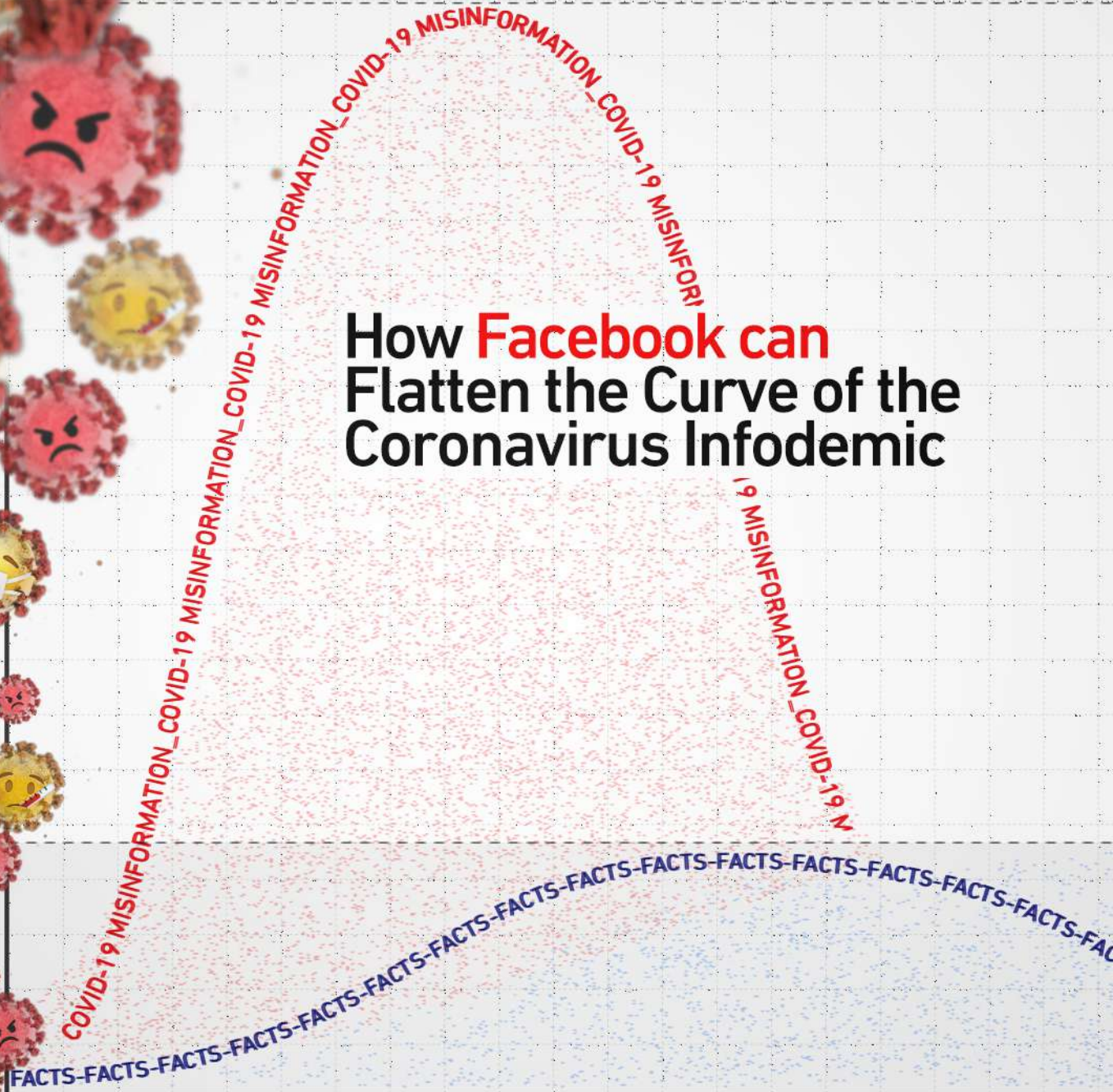


How Facebook can Flatten the Curve of the Coronavirus Infodemic





How Facebook can Flatten the Curve of the Coronavirus Infodemic

Study indicates Facebook is rife with bogus cures and conspiracy theories that remain on the platform long enough to put millions of people at risk

April 15, 2020

Contents

1. [Executive Summary](#)
2. [Key Findings](#)
3. [How Facebook Can Inoculate Users Against Coronavirus Misinformation](#)
4. [Methodology and Data Set](#)
5. [Annex](#)

Executive Summary

As the coronavirus crisis started to spread across the globe, the World Health Organization (WHO) warned that “we’re not just fighting an epidemic; we’re fighting an *infodemic*.” It cautioned that “fake news spreads faster and more easily than the virus, and it is just as dangerous.”¹

In response to the crisis, Facebook CEO Mark Zuckerberg and other company executives launched a media blitz to publicize the company’s expanded efforts to stop the spread of COVID-19 misinformation. Facebook announced that these efforts had been “quick,” “aggressive,” and “executed...quite well.”²

In February, our team began detecting and monitoring widespread misinformation about COVID-19 online. In March, our investigative team set out to analyse and assess the efficacy of Facebook’s efforts to combat this “infodemic” on its main platform.

For this study, of the thousands of pieces of coronavirus-related misinformation content being shared on Facebook, we decided to examine over 100 pieces of misinformation content in six different languages about the virus that were rated false and misleading by reputable, independent fact-checkers and could cause public harm.³

We found that **millions of the platform’s users are still being put at risk of consuming harmful misinformation on coronavirus at a large scale.** Representing only the tip of the misinformation iceberg, we found that the pieces of content we sampled and analysed were shared **over 1.7 million times** on Facebook, and viewed **an estimated 117 million times**.

Even when taking into consideration **the commendable efforts Facebook’s anti-misinformation team** has applied to fight this infodemic, the platform’s current policies were insufficient and did not protect its users.

Of the 41% of this misinformation content that remains on the platform without warning labels, 65% has been debunked by partners of Facebook’s very own fact-checking program⁴. Throughout the timeframe of our research⁵, this content remained

on the platform despite the company’s promise to issue “strong warning labels” for misinformation flagged by fact-checkers and other third party entities, and remove misinformation that could contribute to imminent physical harm.⁶

Secondly, Avaaz found that there are **significant delays in Facebook’s implementation of its anti-misinformation policies.** These delays are especially troubling because they result in **millions of users seeing harmful misinformation content** about the coronavirus before the platform labels it with a fact check and warning screen or removes it. Specifically, we found that **it can take up to 22 days for the platform to downgrade and issue warning labels** on such content, giving ample time for it to go viral.

Our analysis also indicates that **Italian and Spanish-speaking users may be at greater risk of misinformation exposure.** Facebook has not yet issued warning labels on 68% of the Italian-language content and 70% of Spanish-language content we examined, compared to 29% of English-language content.

The scale of this “infodemic” along with **Facebook’s reluctance to retroactively notify and provide corrections to every user exposed to harmful misinformation about the coronavirus is threatening efforts to “flatten the curve” across the world and could potentially put lives at risk.**

In a conversation with members of Facebook’s misinformation team on 13 April 2020, we were strongly encouraged by unprecedented commitments from Facebook to institute retroactive alerts to fight coronavirus misinformation, an important and necessary first step that could potentially save lives⁷.

To ensure that content that could cause imminent physical harm to Facebook users was speedily removed by the platform, we shared with the company on 8 April 2020 a list of misinformation posts that we believe violate its policies and recommended that all be removed. In response, Facebook has, as of 14 April 2020, taken down 17 of the posts we flagged, cumulating 2.4 million estimated views.

Summary of Key Findings of the Analysed Sample:

- **Facebook is an epicentre of coronavirus misinformation, and the platform must do more to protect its users from this infodemic.**
 - The 104 coronavirus-related pieces of misinformation content analysed in this report have reached over **117 million estimated views**⁸, even though this sample represents only the tip of the iceberg.
 - Of the **41% of this misinformation content** that remains on the platform **without warning labels**, **65% had been debunked by partners of Facebook's very own fact-checking program.**
- **Facebook's current anti-misinformation efforts remain slow and insufficient in limiting the spread of coronavirus misinformation, even when the content is flagged by the platform's very own fact-checking partners.**
 - **It can take up to 22 days for the platform to downgrade and issue warning labels** on false content, even when Facebook's fact-checking partners as well as the World Health Organization and local health authorities, move at a significant pace to issue corrections.
 - These delays mean that millions of users see, interact with, and share the misinformation content before it is labeled and/or removed.
- **Italian, Spanish and Portuguese-speaking users appear to receive significantly fewer warning labels from Facebook, and hence are at greater risk of being exposed to misinformation.**
- Over half (51%) of non-English misinformation content had no warning labels.
- 29% of fact-checked English-language misinformation content analysed, as of April 7th 2020, was found **without** clear warning labels from Facebook.
- But our sample indicates Facebook remains significantly less effective in moderating COVID-19 misinformation in Italian, Spanish, and Portuguese.
- **Coronavirus misinformation content mutates and spreads faster than Facebook's current system can track it.**
 - One effective misinformation story can mutate, infect, and then lodge into and remain across hundreds of Facebook groups and pages without detection.
 - For example, a harmful misinformation [post](#)⁹ that claimed that one way to rid the body of the virus is to drink a lot of water and gargle with water, salt or vinegar was **shared over 31k times** before eventually being taken down after Avaaz flagged it for Facebook. But the other [2,611 clones of that false and misleading post remain on the platform with over 92,246 interactions](#). Most of these cloned posts have no warning labels from Facebook.

