June 22, 2022

Meta's Policy Advisory Opinion Request: Removal of COVID-19 Misinformation

Table of Contents

1. Introduction

Meta requests a Policy Advisory Opinion from the Oversight Board on whether we should continue removing COVID-19 misinformation under our harmful health misinformation policy or if, instead, we should apply demotions or labels to this misinformation either directly, or through our third-party fact checking program.

As noted in our Community Standards, Meta's commitment to expression is paramount. We understand, however, that the internet creates opportunities for abuse, including through the spreading of misinformation. Recognizing that it is impossible to precisely define what constitutes misinformation across the whole range of online statements, we generally focus on contextualizing potentially false claims and reducing their reach. Through our third-party fact checking program, we limit the spread of misinformation and direct users to information and context from independent fact-checkers that assists users in deciding what to read, trust, and share. This approach allows us to preserve our users' ability to express themselves while reducing the risk that misinformation proliferates. We also know, however, that a stricter line is appropriate for the most dangerous kind of misinformation, specifically that which trusted experts tell us is likely to directly contribute to a risk of imminent physical harm, which we remove.

Before the COVID-19 pandemic, Meta only removed misinformation when local partners with relevant expertise told us a particular piece of content (e.g., a specific post on Facebook) could contribute to a risk of imminent physical harm. We never applied this policy to entire categories of false claims on a worldwide scale. However, in January 2020, based on the rapidly unfolding COVID-19 pandemic, we took the extraordinary step of removing entire categories of misinformation about the pandemic from our platforms (e.g., the claims that "'X' cures COVID"). Meta took this approach because outside health experts told us that misinformation about COVID-19, such as false claims about cures, masking, social distancing, and the transmissibility of the virus, could contribute to the risk of imminent physical harm, including by contributing to the risk of individuals contracting or spreading the disease, or refusing an eventual vaccine. Today, Meta removes 80 distinct false claims regarding COVID-19 and its associated vaccines from our platforms. We list each of these claims publicly in our COVID-19 Help Center.

Meta has maintained this approach for the past two and a half years despite many changes to the COVID-19 landscape. This is particularly true of the amount of accurate information about COVID-19 that is available to the public. When Meta first decided to remove COVID-19 misinformation from its platforms, the information available about the virus was sparse and sometimes contradictory. While misinformation about the virus still exists today, public health guidance about the

prevention and treatment of COVID-19 is readily available to people looking for it, including in our COVID-19 Information Center. Other circumstances have changed, too. Death rates from COVID-19 are decreasing and more than 60% of people in the world have received at least one dose of a COVID-19 vaccine.[1]

At the same time, Meta acknowledges that the course of the pandemic has been, and will continue to be, different across the globe. While vaccines, medical treatment, and authoritative guidance are increasingly available in high-income countries, experts predict that access will lag for people in low-income countries with less developed healthcare systems. It is important that any policy that Meta implements be appropriate for both sides of this divide, while being consistent and workable globally.

With this in mind, Meta asks the Oversight Board whether continuing to remove COVID-19 misinformation is consistent with our values of promoting users' voice and safety, along with recognized international human rights principles, or whether we should begin addressing this misinformation through less constrictive means, particularly through labels, demotions, or our third-party fact checking program.

Meta understands the difficult nature of this question and thanks the Board for its thoughtful consideration of this topic. We welcome further engagement with the Board on this issue. The Board's guidance will be useful not only in Meta's continued efforts to address the COVID-19 pandemic but in future health emergencies as well.

2. Meta's Approach to Misinformation

As noted in Meta's Misinformation Community Standards section, we remove misinformation where it is likely to directly contribute to the risk of imminent physical harm. In determining whether misinformation meets this standard, we partner with independent experts who possess knowledge and expertise to determine the truth of the content and assess whether it is likely to directly contribute to the risk of imminent physical harm. This includes, for instance, partnering with health organizations during the global COVID-19 pandemic.

We set this high bar for removing misinformation because misinformation, by its nature, is different from other types of speech addressed in the Community Standards. First, there is often no way to articulate a comprehensive list of what constitutes misinformation at any given time. In many cases, facts regarding a particular topic change day-to-day. Because of that, it would not be possible for

[1] Our World in Data, "Coronavirus (COVID-19) Deaths" (Updated June 1, 2022) (link here); Josh Holder, "Tracking Coronavirus Vaccinations Around the World," The New York Times (Updated May 30, 2022) (link here).

Meta to provide adequate notice to users regarding what is, and is not, prohibited on our platforms. Second, people often use misinformation in harmless ways, such as to exaggerate a point or in humor or satire. They also may share their experience through stories that contain inaccuracies. In some cases, people share deeply-held personal opinions that others consider false or share information that they believe to be true but others consider incomplete or misleading. Third, there is no expert consensus that removing misinformation is always the right approach. In fact, some experts believe that it is preferable instead to allow more fulsome discussion that lets people digest and consider the relevant information themselves.[2]

Therefore, for misinformation not meeting our high standard for removal, we focus instead on slowing its spread and directing users to authoritative information. As part of that effort, we partner with third-party fact checking organizations to review and rate the accuracy of the most viral content on our platforms. These fact-checkers are independent from Meta and are certified through the non-partisan International Fact-Checking Network (IFCN). We currently partner with more than 80 organizations fact-checking in more than 60 languages globally.

Under this approach, Meta's employs four exclusive interventions in response to misinformation on its platforms:

- *Remove*: Our general approach to misinformation is to address viral false claims by providing users with additional context to better enable them to decide what to read, trust, and share. This approach comports with our paramount value of promoting expression. For that reason, we maintain a high bar for removing misinformation (i.e., the misinformation must directly contribute to a risk of imminent physical harm as assessed by a relevant external expert).

- *Temporary Emergency Reduction Measures*: Meta scales the demotion of important and repeatedly fact-checked claims when misinformation about a particular crisis spikes on our platforms and our third-party fact-checkers cannot keep up with rating those claims. We may also temporarily reduce other claims that do not rise to the level of what we remove, if warranted during a crisis. These measures are not intended to be steady-state measures and are only meant to be implemented during the emergency or crisis.

