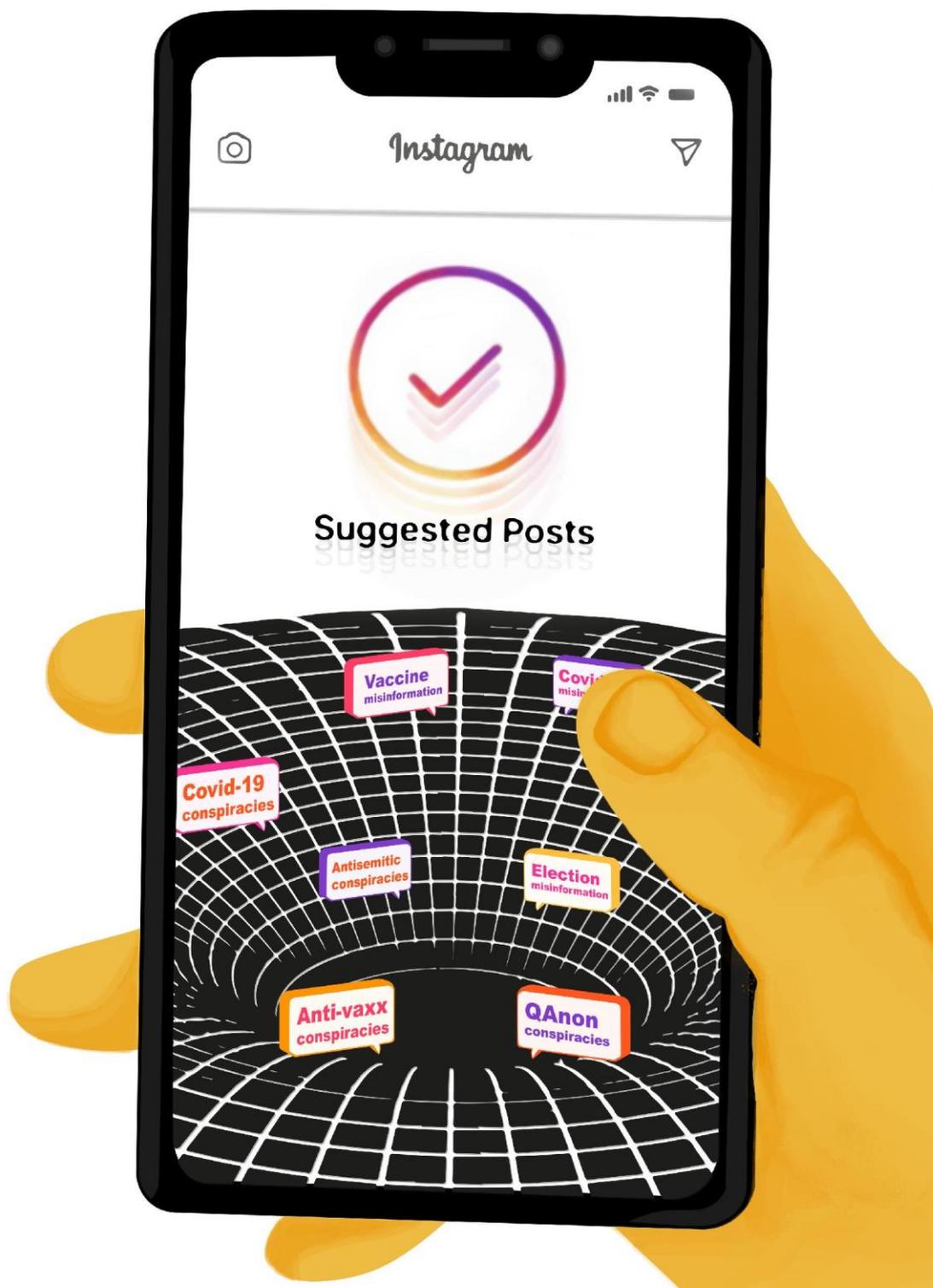


# MALGORITHM

HOW INSTAGRAM'S ALGORITHM PUBLISHES MISINFORMATION AND HATE TO MILLIONS DURING A PANDEMIC



# Contents

Introduction ..... 4

Youth Against Misinformation Foreword ..... 6

Executive Summary..... 8

Instagram recommended 104 posts containing misinformation ..... 9

Methodology .....11

Phase 1: Recommendations from Instagram’s “Explore” feature .....13

    List 1: Anti-vaxxers .....14

    List 2: Wellness influencers linked to anti-vaxx..... 19

    List 3: Health authorities .....24

Phase 2: Recommendations from Instagram’s “Suggested Posts” feature .....25

    List 4: Anti-vaxxers extended..... 26

    List 5: Far-right and wellness..... 30

    List 6: QAnon and health authorities .....32

Recommendations ..... 37

Appendix: Youth Leadership .....39

Appendix: Misinformation recorded ..... 40



The Center for Countering Digital Hate is a not-for-profit NGO that seeks to disrupt the architecture of online hate and misinformation.