Some of the most egregious falsehoods promoted by the content examined in this study include:

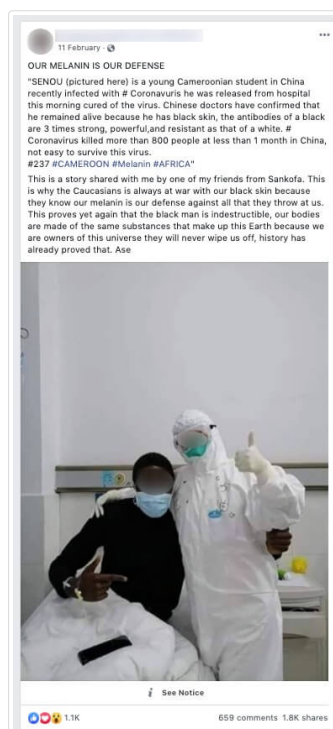
FALSE CLAIM

"Black people are resistant to Coronavirus"¹⁰

The false claim that those with black skin are more resistant to the coronavirus.

3,604 best available Interactions

112,830 estimated views



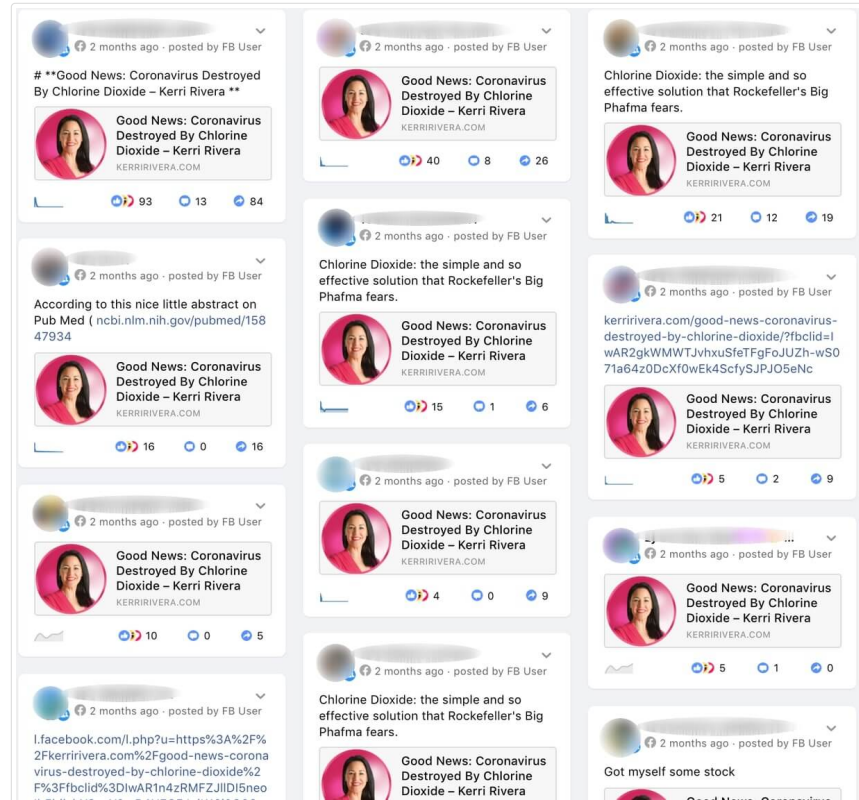
FALSE CLAIM

"Good News: Coronavirus is destroyed by Chlorine Dioxide"¹¹

The false claim that chlorine dioxide cures the virus.
(6,368 best available Interactions and 199,362 estimated views)

6,368 best available Interactions

199,362 estimated views



FALSE CLAIM

"Vice President Pence urges those with Coronavirus to go to the Police"¹²

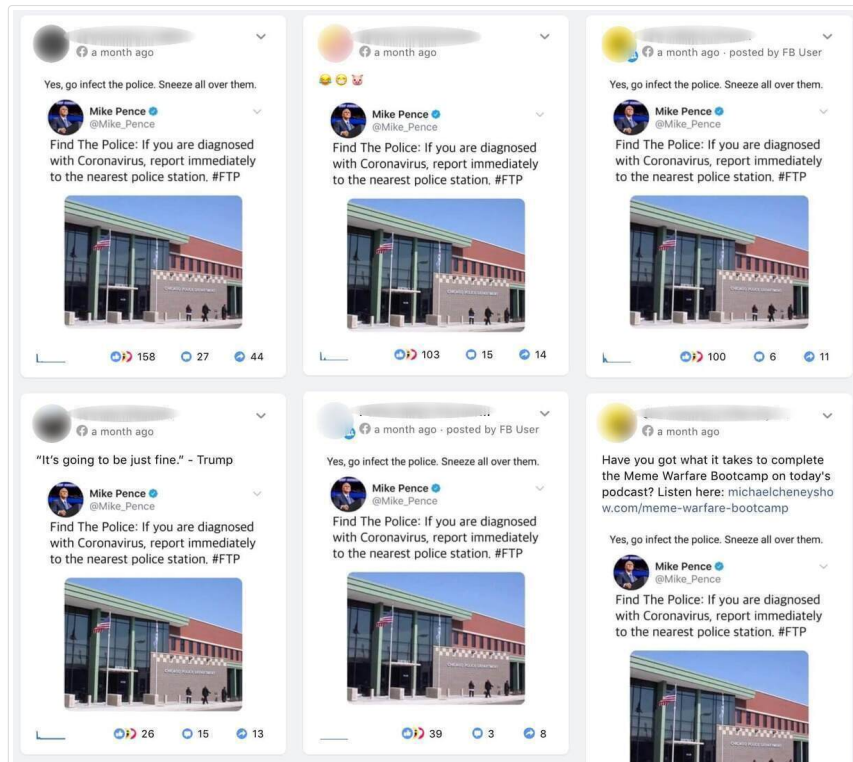
(The tweet in the image is fake, Vice President Pence did not tweet this)¹³

The false claim that Vice President Mike Pence advised those diagnosed with the virus to go to the nearest police station.

694 best available Interactions

21,727 estimated views





FALSE CLAIM

["Hairdryers could be used for coronavirus prevention"](#)¹⁴

The false claim that using hot air from a hair dryer would fight the virus.

119,057 best available Interactions

2,600,000 Facebook views



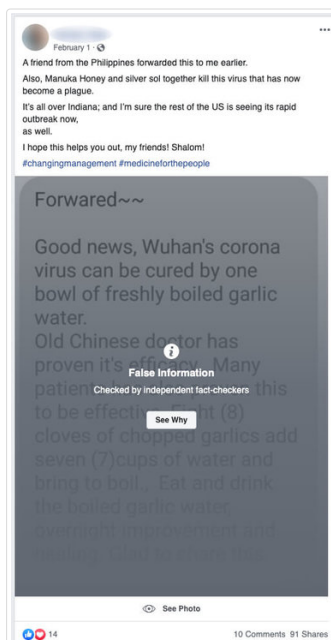
FALSE CLAIM

["Virus can be cured by one bowl of freshly boiled garlic water"](#)¹⁵

The false claim that the virus can be cured with garlic boiled in hot water.

115 best available Interactions

3,600 estimated views



Although Facebook has taken a number of commendable steps to fight the COVID-19 infodemic, such as providing the WHO with free advertising and promoting authoritative content to millions of users through the COVID-19 hub, our findings raise the question of whether the company is doing all that it can and must do to protect its users. Our findings suggest that the platform must adopt stronger steps to fight misinformation to remain in alignment with official statements it has made to the public about its efforts to fight coronavirus-related misinformation.

Notably, Facebook still has an opportunity to “flatten the curve” of this “infodemic.” After consultation with industry experts, academics, doctors, and legislators around the world, **Avaaz recommends that Facebook urgently:**

- **“Correct the Record”:** This means that every user who has been exposed to disinformation and/or misinformation¹⁶ should be alerted and provided with independently fact-checked corrections.
- **Start with “Correcting the record” for coronavirus misinformation:** Facebook can immediately start by **retroactively sending** a notification and a correction from the WHO, the Center for Disease Control and Prevention (CDC), and/or credible, independent fact-checkers to every user who interacted with harmful misinformation posts related to COVID-19 before Facebook moderators detected and labeled the content as false or removed it.
- **Scale up moderation and fact-checking capacities, and continue to expand its work with civil society researchers:** Facebook must bring its moderation capacities to scale throughout the world to swiftly detect and to work with fact-checkers to aggressively fact-check, downgrade and issue strong warning labels on misinformation content before it

goes viral. Facebook must also share more detailed information about the reach of false and misleading content on its platform, the speed at which the company detects and removes this misinformation, and other key data to ensure researchers can work with the platform to innovate better solutions to these challenges.

Findings from a new academic study prove that “Correct the Record” can cut belief in misinformation and/or disinformation **by an average of 50 percent, and by as much as 61 percent.** Avaaz believes that Facebook can be a trailblazer in the fight against misinformation and potentially save lives by implementing this solution in full.