[2] "The online information environment," The Royal Society, Jan. 19, 2022 ("[C]ensoring or removing inaccurate, misleading and false content, whether it's shared unwittingly or deliberately, is not a silver bullet and may undermine the scientific process and public trust.") (link here); Conor Friedersdorf, "*Tolerating COVID Misinformation is Better than the Alternative," The Atlantic, Mar. 7, 2022 (link here).*

- ➢ *Third-Party Fact-Checking*: When an independent fact-checker rates a piece of content, users will see a label showing the fact-checker's rating (i.e., False, Altered, Partly False, or Missing Context). The label will also direct users to the fact-checker's article on the topic. We apply our strongest warning labels for content rated False or Altered, and lighter labels for Partly False and Missing Context.[3] We also notify people before they try to share this content or if they shared it in the past. Content rated "False," "Altered," or "Partly False" is reduced (demoted) in users' Feeds, while content rated "Missing Context" is generally not demoted (Meta, however, currently demotes COVID-19 related content that is rated "Missing Context."). We provide, in Appendix A of this Policy Advisory Opinion request, examples of how fact-checked content appears to users.

- ➢ *Labels*: In limited circumstances, Meta may provide authoritative information on a topic by placing labels on content. For example, Meta employed labels during the U.S. 2020 election, and we currently employ labels on non-violating COVID-19 related content that directs users to our COVID-19 Information Center. These labels do not signal a judgment on whether the post is true or false. Rather, they are another way we connect people with reliable information about the underlying topic. We provide, in Appendix B of this Policy Advisory Opinion request, examples of how labels appear to users.

3. Meta's Approach to Harmful Health Misinformation

In January 2020, Meta convened a team to monitor the rapidly evolving COVID-19 outbreak to identify any related content risks on our platforms. Because the outbreak was still in its very early stages, little was known about the virus. Meta saw new types of misinformation on our platforms at the time, with a significant amount of content focusing on the origins of the virus. This combination of a rapidly evolving health crisis and a lack of information meant that people faced obstacles in making informed decisions on how to protect themselves from the virus.

[3] Content rated "False" or "Altered" receives a warning screen that blocks our users from seeing the content in their Feed unless they click a button on the post, which also directs our users to the fact-checker's article debunking the content. Content rated "Partly False" or "Missing Context" receives a lighter label that does not obstruct the user from seeing the post in their Feed, and also provides a link to the fact-checker's article. Content rated "False," "Altered," or "Partly False" is reduced (demoted) in users' Feeds, while content rated "Missing Context" is generally not demoted (Meta currently demotes COVID-19 related content that is rated "Missing Context.").

Through conversations with health experts at the time, Meta understood the following about the disease: (1) it was deadly, with an estimated fatality rate around 2%[4]; (2) it was highly contagious, particularly because it was an airborne disease; and (3) there was no treatment or cure. Public health experts also told Meta that certain misinformation about COVID-19, such as false claims about masking or social distancing, could contribute to a risk of harm to people and exacerbate the crisis by increasing the risk of spreading, contracting, or mistreating the disease.

Before COVID-19, Meta's mechanism for removing harmful health misinformation was through our misinformation and harm policy.[5] As the widespread impact of COVID-19 became apparent, Meta adopted a policy on January 28, 2020, to remove harmful health misinformation from our platforms, including COVID-19 misinformation. Under this policy, Meta would remove harmful health misinformation if the following criteria are met: (1) there is a public health emergency; (2) leading global health organizations or local health authorities tell us a particular claim is false; and (3) those organizations or authorities tell us the claim can directly contribute to the risk of imminent physical harm.

Under this policy, Meta required the presence of a public health emergency because, as advised by the World Health Organization (WHO), there is a high risk of irreversible physical harm to individuals when the risk of exposure, rate of transmission, association between exposure and risk, and morbidity and mortality rates are unusually high. In addition, public health emergencies often begin with information vacuums where public knowledge and discourse about the emergency is not yet developed enough to allow accurate information to counteract misinformation.

[4] The case fatality rate as of January 24, 2020, was estimated at around 2% (80 deaths out of approximately 3,000 cases). However, health experts told us these statistics likely underestimated the severity of the disease due to a lack of reliable reporting coming out of China.

[5] This policy is now contained in the Misinformation Community Standards section under Part I, Physical Harm and Violence. Before the COVID-19 pandemic, Meta removed (rather than reduced) health content under our misinformation and harm policy for a limited time on two occasions. The first was in June 2019 in Pakistan following a string of deadly attacks targeting healthcare workers who were administering polio vaccinations that were linked to misinformation about the polio vaccine. Meta made the decision to remove this misinformation after consulting with public health authorities, including the country heads of UNICEF, the World Health Organization ("WHO"), and the National Health Ministry of Pakistan. The second occasion was in December 2019, when Meta began removing nine vaccine "widely debunked hoaxes" in Samoa during a severe and deadly measles outbreak. Meta made this decision after consulting the WHO and UNICEF, both of which noted a correlation between vaccine misinformation and the outbreak. In both instances, information from public health authorities enabled Meta to make a link between health misinformation and a direct contribution to the risk of imminent physical harm.

Meta determined it would rely on the following to assess whether a public health emergency exists: (1) whether the WHO declared a public health emergency; (2) whether the WHO designated a disease as communicable, deadly, or high risk; or (3) in the event a WHO risk assessment is unavailable, we would defer to local health authorities' designation of a public health emergency for a given country.

To assess falsity under this policy, Meta would rely on reports and official statements from credible health organizations, such as the WHO and certain governmental health organizations, such as the United States Centers for Disease Control and Prevention (CDC). As with its overall approach to removing misinformation from its platforms, Meta would not make its own truth or falsity assessments regarding harmful health misinformation.