Digital technology has changed forever the way we communicate, build relationships, share knowledge, set social standards, and negotiate and assert our society's values.

Digital spaces have been colonised and their unique dynamics exploited by fringe movements that instrumentalise hate and misinformation. These movements are opportunistic, agile and confident in exerting influence and persuading people.

Over time these actors, advocating diverse causes - from anti-feminism to ethnic nationalism to denial of scientific consensus - have formed a Digital Counter Enlightenment. Their trolling, disinformation and skilled advocacy of their causes has resocialised the offline world for the worse.

The Center's work combines both analysis and active disruption of these networks. CCDH's solutions seek to increase the economic, political and social costs of all parts of the infrastructure - the actors, systems and culture - that support, and often profit from hate and misinformation.



Young people have the power to solve the world's biggest challenges. Restless Development is a global agency that works with young people so that they can lead in solving those challenges.

Whether that means supporting communities to end child marriage or prevent HIV, Restless Development works with young people to change their lives and the lives of people in their communities.

Restless Development works with young people in 74 countries around the world and is led by nine Hubs globally. The agency has been working with young people since 1985 and its work is led by thousands of young people every year.

## **Youth Against Misinformation**

This initiative is being coordinated and supported by Restless Development and the Center for Countering Digital Hate. It aims: to monitor and root out dangerous COVID-19 misinformation online; and build the skillset and agency of young people; to understand how misinformation is spread online, identify fake news; and develop strategies to tackle it. Volunteers are trained in how to deal with harmful online misinformation, based in part on the principles of CCDH's Don't Spread the Virus initiative.

# Introduction

When it launched in 2010, Instagram – the social media giant acquired by Facebook in 2012 – was pitched as a simple and fun app for sharing photos with friends and family. Since then, it has become a networked behemoth profiting from the promotion of dangerous misinformation and conspiracy theories.

Malgorithm, the latest report from the Center for Countering Digital Hate, shows how the Instagram algorithm has been actively pushing radicalising, extremist misinformation to users. Users are being encouraged to view radical material, and then, once hooked, cross-fertilized with content from other limbs of the radical worldview. If a user follows anti-vaxxers, they are fed QAnon conspiracism and antisemitic hate; if they engage with conspiracies, they are fed electoral and anti-vaxx misinformation.

This is a purposeful tactic. Social media companies constantly seek to maximize user engagement. Quite simply, more user time on Instagram means more ad revenue. This is why in August 2020, to extend engagement, Instagram added unsolicited content to users' streams. Once a user exhausts the latest content from accounts they follow, Instagram's algorithms [present](#) new content as "an extension of your feed." "Machine learning" algorithms [identify](#) users' potential interests based on their data and habits, then find high-engagement content in the same ilk, and injects it into users' feeds. Previous [research](#) shows that misinformation is shared and gets more engagement than truths on social media. Worse still, high engagement numbers [increases the likelihood](#) neutral observers engage with the content. For Instagram and its algorithms, a click is a win, no matter the content.

The removal of a few posts in one or another limb of this Digital Counter-Enlightenment is insufficient. By funnelling users towards extreme content, Instagram's recommendation algorithm sends users down paths of radicalization that are increasingly difficult to escape. We conclude that the algorithm must be suspended as it poses an imminent critical threat to lives and national security.

For months, the World Health Organization has warned governments and organizations about our current "infodemic" and how the proliferation of misinformation damages national and global biosecurity. In 2020, the CCDH tracked the movement of misinformation through [The Anti-Vaxx Industry](#), which demonstrated how anti-vaccine campaigners, entrepreneurs, and their social media misinformation networks provide Big Tech with audiences valued at more than \$1 billion annually. Anti-vaxxers monitored in CCDH's [most recent study](#) had amassed more than 59 million followers.

When Instagram chose to promote this content to its users using a revenue-maximizing algorithm, its role as a publisher of user-generated content was unmasked. It cannot claim to be a neutral platform. Facebook owns, controls, and profits from the Instagram algorithm, which in turn is amplifying anti-vaxx, Covid-

19 misinformation, and QAnon conspiracies. As with any publisher or producer of dangerous goods, Facebook and its executives must be held accountable.