Failing to implement the recommendations above to quickly and effectively quarantine coronavirus misinformation means that there will continue to be large numbers of individuals who will be exposed to this harmful content on the platform, believe it, never receive a correction, and thus act upon the misinformation, putting themselves and those around them at risk. It also means that life-saving services provided by the WHO, local health authorities, and the doctors and nurses, who are on the frontlines of this crisis, will continue to become more difficult as unknown numbers of social media users heed misleading medical advice. Lives could potentially be lost.

In the event that Facebook and other social media platforms refuse to act promptly to “Correct the Record”, **Avaaz recommends that the WHO, the CDC, local health authorities, and legislators throughout the world urge the platforms to adopt “Correct the Record” immediately.**

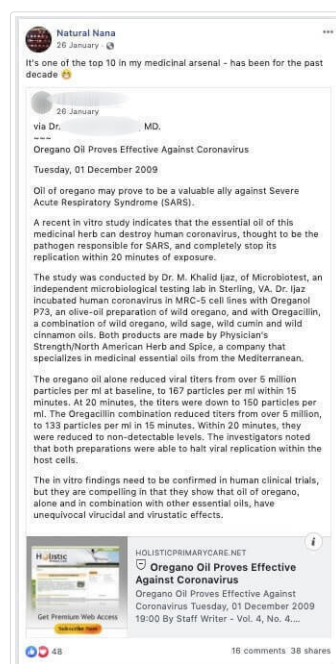
Key Findings

FACEBOOK IS AN EPICENTRE OF CORONAVIRUS MISINFORMATION

In an official statement, Facebook announced that “we remove COVID-19 related misinformation that could contribute to imminent physical harm,” and that “once a post is rated false by a fact-checker, we reduce its distribution so fewer people see it, and we show strong warning labels and notifications to people who still come across it, try to share it or already have.”¹²

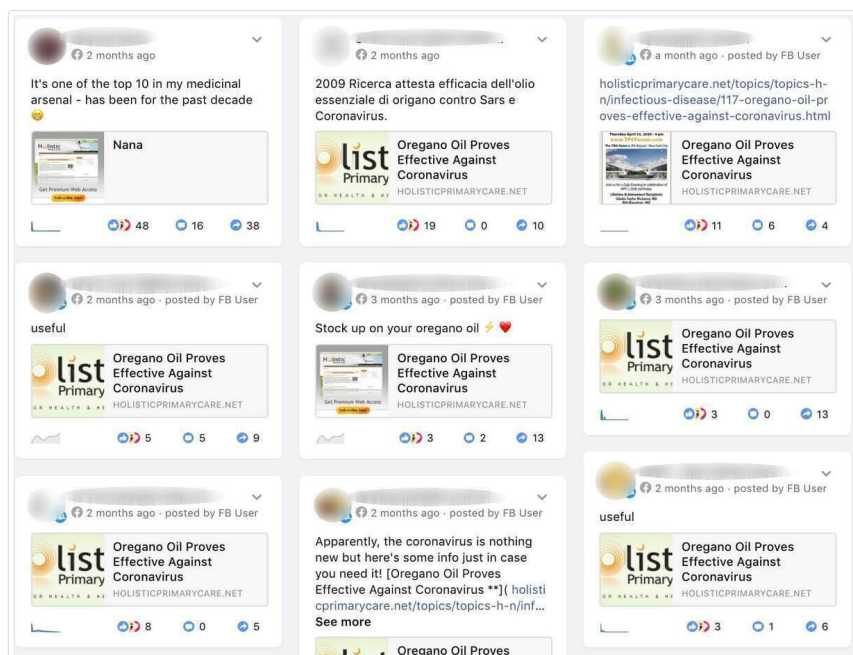
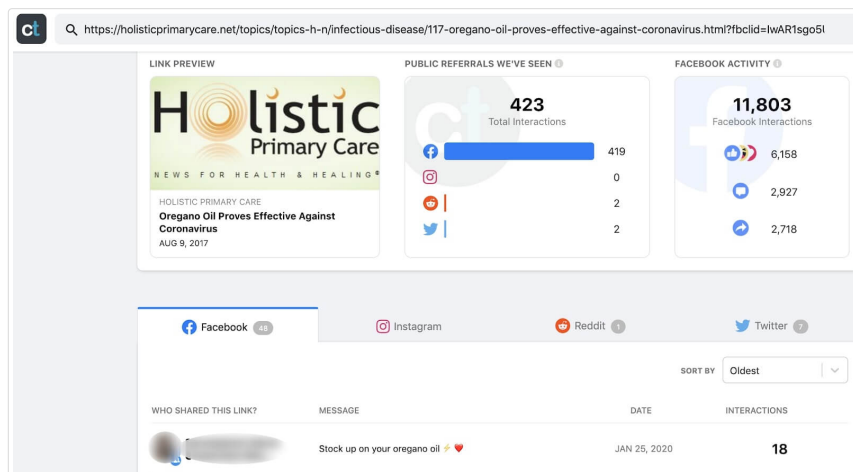
Our findings indicate a gap between these promises and what is actually happening on the platform. 43 pieces of misinformation content that Avaaz analysed on Facebook (41% of our *entire sample*) have, as of April 7 2020, no warning labels to inform users of falsehoods and potential harms even though *fact checks are publicly available for all 104 posts*.

- 28 out of those 43 pieces of content with no warning labels from Facebook (65%) have been fact-checked by one of Facebook’s Third-Party Fact-Checking Program partners. This indicates that even when Facebook fact-checking partners detect and issue a correction for coronavirus virus, the platform still has trouble labeling and downgrading it.
- For example, this [post¹⁸](#) by [Natural Nana¹⁹](#) (a page whose banner reads “Modern medicine will cure you to death”) from January 26, 2020 claims that oregano oil “can destroy human coronavirus” and “completely stop its replication within 20 minutes of exposure”. The link²⁰ shared in this post has 6,158 reactions and 2,718 shares, with 100% of the “Top Referrals” for this



link²¹ on Facebook coming after 25 January 2020. Our findings indicate that the story was still being widely shared in March with the associated claim that oregano oil can kill COVID-19. This post was often shared by actors interested in marketing oregano oil with posts claiming it cures the virus. However, the original claim this post refers to is a decade old, originating from the website holisticprimarycare.net. Scientists have said there is [no such cure for COVID-19](#).

The heavy online traffic this refurbished story received as a result of the pandemic led the editor of the website to issue a clarification debunking the claim on the website itself, yet the claim that “Oregano is a cure for coronavirus” continues to spread on Facebook without any warning labels to inform users who see this post that the information is false and could lead to public harm.



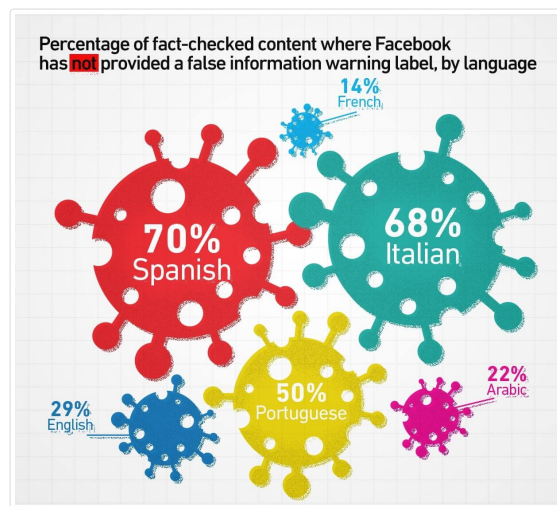
24% of the 104 pieces of misinformation content we analysed could contribute to imminent physical harm (according to Facebook’s content moderation policies)²², yet, as of April 7th 2020, only 8% of the misinformation content has been removed by Facebook.²³

Facebook’s Chief Operating Officer Sheryl Sandberg said “we want people to get accurate information” and “we’re removing things quickly if they’re harmful.”²⁴

ITALIAN, SPANISH AND PORTUGUESE SPEAKING USERS APPEAR TO BE AT GREATER RISK OF UNCHECKED MISINFORMATION EXPOSURE

For 29% of fact-checked, English-language misinformation content in our sample, Facebook has not issued warning labels alerting users to the falsehoods of this content. But the situation is very different for content in other languages, putting users in other countries that have been hit hard by the pandemic at greater risk of misinformation exposure. Avaaz's investigation found that as of 7 April 2020 **over 51% of the non-English content analysed, in total, has yet to receive warning labels by Facebook**. More specifically, these labels have ***not*** been applied to:

- 70% (7 out of 10) of Spanish-language content (to date, shared 510k times overall);
- 68% (13 out of 19) of Italian-language content (to date, shared 307k times overall);
- 50% (7 out of 14) of Portuguese-language content (to date, shared 78k times overall);
- 22% (2 out of 9) of Arabic-language content (to date, shared 453k times overall);
- 14% (1 out of 7) of French-language content (to date, shared 44k times overall), faring better than the scale at which English-language content was appropriately labeled.