In assessing whether a false claim could directly contribute to a risk of imminent physical harm, Meta would rely on public health experts' assessments. Those health experts advised that we should focus on the following types of claims, each of which could exacerbate an ongoing public health crisis by increasing the risk of spreading, contracting, or mistreating the disease:

- *False cures*: Health experts advised that false claims that a cure exists to a particular health emergency could directly contribute to the risk of imminent physical harm because, if a cure is not actually in existence, it could lead someone to engage in riskier behavior or it could lead them to seek ineffective treatment and thereby risk their own safety and/or additionally expose others.

- *False information designed to discourage treatment*: Experts advised that this kind of misinformation could lead someone to not seek effective treatment, thereby risking their own safety and additionally risking exposure for others. This includes false claims about hospital practices or safety.

- *False prevention information*: Experts advised that this type of misinformation could lead someone to not take effective precautions to avoid exposure or infection. This includes claims such as masks not being effective to stop the spread of airborne diseases.

- *False information about availability of or access to health resources*: Experts advised that this type of misinformation could prevent someone from seeking effective treatment, thereby risking their own safety or the safety of others. This includes false claims about

the dates, locations, times, requirements, or methods of administration of medical countermeasures.

- ➢ *False information about the location or severity of a disease outbreak*: Experts advised that this type of misinformation could lead someone to increase their risk of exposure to the disease. This includes false claims that an outbreak has not reached a place where there are already confirmed infections.

Two days after Meta adopted its harmful health misinformation policy, the WHO re-convened an emergency committee meeting on January 30, 2020, and designated COVID-19 a global Public Health Emergency of International Concern. That same day, we published our first Newsroom post on the company's response to COVID-19, stating that we would start removing:

> [C]ontent with false claims or conspiracy theories that have been flagged by leading global health organizations and local health authorities that could cause harm to people who believe them. We are doing this as an extension of our existing policies to remove content that could cause physical harm. We're focusing on claims that are designed to discourage treatment or taking appropriate precautions. This includes claims related to false cures or prevention methods — like drinking bleach cures the coronavirus — or claims that create confusion about health resources that are available. We will also block or restrict hashtags used to spread misinformation on Instagram, and are conducting proactive sweeps to find and remove as much of this content as we can.

Since determining that COVID-19 misinformation is eligible for removal under the harmful health misinformation policy, Meta has continued its work with public health authorities to incorporate additional claims that those authorities advise are false and may directly contribute to the risk of imminent physical harm. Information about how Meta treats COVID-19 misinformation is published in the Misinformation Section of the Community Standards under the subsection titled, "Misinformation about health during public emergencies." That subsection provides:

> We remove misinformation during public health emergencies when public health authorities conclude that the information is false and likely to directly contribute to the risk of imminent physical harm, including by contributing to the risk of individuals getting or spreading a harmful disease or refusing an associated vaccine. We identify public health emergencies in partnership with global and local health

> authorities. This currently includes false claims related to COVID-19 that are verified by expert health authorities, about the existence or severity of the virus, how to cure or prevent it, how the virus is transmitted or who is immune, and false claims which discourage good health practices related to COVID-19 (such as getting tested, social distancing, wearing a face mask, and getting a vaccine for COVID-19).

We also provide a link in this paragraph to our COVID-19 Help Center, which includes a complete list of the COVID-19 and vaccine misinformation we remove from our platforms. This list includes, but is not limited to, false claims about: (1) the existence and severity of COVID-19[6]; (2) COVID-19 transmission and immunity[7]; (3) guaranteed cures or prevention methods for COVID-19[8]; (4) discouraging good health practices;[9] and (5) access to essential health services.[10] We also provide examples of the types of claims prohibited under our COVID-19 removal policies. We also note in our COVID-19 Help that, as the situation surrounding COVID-19 evolves, we will "continue to look at

---

[6] This includes: (1) claims that deny the existence of the COVID-19 disease or pandemic; (2) claims that downplay the severity of COVID-19; and (3) claims about the cause of COVID-19 that are linked to 5G communication technologies.

[7] This includes: (1) claims that any group is immune or cannot die from COVID-19 or that a specific activity or treatment results in immunity; (2) claims that COVID-19 cannot be transmitted through the air, in certain climates, weather conditions, or locations; and (3) claims that COVID-19 can be transmitted from anything other than human-to-human transmission.

[8] This includes claims that, for the average person, something can guarantee prevention from getting COVID-19 or can guarantee recovery from COVID-19 before such a cure or prevention has been approved.

[9] This includes: (1) claims about wearing a facemask, such as claims that wearing a face mask properly does not help prevent the spread of COVID-19; (2) claims that social/physical distancing does not help prevent the spread of COVID-19; (3) claims that can discourage someone from getting a government approved COVID-19 test, such as claims that COVID-19 can be successfully tested without an approved test; (4) claims about COVID-19 vaccines that contribute to vaccine rejection, such as claims that there are no U.S. Food and Drug Administration approved COVID-19 vaccines or that something other than a COVID-19 vaccine can vaccinate you against COVID-19; (5) claims about the safety or serious side effects of COVID-19 vaccines, such as that COVID-19 vaccines kill or seriously harm people or that building immunity by getting COVID-19 is safer than getting the vaccine; (6) claims about the efficacy of COVID-19 vaccines, such as claims that the vaccines are not effective in preventing severe illness or death from COVID-19; (7) claims about how the COVID-19 vaccine was developed or its ingredients, such as claims that the vaccine contains toxic or harmful ingredients or microchips, or that the COVID-19 vaccines are untested; and (8) claims involving conspiracy theories about a COVID-19 vaccine or vaccination program, such as that COVID-19 vaccines are intended for population control.

[10] This includes: (1) claims that hospitals or a specific hospital is closed and will not permit infected people; (2) claims that only certain people are allowed to receive medical care for COVID-19; and (3) claims that hospitals kill patients in order to inflate the number of COVID-19 deaths, to get more money, or in order to sell people's organs.

content on the platform, assess speech trends, and engage with experts like the World Health Organization (WHO), government health authorities, and stakeholders from across the spectrum of people who use our service, and we will provide additional policy guidance when appropriate to keep the members of our community safe during this crisis."