People are being put in danger by their product and the minimum we can expect is that they make an emergency recall to fix the problem. The social media behemoth must suspend the algorithm until it has proven it can prevent the promotion of harmful misinformation. Should social media companies continue their pattern of negligence, governments must use every power - including new legislation, fines and criminal prosecutions - to stop the harms being created. Lies cost lives. There is no time to waste.

**Imran Ahmed**  
**CEO, CCDH**

# Youth Against Misinformation Foreword

The Coronavirus pandemic has affected people's lives in dramatic ways. Being under lengthy social restrictions has made us more reliant on online platforms for information and communication with others. This has subsequently left many, especially young people who make up the bulk of users and are the most active on social media, highly vulnerable to the growing deluge of online misinformation.

In April 2020, a group of young people from across the UK met online for the first time. Known later as Youth Against Misinformation (YAM), this was the beginning of a nine-month initiative that saw young people enter the vanguard of the fight against online hate and fake news.

The YAM initiative was founded in partnership by global youth agency Restless Development and the Center for Countering Digital Hate (CCDH) to empower young people to take a stand against online spreaders of misinformation and to challenge the social media companies who fail to tackle it. Youth volunteers worked with CCDH for several months to research and produce a number of reports on health and Covid-19 misinformation, which have highlighted [social media's failure](#) to tackle the growing infodemic and the dangers it poses. The rationale was to engage young people to make an active difference, for our digital literacy makes us well-placed to help tackle the problem. As digital natives, young people can help create safer online spaces. The research conducted by YAM volunteers helped expose a vicious online industry, emboldened by the pandemic, which is committed to spreading health misinformation across social media. Social media companies failed to remove this material and their algorithms have now been found to spread it further by actively suggesting it to users.

The shocking consequences of this are plain to see, with anti-vaxx conspiracies [contributing to growing rates of vaccine hesitancy](#) and a loss of trust in our health authorities at the very time we need them most. As many of us look desperately towards an end to the Covid-19 crisis, the sacrifices being made to stop this virus are being undermined by those who preside over these crucial public spaces and who are allowing misinformation to proliferate.

This latest report focuses on Instagram, one of the most popular platforms among young people, and its new content suggesting functions. The results are shocking as volunteers were again exposed to dangerous content ranging from anti-Semitic conspiracy theories to anti-vaxx propaganda.

We wholeheartedly support CCDH's conclusion that Instagram's new algorithm poses a very real threat to its users. Promoting lies and hateful content to users isolating in a pandemic has the potential to seriously impact their mental health and, in the event that they don't take the Covid vaccine due to misinformation suggested to them, their physical health as well.

Social media companies continue to permit misinformation to spread unchecked on their platforms, exposing young people and our wider societies to a wealth of dangerous and radicalising propaganda. We hope this research provides the wake-up call which is long overdue.

We would like to thank the incredible staff at Restless Development, CCDH and our amazing Volunteer Team Leaders for their support throughout the research and beyond the programme.

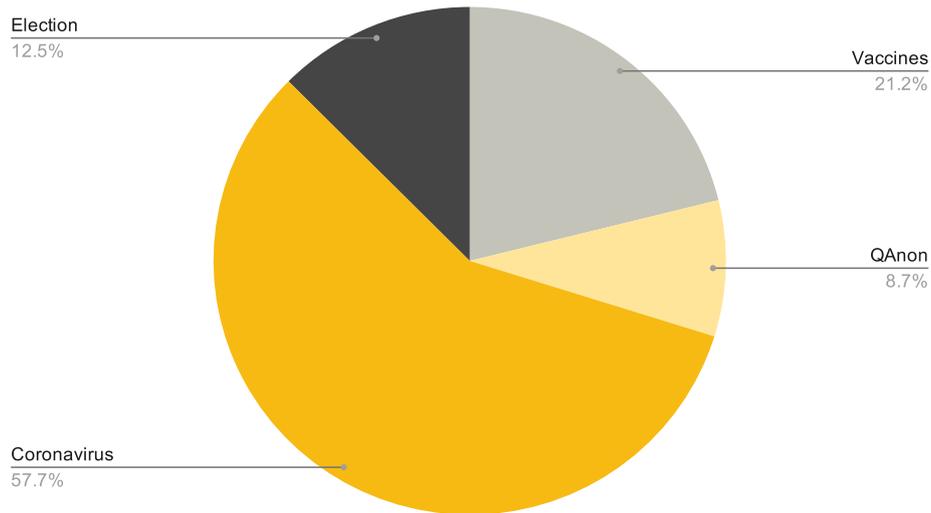
**Youth Against Misinformation Youth Volunteers**