Here is a more in-depth look at each language-specific subset of content analysed:

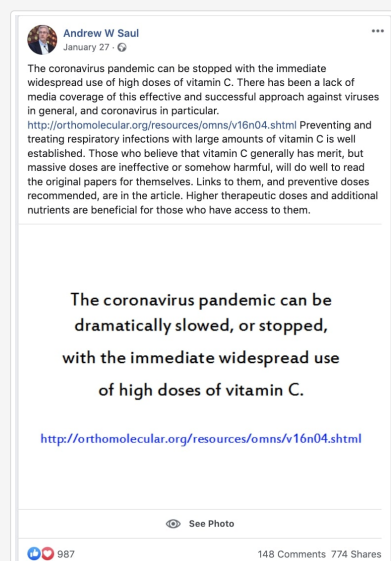
ENGLISH-LANGUAGE CONTENT:

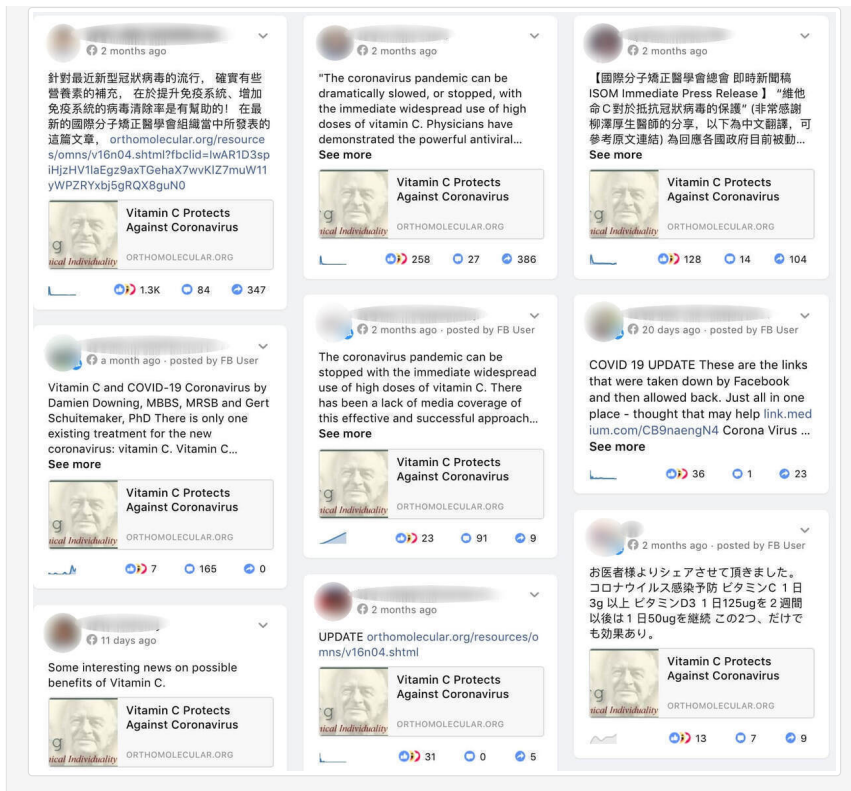
- 29% of the English-language misinformation content analysed (shared 290k times) has not been labeled by Facebook.
- 16% of all the English stories fall into the category of the exact kind of violative harmful content that Facebook has publicly committed to remove.²⁵
- For example, in the [post²⁶](#) below and the accompanying [article²⁷](#), Andrew W Saul claims, "The coronavirus pandemic can be stopped with the immediate widespread use of high doses of vitamin C." As of 7 April 2020, posts linking the article have received an estimated 301,110 views.

As fact-checking website, Lead Stories concludes, "vitamin C does not slow or stop coronavirus"; and "The WHO is warning against misinformation regarding the prevention and treatment of the virus, including claims that taking large doses of vitamin C, [...] will either prevent or slow the spread of it."²⁹

Although the original post has now been removed by Facebook after being flagged by Avaaz,³⁰ a number of posts by others sharing the same story remain on the platform without a warning label.

["The coronavirus pandemic can be stopped with the immediate widespread use of high doses of vitamin C"](#)²⁸





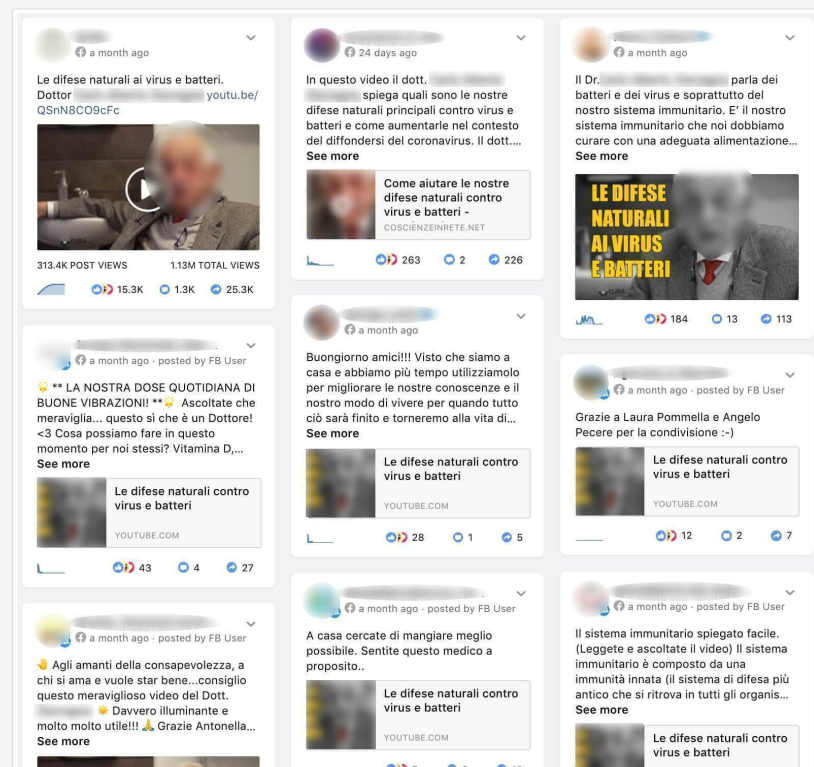
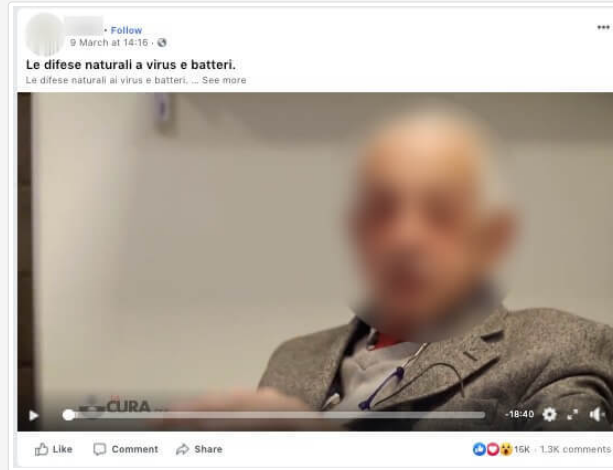
PORTUGUESE-LANGUAGE CONTENT:

- 50% of the Portuguese-language content analysed (shared 78k times overall) has not been labeled by Facebook.
- 43% of the Portuguese samples fall into the category of misinformation that could lead to imminent harm.³¹
- For example, in [this video](#)³² posted by Brazil's President Jair Bolsonaro, he claims that the anti-malarial drug, hydroxychloroquine, effectively cured 78 out of 80 COVID-19 patients in a study and is "already a reality," and, therefore, is a viable medicine to treat the virus. Nonetheless, as numerous experts have highlighted, the use of hydroxychloroquine to treat COVID-19 patients is controversial and there is not yet enough research to provide conclusive evidence of its safety or effectiveness³³. Misinformation about hydroxychloroquine as a viable treatment method has already been linked to several deaths in the US.³⁴ As of 10 April 2020, this video has been viewed **over 1.5 million times** on Facebook.³⁵



ITALIAN-LANGUAGE CONTENT:

- 68% of Italian-language misinformation content we analysed (shared 307k times overall) has not been labeled by Facebook.
- 21% of the Italian misinformation posts sampled fall into the category of harmful content that Facebook has publicly committed to remove, and these posts were still up on the platform as of 7 April 2020.³⁶
- For example, in one [video](#)³⁷ the speaker, who claims to be a medical doctor, says that staying home is useless against the coronavirus and that our natural bacteria combined with burning incense are sufficient to fend against infection. **This video has been viewed over 1.1 million times on Facebook as of 7 April 2020.**



SPANISH-LANGUAGE CONTENT:

- 70% of the Spanish-language content analysed (shared 510k times) has not been labeled by Facebook.
- 20% of the overall Spanish misinformation content sampled falls into the category of harmful content that Facebook publicly committed to remove.³⁸
- For example, one meme reads, "Let's avoid more deaths," (Avaaz translation) and claims that the way to rid one's body of the virus is to drink a lot of water and gargle with water, salt or vinegar. **The original [post](#)³⁹, shared over 31,000 times on Facebook, has now been taken down after Avaaz notified Facebook about it⁴⁰. But the other 2,611 clones of this meme remain on the platform with over 92,246 interactions. Most of these cloned posts have no warning labels from Facebook.** This example points to another gap in Facebook's enforcement of its anti-misinformation policies: Facebook may add a warning label on a specific piece of misinformation content shared as an image, but that the warning label is not replicated when the image is posted or shared again by others.