4. The Oversight Board's Previous Decisions Involving Our COVID-19 Misinformation Policies

As Meta requests the Board's advice on our approach to removing COVID-19 misinformation, we acknowledge the Board has addressed Meta's COVID-19 misinformation policies in two prior decisions. In the first decision, from January 2021, the Board overturned Meta's removal of a post from France claiming that the anti-malarial drug, hydroxychloroquine, when combined with the antibiotic, azithromycin, is a "cure" for COVID-19. Meta removed that post for violating our COVID-19 misinformation policies, specifically Meta's prohibition on content claiming a cure for COVID-19 exists. As Meta explained at the time, leading experts told Meta that content claiming there is a guaranteed cure or treatment for COVID-19 could contribute to imminent physical harm by leading people to ignore preventive health guidance or attempt to self-medicate.

In its decision, however, the Board noted that "[s]erious questions remain about how the post would result in *imminent* harm." (emphasis in original). According to the Board, "[w]hile some studies indicate the combination of anti-malarial and antibiotic medicines that are alleged to constitute a cure may be harmful, experts the Board consulted noted that they are not available without a prescription in France." The Board stated further that "the alleged cure has not been approved by the French authorities and thus it is unclear why those reading the post would be inclined to disregard health precautions for a cure they cannot access." The Board concluded that "Facebook failed to provide any contextual factors to support a finding that this particular post would meet its own *imminent harm* standard" and, as such, Facebook "did not act in compliance with its Community Standards."

Facebook disagreed with the Board's decision, noting that "[o]ur global expert stakeholder consultations have made it clear that, in the context of a health emergency, the harm from certain types of health misinformation does lead to imminent physical harm," and "[t]hat is why we remove this content from our platform." We further explained that the WHO and other public health authorities advised us that "if people think there is a cure for COVID-19 they are less likely to follow safe health practices, like social distancing or mask-wearing." For that reason, while we restored the post in question pursuant to our commitments to follow the

Board's decision, we did not adopt the Board's recommended approach to COVID-19 misinformation at that time.

In a later case from August 2021, the Board upheld Meta's decision not to remove a Brazilian state-level medical council's post claiming that COVID-19 lockdowns are ineffective, against fundamental rights in the Brazilian Constitution, and condemned by the WHO. While the Board recognized, particularly in the context of the COVID-19 crisis in Brazil, that "the spread of COVID-19 misinformation in the country can endanger people's trust in information about appropriate measures to counter the pandemic, which could increase the risk of users adopting risky behaviors," the post "did not meet the threshold of imminent harm, because it discusses a measure that is not suggested unconditionally by public health authorities and emphasizes the importance of other measures to counter the spread of COVID-19 - including social distancing." The Board recommended that this kind of content, instead, be sent to third-party fact-checkers.

5. Current Trends Regarding COVID-19

While Meta ultimately disagreed with the Board's first COVID-19 decision from January 2021, Meta acknowledges today that the landscape surrounding COVID-19 has changed somewhat from its original decision to remove COVID-19 misinformation over two years ago. Those changes may fundamentally impact whether or not Meta should continue removing this misinformation going forward, or allow it to be addressed through other means. In particular, Meta notes the following to the Board:

a. The COVID-19 information ecosystem has changed since the creation of our COVID-19 misinformation policy.

At the beginning of the pandemic, the lack of authoritative guidance created an information vacuum that encouraged the spread of rumors, speculation, and misinformation. As Tanya Lewis writes in Scientific American, "[r]eporting on the pandemic was like building a plane while flying it—at warp speed in a hurricane. The underlying science was evolving daily, so there was no expert consensus or body of established research to draw on. And there were plenty of people willing to exploit this information vacuum, creating a secondary epidemic of misinformation."[11] "When complex emergencies arise," another article notes, "public officials are cautious about making premature pronouncements, instead carefully crafting statements to ensure accuracy and avoid the pitfalls of misinterpretation and exaggeration.

[11] Tanya Lewis, "How the Pandemic Remade Science Journalism," Scientific American, Volume 326, Issue 3, pp. 38-39 (Mar. 2022) (link here).

Somewhat paradoxically, this careful approach may also contribute to the formation of an information vacuum that rumours and falsehoods are all too ready to fill."[12]

While COVID-19 persists today, much has changed about people's access to information about the virus and the ability of public health authorities to effectively inform and shape the behavior of those at risk. For instance, in January 2020, the WHO's website contained only a single webpage devoted to COVID-19.[13] Reflecting the concerns at the time, much of the webpage's information was devoted to how the virus was spread in the "wet markets" of Asia. The WHO webpage did not yet address false claims about the virus. In 2022, however, the WHO's site now has over 1,800 webpages devoted to COVID-19, including a "mythbusters" section that addresses a number of false claims.

Similarly, in January 2020, the CDC had a handful of webpages with information about COVID-19.[14] These pages reflected the uncertainty of public health officials of the time, with guidance denoted as "interim" and statements such as "[a]t this time, it's unclear how easily or sustainably this virus is spreading between people."[15] Today, the CDC has over 900 webpages devoted to COVID-19 full of definitive statements and guidance.[16]

While misinformation about COVID-19 continues to exist, data-driven, factually reported information about the pandemic has been published at an astounding rate. Between January and December 2020, between 100,000 to 200,000 scientific papers were published about COVID-19, representing four percent of total global scientific output.[17] These papers cover everything from epidemiological surveys to the mental

[12] S. Harris Ali & Fuyuki Kurasawa, "#COVID19: Social media both a blessing and a curse during coronavirus pandemic," The Conversation (Mar. 22, 2022) (link here),
[13] "Novel Coronavirus 2019-n-CoV Advice for the Public," Wayback Machine (World Health Organization) (Jan. 28, 2020) (link here).
[14] "Information for Healthcare Practitioners," Internet Archive (Centers for Disease Control and Prevention) (Jan. 30, 2020 (link here)l.
[15] "Novel Coronavirus 2019-n-CoV Advice for the Public," Internet Archive (World Health Organization) (Jan. 28, 2020) (link here); "Information for Healthcare Practitioners," Wayback Machine (Centers for Disease Control and Prevention) (Jan. 30, 2020) (link here).
[16] "Vaccines for COVID-19," Centers for Disease Control and Prevention (link here) ("COVID-19 vaccines are safe, effective, and free. Get answers to frequently asked questions and bust myths about vaccines.").
[17] Holly Else, "How A Torrent of COVID Science Changed Research Publishing–in Seven Charts," Nature (Updated Dec. 17, 2020) (link here).

health implications of the pandemic.[18] Between January 2020 and today, more than 28 million news articles have been published about COVID-19.[19]

b. Due to vaccines and evolution of disease variants, COVID-19 is less deadly than it was in the spring of 2020.