# Executive Summary

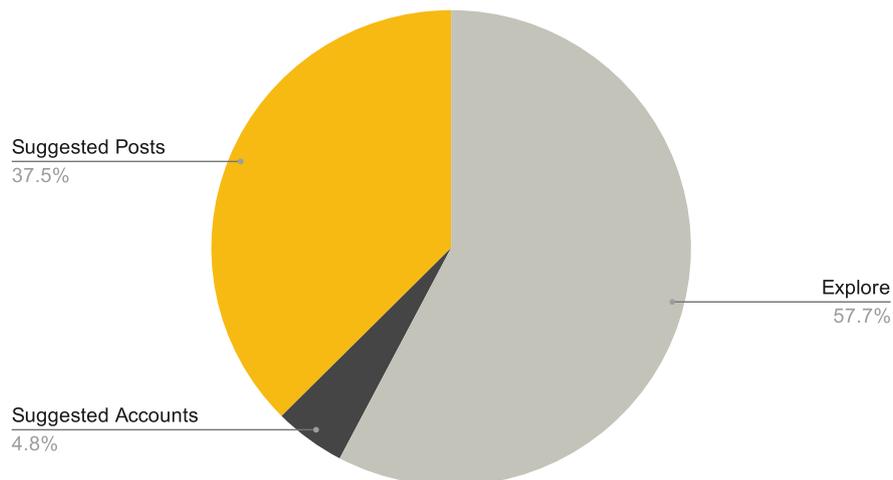
1. To investigate the role of Instagram’s algorithm in publishing misinformation, researchers established new Instagram profiles following prescribed lists of accounts and recorded the misinformation that Instagram’s algorithm recommended to them through the platform’s “Explore” and “Suggested Post” features.
2. These new Instagram profiles followed a range of accounts, from health authorities to anti-vaxxers, and recorded one post containing misinformation every week for the duration of our research, with 104 posts recorded in total. Over half the misinformation recommended by Instagram was about Covid, while a fifth was about vaccines and a tenth contained misinformation about the US elections.
3. Users who follow leading anti-vaxxers receive recommendations for antisemitic content, QAnon conspiracy theories and Covid misinformation. Instagram displayed a number of these recommendations next to adverts for brands.
4. Followers of wellness influencers with links to the anti-vaccine movement receive recommendations for harder anti-vaccine content from leading anti-vaxxers who had been granted “verified” status by Instagram despite being flagged by experts. They also receive recommendations for Covid misinformation and conspiracies.
5. Instagram recommends posts containing misinformation about Covid, vaccines and elections despite applying labels to them that show it has correctly identified their subject matter.
6. Users receive recommendations for posts promoting conspiracies, including the QAnon conspiracy theory. Followers of QAnon accounts receive recommendations for Covid and vaccine misinformation, and Instagram suggests that they follow other QAnon accounts including backups for existing accounts.
7. Users following a mix of health authorities such as the CDC and conspiracy accounts received recommendations for further misinformation about Covid and vaccines. Only users who strictly followed only health authorities did not receive recommendations for misinformation.
8. Instagram must suspend its algorithm until it can show that it is no longer promoting dangerous misinformation. Instagram can improve its algorithm by excluding posts about Covid or vaccines from recommendations, maintaining a blacklist of accounts known to spread misinformation, and limiting the number of suggested posts that a user is exposed to.
9. In addition, Instagram can make its users more resistant to misinformation by preventing spreaders of misinformation from gaining “verified” status, introducing warnings on algorithmically published content, and by “inoculating” users by warning them about the misinformation they might encounter.
10. Instagram should remedy the damage already done by its algorithm by displaying corrective posts to users already exposed to misinformation, and by donating its profits from adverts placed next to suggested posts containing misinformation to health organisations.

# Instagram recommended 104 posts containing misinformation

More than half contained Covid misinformation about Covid, while a fifth contained misinformation about vaccines.



57% of recommendations were taken from Instagram's Explore page, while 37% were taken from its new Suggested Posts feature.



With Instagram's help, these posts with misinformation gained **over 233,000 likes**.