ARABIC-LANGUAGE CONTENT:

- 22% of the Arabic-language content analysed (shared 453k times) has not been labeled by Facebook.
- 56% of the overall Arabic content falls into the category of harmful content that Facebook publicly committed to remove, and 80% of the posts that may cause physical harm were still on the platform as of 7 April 2020⁴¹.
- For example, a [video](#) in Arabic⁴², posted on 20 March 2020, claimed that hair dryers could be used for coronavirus prevention. The fact-checked video was taken down 11 days after it was posted. It had been viewed at least 2.6 million times between posting and removal.



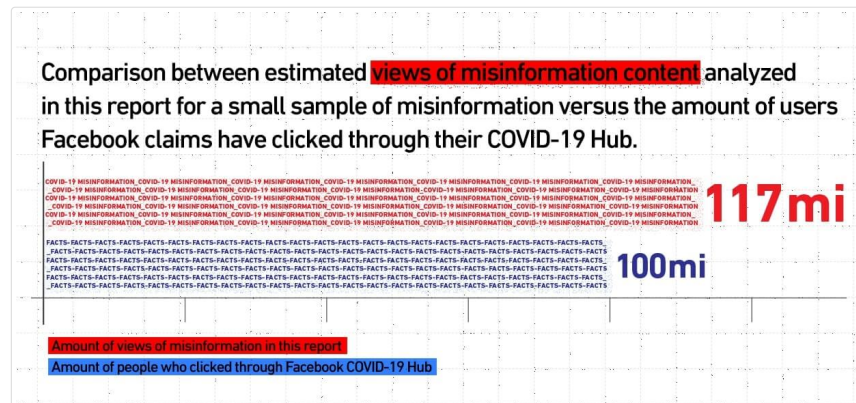
FRENCH-LANGUAGE CONTENT:

- 14% of the French-language content we analysed (shared 44k times) has not been labeled by Facebook.
- 14% of the overall French content sampled falls into the category of harmful content that Facebook committed to remove.⁴³
- For example, the [article](#)⁴⁴ linked in this [post](#)⁴⁵ claims that coronavirus dies at temperatures above 26/27 degrees Celsius and one can prevent infection by gargling with disinfectant.



VIRUS MISINFORMATION IS STILL GOING VIRAL - REACHING TENS OF MILLIONS OF FACEBOOK USERS

- The **104 pieces of coronavirus misinformation** content Avaaz analysed on Facebook have been shared almost **1.7 million times** and have reached over **117 million estimated views**. It is important to note that these numbers represent the tip of the iceberg given this investigation analysed a small sample of misinformation in only 6 languages and was designed to investigate whether Facebook's response was as effective as the company claims, and not the overall scale of misinformation on the platform.
- The significant reach of the 104 pieces of content suggests that Facebook's efforts to stem the spread of coronavirus misinformation are not as effective as company executives have suggested.⁴⁶
- On 25 March 2020, Facebook announced that 100 million users clicked through its newly released "COVID-19 Information Center."⁴⁷ If the small selection of misinformation content sampled for this analysis (104 posts) were viewed an estimated 117 million times, this strongly suggests that hundreds of millions of Facebook users across the world have likely clicked through the vast amount of misinformation content spreading across the platform.



A better understanding of the true scale of this infodemic requires more openness from Facebook. Specifically, Facebook continue to update researchers with answers to the following questions:

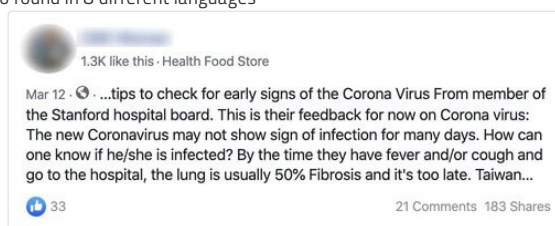
- How many of its users have seen and clicked on harmful misinformation about COVID-19?
- How many users have seen a warning label from one of its fact-checking partners?
- Will Facebook begin to retroactively alert and provide independently fact-checked corrections from the WHO and its fact-checking partners to every user who has been exposed to misinformation on its platform?

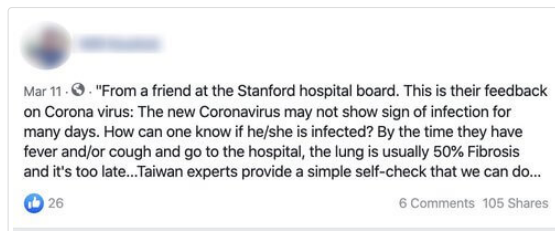
One piece of misinformation content can mutate then infect hundreds of Facebook accounts, groups and pages and thrive for days to weeks without detection.

- For example, our team **detected 28 variations of one distinct misinformation post that contained false and misleading medical advice**. Each mutation of this post maintains the same overall content and form, but the source attribution and language vary. Posts in this chain of deception claim that the information they are promoting is from the WHO, UNICEF, local health authorities (including the Canadian Department of Health), and/or renowned hospitals (including the Stanford Hospital Board) and doctors across at least 14 countries (USA, Nigeria, Egypt, France, Japan, Australia, Italy, Philippines, Ghana, Ethiopia, India, Marshall Islands, Burma, Brazil). They are also found in 8 different languages

(English, Portuguese, Arabic, French, Chinese, Burmese, Punjabi, and Italian).

- The 28 posts promoted an almost identical list of misleading medical advice, including the debunked claim that "taking a deep breath and holding it for more than 10 seconds" is a viable way to find out if you have COVID-19.
- Many of these posts are still being shared on Facebook without any warning labels. For example, up until 6 April 2020, we continued to find the **false post** citing the "Stanford Hospital Board" still active across dozens of profiles and groups on Facebook without any fact checks or labels despite the fact that it was debunked by the Associated Press⁴⁸ a Facebook fact-checking partner, on 12 March 2020. The Stanford Health Center also debunked the claim.⁴⁹

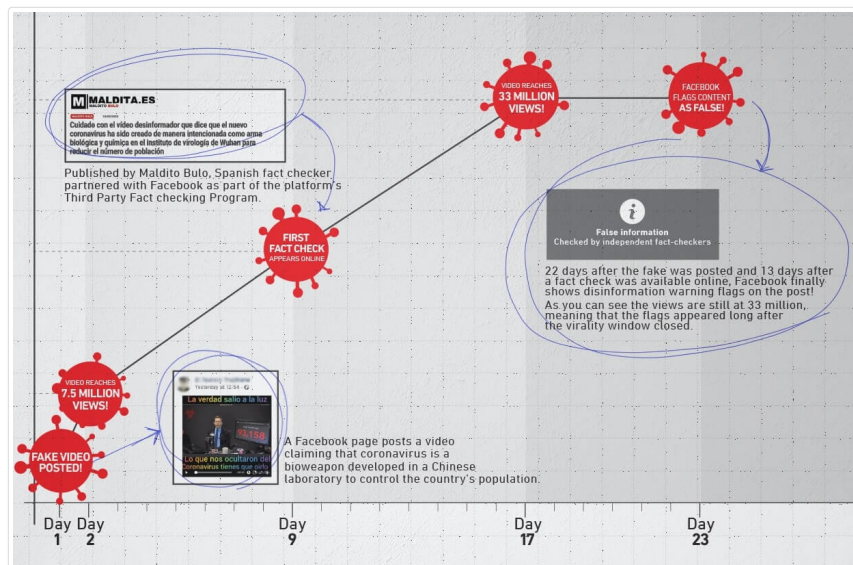




Facebook's current counter-misinformation efforts remain slow *and insufficient* in limiting the spread of coronavirus misinformation, even when the content is flagged by the platform's very own fact-checking partners

- For the subset of coronavirus-related misinformation content in our study for which Facebook added a warning label, there was an average **5-day delay between when harmful misinformation content was posted and the release of a fact check** by a member of Facebook's Third-Party Fact-Checking Program.
- To assess how quickly Facebook issues warning labels in response to its fact-checking partners' fact checks, Avaaz followed 4 false stories that were debunked by Facebook's program partners and measured the number of days between the moment the stories were posted on the platform and when they were appropriately labeled as false. For each piece of content, **it took 1, 4, 10 and 13 days respectively for the platform to issue warning labels after its partners had issued a correction**. This means that, on average, it can take up to **7 days for a post to be labeled by Facebook after the platform had a fact check available from its partners**.
- If the anecdotal results above reflect what's happening across Facebook, this would mean that **it could take on average a total of 12 days for Facebook to issue warning labels on coronavirus misinformation content, or to remove content that may cause imminent harm**.
- This means that millions of users who see the content before these labels are enforced will likely never know what they saw contained false and misleading content.
- For example, it **took 22 days** for Facebook to label a misinformation [video in Spanish](#)⁵⁰ claiming that the virus was deliberately created in a lab at the Institute of Virology in Wuhan. **By that time, the video had reached 33 million views**. It took **9 days for the video to be fact-checked by a Facebook partner** and then **another 13 days for the video to be labeled by Facebook as misinformation**.