Perhaps the biggest change in the global outlook for the future of the pandemic is the availability of vaccines that prevent and reduce the severity of COVID-19 symptoms. As has been widely reported, the development and distribution of effective vaccines continues to be a massive project undertaken, in most parts of the world, by a partnership of governments, non-governmental organizations, and private entities. Meta itself has committed resources to increasing the awareness and availability of vaccines.[20] To date, more than 5 billion people worldwide have received a dose of a COVID-19 vaccine, equal to about 67 percent of the world population. However, only 16% of people in low-income countries have received at least one dose.[21] Therapeutic treatments for COVID-19 are also evolving rapidly, with a number of effective treatments becoming increasingly widely available, including Paxlovid, Remdesivir, and monoclonal antibodies.[22]

Additionally, according to the CDC, COVID-19 infection caused by the Omicron variant of the virus, which is currently the most prevalent throughout the world, "generally causes less severe disease than infection with prior variants."[23] The Coronavirus Resource Center at Johns Hopkins University notes that "[t]he data trends clearly demonstrate that Omicron is a much less deadly variant, which is

[18] Id.
[19] Reuben Ng, Ting Yu Joanne Chow, and Wensu Yang, "News Media Narratives of COVID-19 Across 20 Countries: Early Global Convergence and Later Regional Divergence," PLOS One, (Sept. 1, 2021) (link here).
[20] Furthermore, since April 2020, Meta has collaborated with Carnegie Mellon University and the University of Maryland on a global survey to gather insights about COVID-19 symptoms, testing, vaccination rates and more. We received over 70 million total responses, and more than 170,000 responses daily across more than 200 countries and territories. For people in the United States on Facebook, vaccine hesitancy had declined by about 50% during the first months of 2021, and today, acceptance is high - 84% of Facebook users in the United States have been or want to be vaccinated against COVID-19.
[21] Josh Holder, "Tracking Coronavirus Vaccinations Around the World," The New York Times (Updated May 30, 2022) (link here); Our World in Data, "Coronavirus (COVID-19) Vaccinations" (Updated June 1, 2022) (link here).
[22] Kathy Katella, "COVID-19 Treatments: What We Know So Far," Yale Medicine (May 19, 2022) (link here).
[23] "Omicron Variant: What You Need to Know," Centers for Disease Control and Prevention (Updated Mar. 29, 2022) (link here).

critical for downgrading COVID-19 to an endemic disease."[24] Experts note that the lessening of severe symptoms may be connected to increased vaccination rates.[25]

c. Public health authorities are actively evaluating whether COVID-19 has evolved to a less severe state.

With the rise in vaccinations, and the emergence of a less severe strain of COVID-19 (Omicron), some public health authorities are noting that certain regions of the world have begun transitioning to a less severe state of the pandemic, and are moving towards an endemic state. For instance, Dr. Anthony S. Fauci, the Director of the United States National Institute of Allergy and Infectious Diseases (NIAID), noted in April 2022 that, while the pandemic is not over in the United States, the country is "out of the full-blown explosive pandemic phase" and is "really in a transitional phase, from a deceleration of the numbers into hopefully a more controlled phase and endemicity."[26] Similarly, the European Union announced in April that it was moving out of the emergency phase of the pandemic to one focusing on vaccination, pandemic surveillance, and testing.[27] Moreover, in June, the Public Health Minister of Thailand announced his ministry's intention to work toward classifying COVID-19 as an endemic disease.[28]

Meanwhile, the long-term future state of COVID-19 is uncertain.[29] Some experts believe that people should expect to contract COVID-19 multiple times over their lives – perhaps as often as once every three years, which would be similar to flu viruses – and that the typical infection will get less dangerous over time. However, it is unclear what the effects of multiple reinfections could be, how immunity may change over time, and the effects of long COVID-19.

It is important to note that the course of the COVID-19 pandemic will likely continue to vary across the world. The most significant variation right now is between developed nations, which already have high vaccination rates, and less developed nations, which do not. While the WHO set a goal of vaccinating 70 percent of the

[24] "Comparing Cases, Deaths, and Hospitalizations Indicates Omicron Less Deadly," John Hopkins Coronavirus Research Center (Mar. 7, 2022) (link here).
[25] Id.
[26] Joel Achenbach and Bryan Pietsch, "U.S. no longer in 'full-blown' pandemic phase, Fauci says," Washington Post (Apr. 27, 2022) (link here); see also Naomi Thomas and Brenda Goodman, "The US is in 'Transition Phase' of Pandemic, Fauci Says," CNN, April 27, 2022 (link here).
[27] Monika Pronczuk, "The European Union says the emergency phase of the pandemic is over.," N.Y. Times, April 27, 2022 (link here).
[28] "Anutin: Time to 'Move On', Declare Covid-19 Endemic," Bangkok Post (June 8, 2022) (link here).
[29] Katherine J. Wu, "You Are Going to Get COVID Again…And Again…And Again," The Atlantic (May 27, 2022) (link here).

