be stated clearly in the policy.



Efficacy and Necessity	<ul style="list-style-type: none"> • False claims that COVID-19 is not real or not serious, and therefore that vaccinations are unnecessary. • The safety or efficacy of treatments or preventative measures that are not approved by health authorities, or that are approved by health authorities but not safe to administer from home 	<p>Comprehensive. Twitter's policy clearly addresses claims of both efficacy and necessity.</p>
Vaccine Development and Government Distribution	<ul style="list-style-type: none"> • How vaccines are developed, tested, and approved by official health agencies as well as information about government recommendations. 	<p>Non-Comprehensive. Twitter's policy clearly addresses claims related to vaccine development, however it's policy does not fully address government distributions. While government recommendations could include false government mandates, it is not clear if the policy will address other claims about vaccine distribution.</p>
Conspiracy	<ul style="list-style-type: none"> • The pandemic or COVID-19 vaccines that invoke a deliberate conspiracy by malicious and/or powerful forces. • Vaccines and vaccination programs which suggest that COVID-19 vaccinations are part of a deliberate or intentional attempt to cause harm or control populations. 	<p>Comprehensive. Twitter's policy states it will address known conspiracies related to the COVID-19 vaccine and provides an example of the types of conspiracy theories the platform will address.</p>
Platform intervention (content-based level)	<p>Twitter provides some clarity in its blog post on which actions it will take to vaccine-related misinformation:</p> <ul style="list-style-type: none"> • False claims that suggest immunizations and vaccines are used to intentionally cause harm to or control populations, including statements about vaccines that invoke a deliberate conspiracy; 	



	<ul style="list-style-type: none"> • False claims which have been widely debunked about the adverse impacts or effects of receiving vaccinations; or • False claims that COVID-19 is not real or not serious, and therefore that vaccinations are unnecessary <p>For unsubstantiated rumors, disputed claims, or incomplete, or out-of-context information Twitter will place a label or a warning.</p>
Twitter's Ad Policy	<p>Ads allowed on Twitter:</p> <ul style="list-style-type: none"> • The promotion of COVID-19 vaccines, treatments and test kits from approved Healthcare advertisers. Contact Twitter if you are interested in this option. <p>Ads not allowed on Twitter:</p> <ul style="list-style-type: none"> • distasteful references to COVID-19 (or variations) are prohibited • content may not be sensational or likely to incite panic • prices of products related to COVID-19 may not be inflated • the promotion of certain products related to COVID-19 may be prohibited (ie facemasks and hand sanitizers)

YouTube

In February 2019, BuzzFeed News [broke the news](#) that YouTube will demonetize channels that promote anti-vaccination content (we could not find any formal statement announcing this change from YouTube such as a blog post). YouTube also announced it would roll out a feature to a few users in India that shows fact-checks when people search for sensitive topics such as false health information. This fact-check would show up underneath the search bar. However, this product feature has not expanded to other countries since its introduction. In December 2020, YouTube expanded its [COVID-19 Medical Misinformation Policy](#) to include more specific examples of vaccine misinformation that YouTube wouldn't allow on its platform.

As per its ongoing policies, YouTube specifically mentions that it does not allow information that contradicts local health authorities of the World Health Organization's medical information. YouTube also has a clear Strikes system in which channels are terminated after three cases of content violating this policy. YouTube also provides a list of updates to its COVID-19 policies which can be found [here](#).



Category	Policy Language	Corresponding Rating
Safety	<ul style="list-style-type: none"> ● Claims that an approved COVID-19 vaccine will cause death, infertility, or contraction of other infectious diseases ● Claims that an approved COVID-19 vaccine will contain substances that are not on the vaccine ingredient list, such as fetal tissue ● Claims that any vaccine causes contraction of COVID-19 ● Claims that an approved COVID-19 vaccine will alter a person's genetic makeup ● Claims that the COVID-19 vaccine will kill people who receive it 	Comprehensive. YouTube's policy addresses a breadth of claims related to vaccine safety including side effects and vaccine ingredients.
Efficacy and Necessity	<ul style="list-style-type: none"> ● Claims that the COVID-19 vaccine causes contraction of other infectious diseases or makes people more vulnerable to contraction of other infectious diseases ● Claims that any vaccine is a guaranteed prevention method for COVID-19 	Non-Comprehensive. YouTube's policies do not address the breadth of claims related to the efficacy or necessity of the COVID-19 vaccine.
Vaccine Development and Government Distribution	<ul style="list-style-type: none"> ● Claims that a specific population will be required (by any entity except for a government) to take part in vaccine trials or receive the vaccine first 	Non-Comprehensive. YouTube's policies do not address claims about the development of the vaccine besides specific individual involvement. YouTube's policies are also unclear if they will apply to false claims about government mandates.
Conspiracy	<ul style="list-style-type: none"> ● Claims that an approved COVID-19 vaccine will contain substances or devices meant to track or identify those who've received it ● Claims that the COVID-19 vaccine will be used as a means of population reduction 	Comprehensive. YouTube's policies address known conspiracy theories related to vaccines.