Here's a closer look at the sequence and timing of Facebook's actions:



- Independent fact-checkers have moved quickly to respond to the deluge of misinformation around COVID-19, issuing fact checks within, **on average, 5 days from when the misinformation content is posted online**. Fact-checking organisations alongside the WHO and the CDC are performing a Herculean public service by issuing corrections to this flood of misinformation - but Facebook is not doing enough to match these efforts by ensuring that fact checks reach those who are exposed to and potentially harmed by the misinformation.
- Facebook has often sought to shift the responsibility for delays in catching and labeling misinformation to independent fact-checkers. By doing so, the company is deflecting its responsibility as the platform has the ability to retroactively notify *all* its users who have seen harmful misinformation.

How Facebook Can Inoculate Users Against Coronavirus Misinformation

From fake cures to harmful advice that weakens efforts to sustain social-distancing, misinformation about the coronavirus on Facebook could potentially cost lives and hinder attempts to fight this pandemic.

The findings above show that Facebook's current efforts to fight misinformation are inadequate. Misinformation actors know that they can spread sophisticated and malicious lies such as the fake "Stanford Medical Advice" post claiming that "holding your breath for ten seconds is a test for whether one has Coronavirus" or to market a fake cure such as chlorine dioxide. They know that before Facebook gets to fact-checking and labeling or removing this harmful content, millions of people will have seen the story and possibly believed it.

But Facebook can, today, act to retroactively inform those victimized by misinformation so that even if it takes a week to catch a malicious misinformation story about COVID-19, all those who have seen this misinformation are notified and receive a correction from the WHO, the CDC, or a relevant independent fact-checker. This solution, called "Correct the Record", could save lives and significantly disincentivize misinformation actors by retroactively inoculating Facebook users against their content.

A new academic study proves that providing social media users who have seen false or misleading information with corrections from fact-checkers can

decrease belief in disinformation⁵¹ by an average of **50 percent, and by as much as 61 percent.**

The academic study on corrections⁵², commissioned by Avaaz, was conducted by Dr. Ethan Porter of George Washington University and Dr. Tom Wood of Ohio State University, who are leading experts in the study of corrections of false information.

In order to test the effectiveness of corrections, a hyper-realistic visual model of Facebook⁵³ was designed to mimic the user experience on the platform. Then a representative sample of the American population, consisting of 2,000 anonymous participants, chosen and surveyed independently by YouGov's Academic, Political, & Public Affairs Research branch, were randomly shown up to 5 pieces of false news that were based on real, independently fact-checked examples of false or misleading content being shared on Facebook.

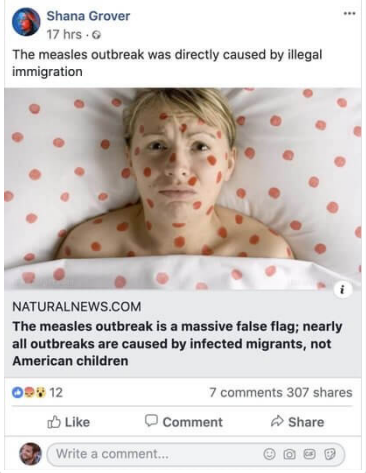
Through a randomized model, some of the users, after seeing the false news, were shown corrections. Some users saw only the false or misleading content, and some saw neither. Then the surveyed participants answered questions designed to test whether they believed the false news.

In addition to the conclusion that corrections can effectively counter American social media users' belief in disinformation, cutting it in half, the study also found that:

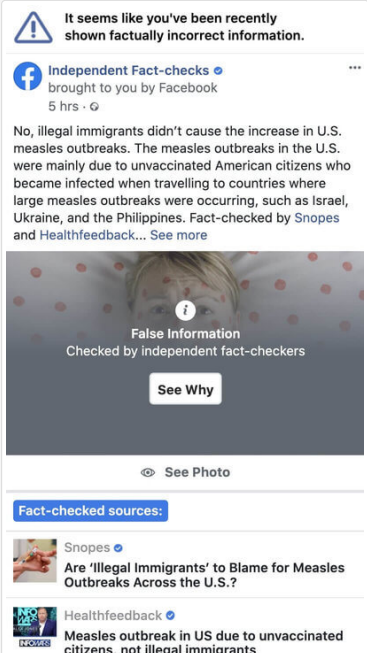
- The average drop in belief in false or misleading information after seeing a correction was **50% for Republicans and 47% for Democrats.**
 - **61.5% fewer people believed a false news story** that claimed President Trump said Republicans are the dumbest group of voters.
 - **50.7% fewer people believed** a false news story that claimed the measles outbreaks were caused by illegal immigrants after being exposed to a correction versus those that saw only the fake item.
 - **50% fewer people believed** a false news story that claimed Greta is the highest paid activist in the world and had made millions from her activism.
 - **42.4% fewer people believed** a false news story that claimed that a photo of a female Somali soldier circulating on social media showed US Representative Ilhan Omar at a terrorist training camp in Somalia.
 - **42% fewer people believed** a false news story that claimed that installing 5G towers required a hazmat suit.

AN EXAMPLE OF HOW THE CORRECTIONS TESTED WOULD LOOK IS PRESENTED BELOW:

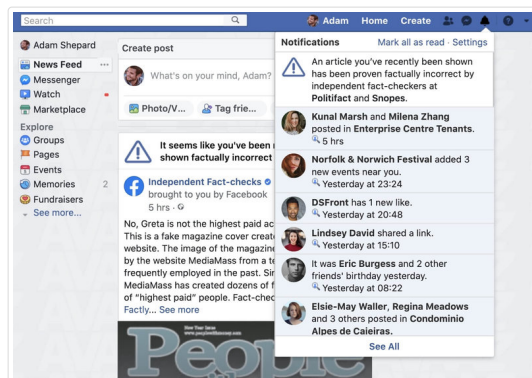
Users who happen to see a misinformation post like this:



will see a notification like this in their newsfeed and via their notification centre, when the post is eventually fact-checked by an independent fact-checker:



And will also receive a notification like this notifying them them to review the correction:



Facebook has the capacity to design and implement a correction model similar to the above, and possibly more sophisticated. The platform has a social responsibility to inoculate users against misinformation and disinformation by correcting the record. But if the platform refuses to implement this solution, the WHO, the CDC, and local health authorities throughout the world have significant public support to oblige Facebook to protect its users:

- Out of 1,000 Americans polled who participated in the study, **68.4% agree or strongly agree** that Facebook should inform users when they have been targeted with disinformation, and provide them with corrections from independent fact-checkers.
- **60.8% agree or strongly agree** that social media companies should be required by law to protect their users from disinformation, with only 18.1% disagreeing.
- **On average 87% of citizens in Germany, France, Spain and Italy** want social media platforms to work with fact-checkers to issue corrections to anyone exposed to false or misleading content.³⁴

Avaaz recommends that the WHO, the CDC, local health authorities, and lawmakers throughout the world use their leverage, and significant public support, to engage with Facebook and push the platform to "Correct the Record."

Methodology and Data Set

To better understand the scale of coronavirus-related misinformation on Facebook and the platform's handling of the "infodemic," Avaaz collected and monitored relevant misinformation content in 6 languages: English, Italian, Spanish, French, Portuguese, and Arabic.

For the purpose of measuring Facebook's stated claims about its fact-checking efforts, the investigative team analysed 104 verifiably false or misleading posts about the coronavirus that, cumulatively:

1. Were fact-checked by Facebook's Third-Party Fact-Checking partners or other reputable fact-checking organisations⁵⁵. To double-check the accuracy of all fact checks, Avaaz hired Health Feedback to review all samples chosen for this study. Health Feedback verifies scientific claims and is a member of the WHO-led project Vaccine Safety Net (VSN). They are part of the non-profit organisation, Science Feedback - a third-party fact-checking partner of Facebook.^{56 57}
2. Could cause public harm by undermining public health. Avaaz has included false and misleading content which impacts public health in the areas of:
 - A. *Preventing disease*: e.g. false information on diseases, epidemics and pandemics, and anti-vaccination misinformation.
 - B. *Prolonging life and promoting health*: e.g. bogus cures and/or encouragement to discontinue recognised medical treatments.
 - C. *Creating distrust in health institutions, health organisations, medical practice, and their recommendations*: e.g. false information implying that clinicians or governments are creating or hiding health risks.
 - D. *Fearmongering*: health-related misinformation which can induce fear and panic, e.g. misinformation stating that the coronavirus is a human-made bio-weapon being used against certain communities or that Chinese products may contain the virus.
 - E. *Health misinformation with the potential of inducing discrimination against minorities*: e.g. migrants are spreading the virus.

For each of the false and misleading posts and stories sampled based on the above criteria, Avaaz researchers recorded and analysed, using both direct observation and Crowdtangle software:

- The number of times each of these stories was shared and the total number of interactions it received;
- An estimate of the number of times each was likely viewed on the platform;⁵⁸
- Whether each had a warning label as false and misleading added to it by Facebook;⁵⁹

- When misinformation posts would receive a fact-check warning label or be removed (researchers monitored misinformation posts with no warning labels daily);
- The delay between when the misinformation content was posted and the publication of a fact-check by a Facebook partner website;
- Where applicable, researchers measured the delay between when the original misinformation content was posted and the date when Facebook first applied its moderation policies on the post by either flagging the content as misinformation or removing it from the platform;

A list of all the misinformation content analysed for this report can be found in the appendix.