global population by the end of 2022, it has acknowledged that the world is unlikely to meet this goal without implementing new policies for distributing vaccines.[30] Eighty percent of people in high-income countries have received at least one dose of the vaccine, as opposed to only 13 percent of people in low-income countries.[31] Low-income countries are also more likely to have health care systems with less capacity, less robust economies, and lower trust in government guidance, all of which will add challenges to vaccinating people and treating those that contract COVID-19.[32] In the years ahead, the disease is likely to disproportionately affect people in those countries as interest wanes in richer countries. This has happened before with malaria, tuberculosis, and HIV/AIDS.[33]

6. Options for Addressing COVID-19 Misinformation on Our Platforms

In line with Meta's approach to misinformation, as described in Section 2 above, the following options are available to address COVID-19 misinformation on our platforms, depending on the Board's ultimate guidance:

- ➢ *Continue Removing Certain COVID-19 Misinformation*. This option is available to the Board if the Board determines that, applying Meta's policy for harmful health misinformation, certain COVID-19 misinformation that Meta removes still directly contributes to a risk of imminent physical harm. Under this option, we would eventually stop removing this misinformation when it no longer satisfied the requirements of our harmful health misinformation policy. Should the Board choose this option, Meta requests the Board's guidance regarding how Meta should determine when COVID-19 misinformation no longer meets that threshold.

- ➢ *Temporary Emergency Reduction Measures*. Under this option, Meta would cease removing COVID-19 misinformation and, instead, reduce the distribution of those claims. This measure would only be temporary. If the Board decides that Meta should adopt this approach for all COVID-19 misinformation claims we currently remove, Meta requests the Board's

[30] "COVID vaccines: Widening inequality and millions vulnerable," United Nations (Sept. 19, 2021) (link here).

[31] Nadia A. Sam-Agudu, Boghuma Kabisen Titanji, Fredros Okumu, and Madhukar Pai, "The Pandemic Is Following a Very Predictable and Depressing Pattern," The Atlantic (Mar. 4, 2022) (link here).

[32] Joseph Stiglitz, "The Pandemic Has Laid Bare Deep Divisions, But It's Not Too Late to Change Course," International Monetary Fund (Sept. 2020) (link here).

[33] Nadia A. Sam-Agudu, Boghuma Kabisen Titanji, Fredros Okumu, and Madhukar Pai, "The Pandemic Is Following a Very Predictable and Depressing Pattern," The Atlantic (Mar. 4, 2022) (link here).

guidance on when it should eventually cease using this measure in the future. We provide a discussion of how we demote certain non-violating vaccine related content during the pandemic in Appendix C of this PAO.

- ➢ *Third-Party Fact Checking*. Under this option, instead of removing COVID-19 misinformation, Meta would defer to independent third-party fact checkers to find and rate the falsity of those claims. Since the beginning of the pandemic, we have displayed warnings on more than 195 million pieces of COVID-related content on Facebook that our fact checking partners rated. That said, the number of fact checkers available to rate content will always be limited. If Meta were to implement this option, fact-checkers would not be able to look at all COVID-19 content on our platforms, and some of it would not be checked for accuracy, demoted, and labeled.[34]

- ➢ *Labels*. Under this approach, Meta would affix labels underneath content on its platform. The label would not obstruct users from seeing the content and would direct those users to authoritative information. We provide examples of how labels appear to users in Appendix B to this PAO request.[35] As with the Temporary Emergency Measure described above, this measure would only be a temporary and we would welcome the Board's guidance, if it selects this option, on what factors Meta should consider in deciding to cease using these labels.

Each of these enforcement options have advantages and disadvantages, particularly in terms of scalability, accuracy, and in terms of the amount of content affected.[36] For

[34] The fact-checkers' ratings would apply to shares and re-shares of identical content. Their ratings, however, would not apply to other pieces of content that are not identical but make the same claim.

[35] Given the sheer number of posts on this topic, it would not be feasible for humans to review each post and apply a label. We would, therefore, rely on artificial intelligence or content rules to identify posts related to COVID-19 and vaccines. These systems, however, often present issues with precision/recall given that COVID-19 is a constantly evolving topic and posts that are irrelevant to the particular label will often get labels attached to them. There are also major challenges with discerning differences between misleading, debunking, supportive, or neutral content and this approach applies the label to all of these categories of content. We can discuss these issues in greater depth with the Board during subsequent Q&A sessions, or through additional briefing, if the Board wishes to pursue this option.

[36] As noted in Footnote 1, Meta sometimes scales repeatedly fact-checked claims during critical events. This option is intended to supplement fact checkers' efforts to address misinformation claims that are going viral during a critical event. If important claims during a critical event have been repeatedly fact-checked and we find our partners cannot keep up with applying their ratings to new content making the same claim, Meta can search for content on its platform containing those claims and directly apply a misinformation label to, and demote, the content, allowing us to scale the impact of fact checks. This process, however, is extremely resource and labor intensive, and it is not possible to apply this process

technical reasons, we strongly recommend maintaining global policies regarding COVID-19, as opposed to country or region specific policies.[37] Meta welcomes the opportunity to discuss these options with the board in greater detail.[38]

7. Conclusion

Enforcing on COVID-19 misinformation on Meta's platforms is uniquely challenging given the incomplete and continuously evolving understanding of the virus and its treatment. However, Meta's goal remains the same: to connect people to accurate information and stop harmful misinformation from spreading. It is important to Meta, however, that our pursuit of this goal does not overly restrict our users' ability to express themselves on these important topics. We appreciate the Oversight Board's insight on whether certain COVID-19 misinformation still satisfies our standard for removing harmful health misinformation, or whether Meta should address that information through alternative enforcement options in the future.

to every COVID-19 misinformation claim published on our platforms. Therefore, this is not a viable option for addressing COVID-19 misinformation going forward. If the Board wishes to enquire further into this option, Meta can provide additional information during a future Q&A session on this PAO.