	<ul style="list-style-type: none"> • Claims that the COVID-19 vaccine will contain a microchip or tracking device 	
Platform intervention (content-based level)	<p>Strikes system: 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 get a warning with no penalty to your channel. If it's not, we'll issue a strike against your channel. If you get 3 strikes, your channel will be terminated.</p>	

Pinterest

Pinterest was one of the earliest platforms to combat anti-vaccination content on its platform; anti-vaccination advice and other health misinformation have been explicitly against community guidelines [since 2017](#). In 2018, Pinterest took a more aggressive approach to health misinformation by not showing any search results for vaccines at all— a step Pinterest was the only platform to take. Most of the content related to vaccination cautioned against it, as reported by the [Wall Street Journal](#). Now since mid-2019, the platform has produced [reliable results](#) from public health organizations such as the WHO, CDC, and others that provide vaccine safety information in various languages. Below is the platform's policies as it relates to ads, content, and platform intervention.

	Policy Language
Policy related to Ads	<ul style="list-style-type: none"> • Does not allow ads on pages when you search for vaccines.(Blog post,



	<ul style="list-style-type: none"> ● 8/28/19) ● It is unclear if Pinterest bans ads about vaccines
Categories of Vaccine Narratives (Safety, Efficacy and Necessity, Vaccine Development and Government Distribution, and Conspiracy)	<ul style="list-style-type: none"> ● 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 (community guidelines)
Platform intervention (content-based level)	<ul style="list-style-type: none"> ● Pinterest will remove or limit distribution of misleading content ● when you search for vaccination-related /health-related terms, such as measles, vaccine safety, Pinterest will only show content from leading public health institutions such as the WHO, CDC, AAP, VSN. (Blog post, 8/28/19)

TikTok

As a relatively new platform, TikTok hasn't gone through earlier policy changes related to vaccines as other platforms tested here. That said, though not explicitly stated in TikTok's Community Guidelines, according to its [safety center](#), TikTok prohibits content that's false or misleading about the COVID-19 vaccine and anti-vaccination disinformation more broadly.

	Policy Language
Policy related to Ads	TikTok does 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.”
Categories of Vaccine Narratives (Safety, Efficacy and Necessity, Vaccine Development and Government Distribution, and Conspiracy)	Do not post, upload, stream, or share: <ul style="list-style-type: none"> • Medical misinformation that can cause harm to an individual's physical health (community guidelines)
Platform intervention (content-based level)	Distribution <ul style="list-style-type: none"> • TikTok will limit distribution into anyone’s For You feed of content that has an inconclusive fact-check Friction <ul style="list-style-type: none"> • TikTok redirects searches associated with vaccine or COVID-19 disinformation to its Community Guidelines and do not autocomplete anti-vaccine hashtags in search

Nextdoor

Nextdoor’s misinformation policy does not bind the platform to making any promises about removing COVID-19 vaccination-related mis- and disinformation. Its policy [states](#): Nextdoor reserves the right to take action (including through removing posts and/or suspending or deactivating member accounts) where we deem appropriate to prohibit the organizing of activities or prevent the spread of false claims, and conspiracy theories about COVID-19, including its causes, cures, and prevention methods that might be misleading or harmful to our neighbors and communities.



Reddit

Reddit's [only post](#) from its admin regarding COVID misinformation was from April 2020 and cannot include guidelines about vaccine information. While Reddit offers several suggested links from authoritative voices such as the CDC, the UK government, and Canadian government, Reddit leaves content moderation to the moderators of specific subreddits. As a result, subreddits with less content moderation, such as “r/china_flu”, are becoming more extreme than those with stricter content moderation, such as the authoritative “r/coronavirus”.¹

Google Search

In 2014, Google recognized the impact search results have in leading people to lower quality sources of information and as a result launched a policy called, [Your Money Your Life](#) to elevate authoritative sources in health, finance, and news. While Google Search does not have any explicit policies about their moderation of COVID-19 vaccine results, recently launched features, such as red alert banners, show that they are attempting to promote authoritative voices. When users search terms related to COVID-19 and/or vaccines, Google Search prompts an information panel with authoritative information on the disease and common vaccine questions. Further, in a [December 2020 blog post](#), Google announced plans to roll out an information panel specifically for the COVID-19 vaccine which will provide a list of authorized vaccines and information on them. This is currently available in the United Kingdom and will spread to other countries as they authorize vaccines.

Bing Search

In April 2020, Microsoft announced their steps to promote trusted information on COVID-19 on their products, including Bing Search. This includes presenting reliable health information near the top of search results and piloting algorithmic responses. While Bing has not announced updates to these steps specific to vaccines, a search on Bing about the COVID-19 vaccine yields an information panel with news and answers to frequently asked questions from CDC, WHO, and ECDC.

¹ <https://arxiv.org/pdf/2006.04816.pdf>