It is important to note that although fact checks from reputable fact-checking organisations provide a reliable way to identify misinformation content, fact-checkers have limited access to misinformation spreading in private Facebook groups, on private Facebook profiles, and via Facebook messenger.

Similarly, engagement data for Facebook posts analysed in this study are only indicative of wider engagement with and exposure to misinformation. In many cases, the misinformation content sampled in this study were cloned, and spread by many accounts across Facebook, yet we can only measure the interactions with the sampled piece of content in this report. Still, the engagement data we estimate for our sample provides some indication of the relative reach of different claims.

It is important to note that, while we collect data and compute numbers to the best of our ability, this analysis is neither comprehensive nor is it exhaustive (we looked only at a sample of fact-checked misinformation posts in six languages). Moreover, this research is made significantly more challenging because Facebook does not provide investigators with access to the data needed to measure the total response rate, moderation speed, number of fact checks, and the amount of users who have seen or been targeted with misinformation. Nonetheless, Facebook is becoming more cooperative with civil society organisations, and we hope the platform continues this positive trend.

Importantly, this study achieved its purpose by taking a step towards a better understanding of the scale and scope of the COVID-19 misinformation infodemic on Facebook.

Cooperation across fields, sectors, and disciplines is needed more than ever to fight disinformation and misinformation. All social media platforms must become more transparent with their users and with researchers to ensure that the scale of this problem is measured effectively, and to help public health officials respond in a more effectual and proportional manner to both the pandemic and "infodemic."

VIEWERSHIP CALCULATION FOR VIDEOS ON FACEBOOK

Our fake news selection contained 13 examples of Facebook fake news video content accumulating a total of 47,317,476 views and 1,511,408 interactions between the period of 21 January 2020 and 7 April 2020.

VIEWERSHIP ESTIMATION FOR IMAGES AND TEXT

Facebook discloses the number of views for videos, but for posts containing only text and image content the platform displays only the number interactions (which are shares, reactions and comments). Therefore, in order to estimate viewership for text and image content, we designed a ratio based on the actual views divided by the total interactions of the 13 Facebook misinformation videos present in this study. The ratio applied to all images and text posts is:

$$47,317,476 / 1,511,408 = 31,31.$$

Estimated views calculation: For each image and text posts, we multiplied our views/interaction ratio of 31,31 per the highest available number of interactions for those posts or web links provided by CrowdTangle. Adding all those estimated views together to the real Facebook video views we obtain the final estimate of 117,714,168 estimated views for the sampled posts between the dates of 21 January 2020 and 7 April 2020.

For more information and interviews:

- media@avaaz.org
- Andrew Legon (CET Timezone) - andrew.legon@avaaz.org / +34 600 820 285
- Ahmed Alsalman (EST Timezone) - ahmed@avaaz.org / +1 313 247 3679

MORE INFORMATION ON AVAAZ'S DISINFORMATION WORK:

Avaaz is a global democratic movement with more than 56 million members around the world. All funds powering the organisation come from small donations from individual members.

This report is part of an ongoing Avaaz campaign to protect people and democracies from the dangers of disinformation and misinformation on social media. As part of that effort, Avaaz investigations revealed a disinformation network with half a billion views ahead of the European Union elections in 2019; prompted Facebook to take down a network reaching 1.7m people in Spain days before the 2019 national election; released a report on the fake news (reaching 105 million views) that fueled the Yellow Vests crisis in France; exposed a massive disinformation network (28 pages with 12.6 million interactions) during the Brazil presidential elections

in 2018; revealed the role anti-vaccination misinformation is having on reducing the vaccine rate in Brazil; and released a report on how YouTube is driving millions of people to watch climate misinformation videos.

Avaaz's work on disinformation is rooted in the firm belief that fake news proliferating on social media poses a grave threat to democracy, the health and well-being of communities, and the security of vulnerable people. Avaaz reports openly on its disinformation research so it can alert and educate social media platforms, regulators, and the public, and to help society advance smart solutions to defend the integrity of our elections and our democracies. You can find our reports and learn more about our work by visiting:

https://secure.avaaz.org/campaign/en/disinfo_hub/.

Annex

Table with the 20 most significant examples of misinformation content referenced in this report.

Fake News Post	Language	Facebook warning label showing on post?	Delay of fact check (where applicable)	Fact check (linked to article) * = Facebook Third Party Fact Checker
Gaza is the only place in the world free of the Coronavirus	Arabic	Yes (4 days after fact check)	2	Fatabyyano [*] فتابيانو
The hot air produced by hair dryers can prevent coronavirus	Arabic	Yes	6	Fatabyyano [*] فتابيانو
Hold your breath for 10 seconds to test yourself for coronavirus	Arabic	No	-	AFP [*]
UNICEF instructions to protect from coronavirus: gargle with salt water...	Arabic	Yes	11	AFP [*]
The coronavirus pandemic can be stopped with high doses of vitamin C	English	Yes	10	Lead Stories [*]
Coronavirus destroyed by chlorine dioxide	English	Yes	14	Factcheck.org [*]
Oregon oil proves effective against coronavirus	English	No	-	Washington Post
Black people are resistant to Coronavirus	English	Yes	1	AFP [*]
VP Pence urges people with coronavirus to go to the police	English	Yes	1	Lead Stories [*]

Garlic cures coronavirus	English	Yes	10	Factcheck.org*
Information regarding Coronavirus: temperatures above 26/27° kill the virus	French	Yes	4	Les Décodeurs du Monde*
Coronavirus was created and planned by the WHO and the Gates Foundation	Italian	No	-	Open
Use natural defenses against viruses and bacteria - staying home is useless	Italian	No	-	Open
Warning: coronavirus found on packages from China	Portuguese	Yes	27	AFP*
Vaccine that cures coronavirus in 3 hours is ready	Portuguese	Yes (10 days after fact check)	-	Agência Lupa*
793 deaths in Italy today, 232 children among them	Portuguese	Yes (1 day after fact check)	4	Aos Fatos*
Bolsonaro says that hydroxychloroquine has been proved as cure for coronavirus	Portuguese	No	-	Boatos
Doctor Mario Pesaresi reveals inhaling steam as a natural remedy for coronavirus	Spanish	No		Maldito Buló*
Avoid more deaths! Gargle with disinfectant	Spanish	Yes	1	AFP*
Coronavirus was created in the Institute of Virology in Wuhan	Spanish	Yes (13 days after fact check)	9	Maldito Buló*

Endnotes

- WHO, Munich Security Conference, 15 February 2020, <https://www.who.int/dg/speeches/detail/munich-security-conference> ❶
- CBS interview with Sheryl Sandberg: [Facebook is removing fake coronavirus news "quickly". COO Sheryl Sandberg says](#), Facebook Announcement on: [Combating COVID-19 Misinformation Across Our Apps](#), Mark Zuckerberg Press Briefing - March 18th, <https://about.fb.com/wp-content/uploads/2020/03/March-18-2020-Press-Call-Transcript.pdf> ❶
- See [Methodology and Data Set section](#) for detail. ❶
- For our study, we analysed only misinformation content that either had primary or direct fact checks - namely, fact checks that linked or referred to specific posts or articles). For those pieces of content that did not have primary or direct fact checks, we applied findings from primary fact checks to identical health claims made in those posts. ❶
- Our research period covered the actions and statements issued by Facebook between 16 January 2020 and 14 April 2020. However, based on conversations with representatives of the platform, we are encouraged by efforts the platform is experimenting with to improve the policies and processes that reflect some of the recommendations mentioned below. ❶
- Facebook, Combating COVID-19 Misinformation Across Our Apps, 25 March 2020, <https://about.fb.com/news/2020/03/combating-covid-19-misinformation/> ❶
- By moving to alert those who see misinformation, the platform is now doing more than any of the leading social media platforms to inform users about false content they have encountered and starts fixing many of the policy gaps highlighted above. ❶
- See [Methodology and Data Set section](#). ❶
- The original post was taken down after Avaaz shared it with Facebook https://www.facebook.com/permalink.php?story_fbid=2733027656745577&id=1555031211211900&_tn=-R ❶
- The original post was taken down after Avaaz shared it with Facebook <https://www.facebook.com/photo.php?fbid=2937680962950647&set=a.152802728105165&type=3&theater> ❶
- The original post was taken down after Avaaz shared it with Facebook <https://www.facebook.com/groups/499554396786126/permalink/3476737632401106> ❶
- <http://archive.ph/wip/d78wjl> ❶
- Lead Stories, Fact Check: VP Pence Did NOT Urge People With Coronavirus To Go To The Police, 2 March 2020, <https://leadstories.com/hoax-alert/2020/03/Fact-Check-VP-Pence-Did-NOT-Urge-People-With-Coronavirus-To-Go-To-The-Police.html> ❶
- https://www.facebook.com/Hayfaalhusseini/videos/255658768781237/?_tn=-R ❶
- The original post was taken down after Avaaz shared it with Facebook https://www.facebook.com/holistic8healing/photos/a.362593477866554/605012106958022/?type=3&_tn=-R ❶
- For the purposes of the "Correct the Record" academic study, referenced in full later in this report, we used the terms "disinformation", "misinformation", and "false news" interchangeably to mean verifiably false or misleading information that has the potential to cause public harm such as undermining democracy or public health. For the purposes of the investigation presented in this briefing, we relied on a definition of "misinformation" described in full in the Methodology and Data Set section below. ❶
- Facebook announcement, 25 March 2020: [Combating COVID-19 Misinformation Across Our Apps](#) ❶
- <http://archive.ph/S9n1z> ❶
- <http://archive.ph/hqKF5> ❶
- <http://archive.ph/OFOUj> ❶
- As available from CrowdTangle ❶
- Facebook announcement, 25 March 2020: [Combating COVID-19 Misinformation Across Our Apps](#) ❶
- Avaaz shared the list of misinformation posts that could potentially lead to imminent physical harm with Facebook on 8 April 2020 to help ensure that they would be speedily removed. In response, Facebook has, as of 14 April 2020, taken down 17 of the posts we flagged. ❶
- CBS News interview with Facebook's COO, Sheryl Sandberg, 18 March 2020, <https://www.cbsnews.com/news/facebook-coronavirus-fake-news-coo-sheryl-sandberg/> ❶
- "We remove coronavirus-related misinformation that could contribute to imminent physical harm. We've removed harmful misinformation since 2018, including false information about the measles in Samoa where it could have furthered an outbreak and rumors about the polio vaccine in Pakistan where it risked harm to health aid workers. Since January, we've applied this policy to misinformation about the coronavirus to remove posts that make false claims about cures, treatments, the availability of essential services or the location and severity of the outbreak. We regularly update the claims that we remove based on guidance from the WHO and other health authorities. For example, we recently started removing claims that physical distancing doesn't help prevent the spread of the coronavirus. We've also banned ads and commerce listings that imply a product guarantees a cure or prevents people from contracting COVID-19." Facebook, Combating COVID-19 Misinformation Across Our Apps, 25 March 2020, <https://about.fb.com/news/2020/03/combating-covid-19-misinformation/> ❶