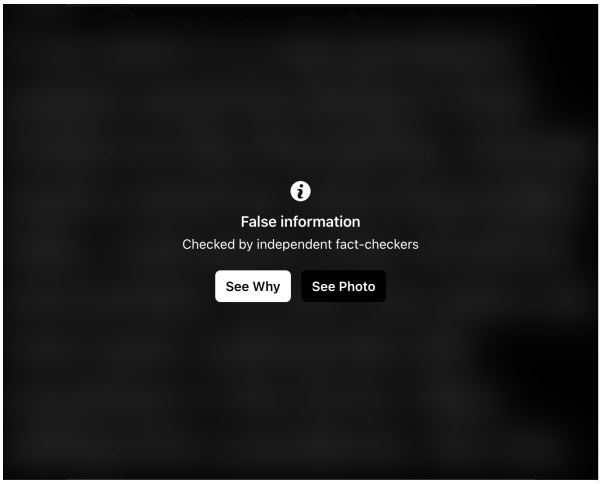
[37] Enforcing policies at the country level can lead to both over-enforcement when one set of market reviewers covers multiple countries, and under-enforcement because content can spread across countries and regions. Meta can provide greater explanation on the difficulties of operating a country or region specific policy during our subsequent Q&A sessions or in additional briefing if the Board so wishes.

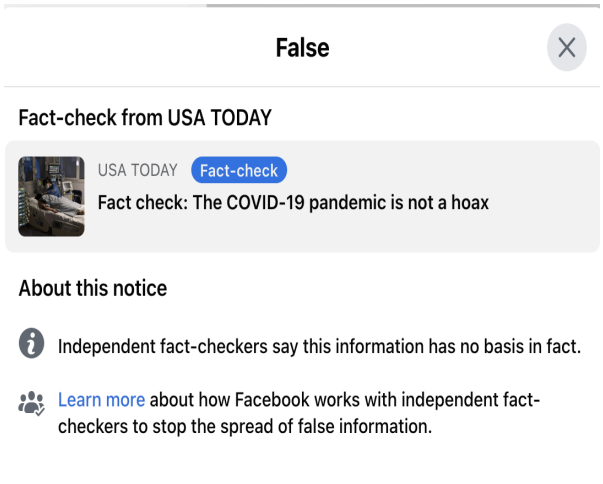
[38] We note that any change to how we enforce on COVID-19 misinformation will likely take our Product and Operations teams approximately 6 months to operationalize.

Appendix A

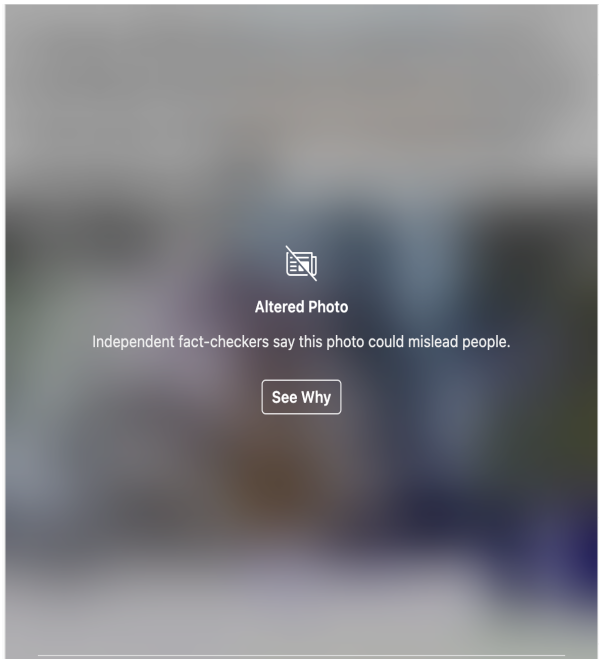
Examples of How Third-Party Fact Checks Appear to Users

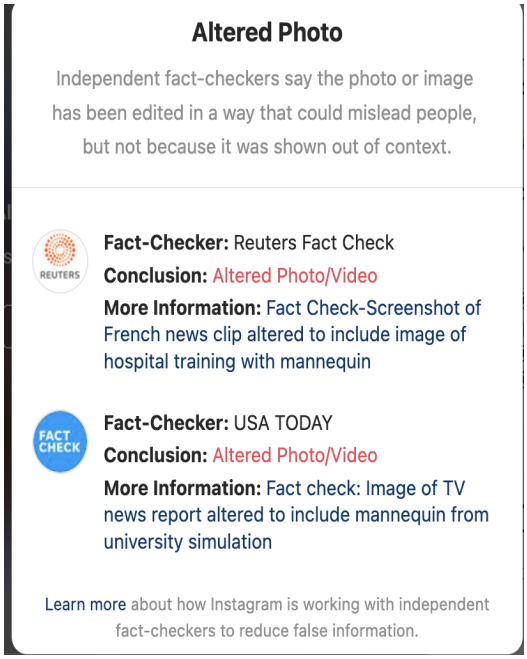
False



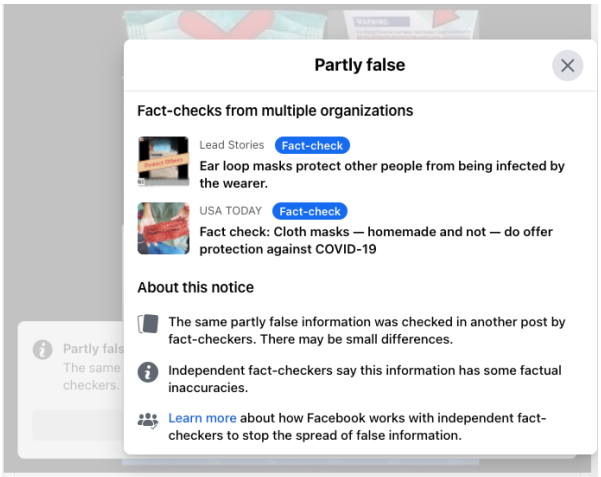


Altered

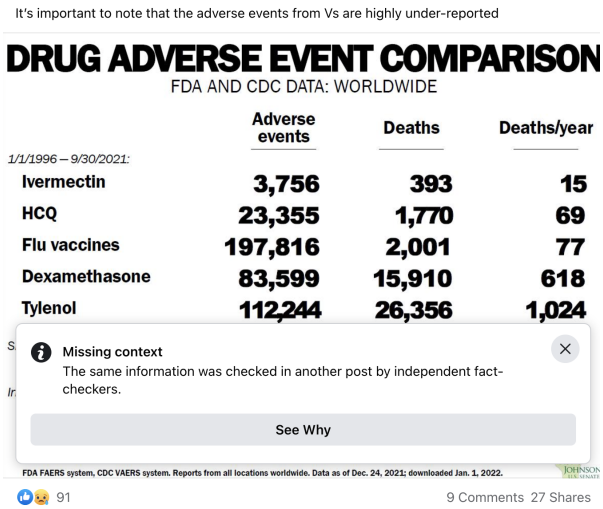


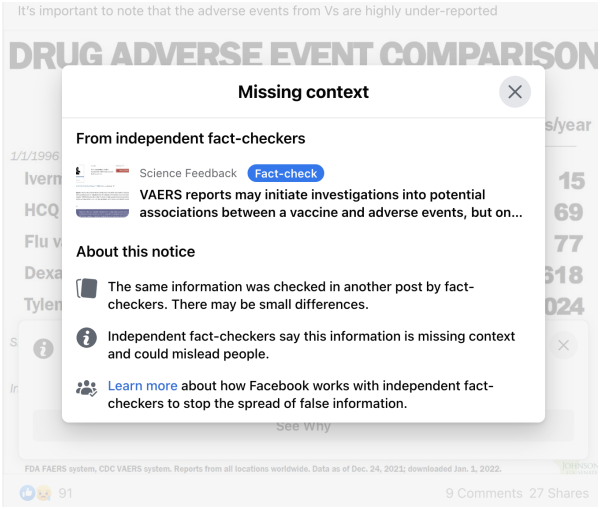


Partly False



Missing Context

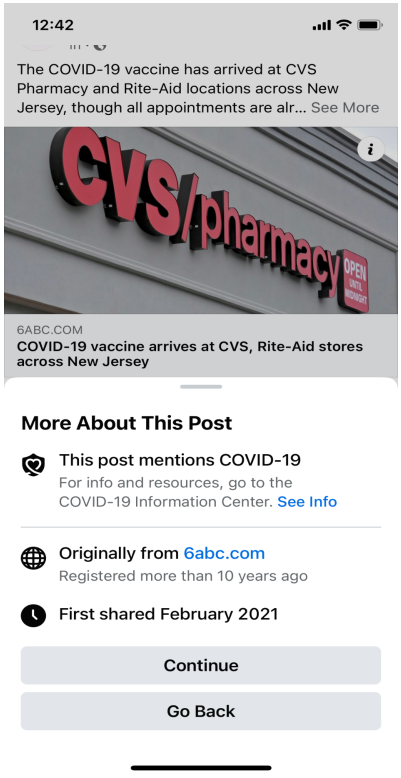
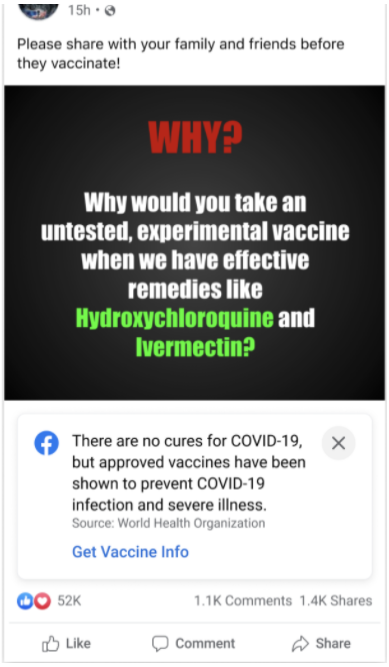




Appendix B

Examples of How Labels Appear to Users

As background, Meta generally employs two types of labels on content: (1) "neutral inform treatments" ("NITs") (e.g., "Visit the COVID-19 information center for vaccine resources"); and (2) "facts about 'X' inform treatments" ("FAXITs") that can be customized and applied to any piece of content on Facebook or Instagram (e.g., "COVID-19 vaccines go through many tests for safety and effectiveness before they are approved.").

Neutral Inform Treatments ("NITs")	Facts About 'X' Informed Treatments ("FAXITs")
	

Appendix C

Reducing the Distribution of Certain Non-Violating Vaccine Content that May Contribute to Vaccine Hesitancy

Potentially problematic content that does not meet the standard for removal under our policies is left on our platforms to preserve user voice and dialogue. However, Meta may reduce the distribution of such content by demoting it in users' feeds. Additionally, Meta may limit its eligibility to appear in our recommendations experiences, which include Pages you may like, "suggested for you" posts, Instagram Explore, people you may know, or groups you should join.

As explained in the COVID-19 Help Center, Content Distribution Guidelines, and Content Borderline to the Community Standards, Meta reduces the distribution of certain content about vaccines that public health experts have advised us could contribute to individuals being hesitant to get vaccinated against COVID-19. This content does not meet our misinformation removal policy because it may not make a false claim or it may not be verifiably false, as determined by third party fact-checkers or outside health experts. Instead, we reduce the distribution of this content and limit its recommendation systems' eligibility to be consistent with our overall approach to ranking problematic content and our efforts to maintain the quality of discourse on our platforms.

Under this approach, Meta reduces the following non-violating, vaccine-related content: (1) sensationalist or alarmist vaccine content, such as content using conspiratorial terms to suggest that vaccines are unsafe; (2) content that criticizes a person's choice to receive or provide vaccines, such as content disparaging a parent for vaccinating their child; (3) content that promotes vaccine refusals or alternatives to vaccination; and (4) shocking stories about vaccines, including potentially true events that raise safety concerns about vaccines (e.g., a story about someone dying after receiving a vaccine). We also may reduce the distribution of Pages, groups, and Instagram accounts that are focused on spreading this type of content.