26. The original post was taken down after Avaaz shared it with Facebook https://www.facebook.com/themegavitaminman/photos/a.154722701400928/1234263616780159/?type=3&__tn__=-R ❶
27. <http://archive.ph/7ZsU5> ❶
28. The original post was taken down after Avaaz shared it with Facebook https://www.facebook.com/themegavitaminman/photos/a.154722701400928/1234263616780159/?type=3&__tn__=-R ❶
29. Lead Stories, Fake News: Vitamin C Does NOT Slow Or Stop Coronavirus, 6 February 2020, <https://leadstories.com/hoax-alert/2020/02/Fake-News-Vitamin-C-Does-NOT-Slow-Or-Stop-Coronavirus.html> ❶
30. Facebook removed the original post after Avaaz shared the link with the company's team dedicated to enforcing its anti-misinformation policies on 8 April 2020. ❶
31. Facebook, Combating COVID-19 Misinformation Across Our Apps, 25 March 2020, <https://about.fb.com/news/2020/03/combating-covid-19-misinformation/> ❶
32. <http://archive.vn/C7ZaU> ❶
33. New York Times, 7 Answers to Questions About the Malaria Drug Trump Keeps Pushing, 9 April 2020, <https://www.nytimes.com/article/coronavirus-hydroxychloroquine-malaria.html> ❶
34. Vox, Trump's reckless promotion of hydroxychloroquine to fight coronavirus, explained, 26 March 2020, <https://www.vox.com/2020/3/26/21193912/trump-hydroxychloroquine-coronavirus-treatment> ❶
35. This video was also posted on YouTube and appears to have been taken down by the platform. Original YouTube link: <https://www.youtube.com/watch?v=t1oXg5GlySM> ❶
36. Facebook, Combating COVID-19 Misinformation Across Our Apps, 25 March 2020, <https://about.fb.com/news/2020/03/combating-covid-19-misinformation/> ❶
37. <http://archive.is/njLDZ> ❶
38. Facebook, Combating COVID-19 Misinformation Across Our Apps, 25 March 2020, <https://about.fb.com/news/2020/03/combating-covid-19-misinformation/> ❶
39. The original post was taken down after Avaaz shared it with Facebook https://www.facebook.com/permalink.php?story_fbid=2733027656745577&id=1555031211211900&__tn__=-R ❶
40. Avaaz shared the list of misinformation posts that could potentially lead to imminent physical harm with Facebook on 8 April 2020 to help ensure that they would be speedily removed. In response, Facebook has, as of 14 April 2020, taken down 17 of the posts we flagged. ❶
41. Facebook, Combating COVID-19 Misinformation Across Our Apps, 25 March 2020, <https://about.fb.com/news/2020/03/combating-covid-19-misinformation/> ❶
42. https://www.facebook.com/Hayfaalhusseini/videos/255658768781237/?__tn__=-R ❶
43. Facebook, Combating COVID-19 Misinformation Across Our Apps, 25 March 2020, <https://about.fb.com/news/2020/03/combating-covid-19-misinformation/> ❶
44. <https://archive.md/QJOsV> ❶
45. The original post was taken down after Avaaz shared it with Facebook <https://www.facebook.com/groups/110439175991313/permalink/1060218944346660> ❶
46. CBS interview with Sheryl Sandberg: Facebook is removing fake coronavirus news "quickly", COO Sheryl Sandberg says, Facebook Announcement on: [Combating COVID-19 Misinformation Across Our Apps](https://about.fb.com/wp-content/uploads/2020/03/March-18-2020-Press-Call-Transcript.pdf), Mark Zuckerberg Press Briefing - March 18th, <https://about.fb.com/wp-content/uploads/2020/03/March-18-2020-Press-Call-Transcript.pdf> ❶
47. Facebook, Combating COVID-19 Misinformation Across Our Apps, 25 March 2020, <https://about.fb.com/news/2020/03/combating-covid-19-misinformation/> ❶
48. Associated Press, Ability to hold your breath 10 seconds is not a test for coronavirus, 12 March 2020, <https://apnews.com/article:Content:8635070312> ❶
49. Stanford Health Center, Novel Coronavirus (COVID-19) Resource Center, <https://stanfordhealthcare.org/stanford-health-care-now/2020/novel-coronavirus.html> ❶
50. <http://archive.vn/vP4fB> ❶
51. For the purposes of the "Correct the Record" academic study, we used the terms "disinformation", "misinformation", and "false news" interchangeably to mean verifiably false or misleading information that has the potential to cause public harm, for example by undermining democracy or public health. For the purposes of the investigation presented in this briefing, we relied on a definition of "misinformation" described in full in the Methodology & Data section below. ❶
52. **Note:** [This white paper](#) explains the findings of the research conducted by Dr. Wood and Dr. Porter. The academic study will be published independently by the authors. ❶
53. Visit Factbook, <https://factbook.org/> - to see how the hyper-realistic model used to mimic Facebook in the study. ❶
54. See https://secure.avaaz.org/act/media.php?press_id=932 ❶
55. Spanish
[Chequeado](#) (Argentina)
[Maldito Buló](#) (Spain)
[AFP Factual](#) (various Spanish speaking countries)
Portuguese
[AFP Checamos](#)
E farsas
[Boatos](#)
[Política Estadão](#)
[Aos Fatos](#)
[Lupa](#) (Brazil)
[Observador](#) (Portugal)
French
[AFP France](#)
[Le Monde](#) (France)
Italian
[Open](#)
[Butac](#)
[Bufale](#) (Italy)
Arabic
[Fataby, yano](#)
[AFP MENA](#) (MENA)
English
[Full Fact](#) (UK)
[Washington Post](#)
[Factcheck](#)
[Politifact](#)
[Lead Stories](#)
[Health Feedback](#)
[AFP United States](#)
[Snopes](#) (USA) ❶
56. <https://sciencefeedback.co> ❶
57. WHO, Global Vaccine Safety - Health Feedback, https://www.who.int/vaccine_safety/initiative/communication/network/healthfeedback/en/ ❶
58. For videos we include the views reported by Facebook. Our misinformation content selection contained 14 examples of false video content that have accumulated a total of 47,327,476 views. We followed Facebook's definition of what constitutes a "view," that is, a viewing experience that lasts 3 seconds or longer. For posts containing only text and image content the platform displays only the number of shares and other interactions, such as likes or comments. Therefore, in order to estimate viewership for text and image content we designed a metric based on the publicly available statistics of the Facebook videos analysed in our study, taking into account the total number of video views and dividing it by the total number of video shares. Facebook reports a "video view" only after 3 seconds, while an image or text can be considered as "viewed" and having an actual impact in less than 3 seconds (e.g. just scrolling by an image of a bleeding woman, police brutality or

claims of censorship). Therefore, the estimation of total views in this study is very likely lower than the actual total viewership of all content and therefore is to be seen as a conservative figure. ❶

59. 1. Fact checking articles shown as "Related Articles" below the post, 2. a pop-up notification titled "False information in this post" linking the fact checking article and 3. a gray overlay titled "False Information Checked by Independent Fact Checkers" linking to (a) fact checking article(s) ❷