# Instagram is recommending misinformation and hate to its users

Users following these kinds of accounts

## Wellness

Wellness influencers adjacent to anti-vaxxers

## Anti-vaxxers

Leading anti-vaxxer accounts

## Anti-vaxxers+

Extended list of leading anti-vaxxers

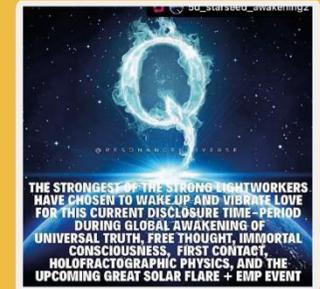
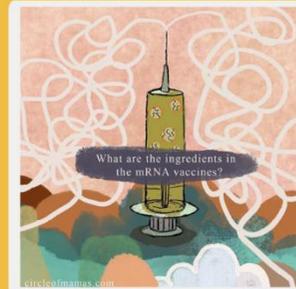
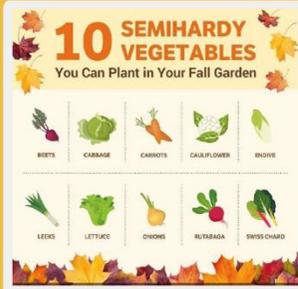
## Far-Right

Far-right and militia groups

## QAnon

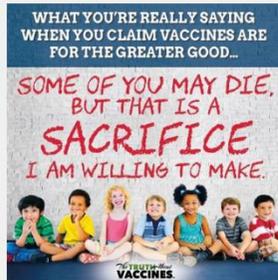
QAnon accounts and health authorities

With this typical content



Are recommended this misinformation by Instagram

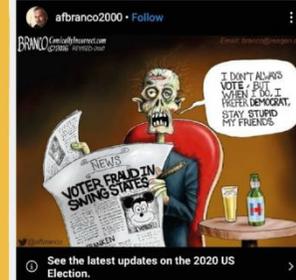
63% Covid  
37% Anti-vaxx



63% Covid  
30% Anti-vaxx



52% Covid  
26% US Election  
22% Anti-vaxx



67% Covid  
33% Anti-vaxx



54% Covid  
25% US Election  
21% QAnon



Key points

- ▶ Harder anti-vaxx content from "verified" accounts
- ▶ Conspiracies about Covid and other subjects

- ▶ Antisemitic conspiracies
- ▶ Covid and vaccine misinformation
- ▶ QAnon content

- ▶ US Election misinformation
- ▶ Covid and vaccine misinformation

- ▶ Covid and vaccine misinformation

- ▶ Covid, vaccine and election misinformation recommended despite also following trustworthy health accounts

# Methodology

This report began with a simple question: what sort of content does Instagram recommend to users?

In order to explore this question, and how it relates to the accounts we follow and the posts we like, we established new Instagram profiles to record Instagram's recommendations from the point of view of an average user.

## How Instagram recommends content to its users

Facebook has used its AI blog to explain how it generates recommendations served to users through Instagram's "Explore" feature.<sup>1</sup>

In outline, Facebook says that it starts by identifying groups of accounts with similar interests, for example calligraphy. It then identifies a user's interests based on their interactions with other posts and accounts on Instagram, and draws up a personalised long list of potential post recommendations from accounts concerning the same interests. Finally, Facebook ranks this long-list analyses which posts the user is most likely to engage with, and presents the top-ranked posts from a variety of accounts.

Importantly, Facebook promises its methods "filter out content we can identify as not being eligible to be recommended before we build out eligible inventory for each person. In addition to blocking likely policy-violating content and misinformation, we leverage ML systems that help detect and filter content like spam."

This is the complex process, only presented in summary above, that governs which posts are recommended to users on the "Explore" tab of the Instagram app. Such processes are often referred to as "the algorithm" used by Facebook to determine which content is presented to which users, and in which order. In truth, Facebook uses a variety of such techniques across its different platforms, apps and features.

What all of these features have in common is that they use data about what users like, share, follow or post to present them with present them with further content that they are likely to find engaging. This is of critical importance to businesses like Facebook and Instagram: more engaging content means users are on their platforms for longer, during which time they can be presented with more adverts, driving up revenue.

## How Instagram presents users with recommended content

At present there are three main ways in which Instagram presents users with content based on algorithmic analysis of their habits:

- **Suggested Accounts:** Instagram will occasionally present users with a carousel of suggested accounts to follow between posts in the user's main feed or as a short list inserted between "Stories".
- **Explore:** Instagram's "Explore" feature displays suggested posts, videos and stories below the app's search bar. Instagram states that in Explore, "posts are selected automatically based on things like the people you follow or the posts you like."<sup>2</sup> Posts in Explore carry a short explanation of why the post was suggested below the main image or video, such as "based on posts you like".

- **Suggested Posts:** Since August 2020, users reaching the end of their Instagram news feed will be presented suggested posts from users they do not follow.<sup>3</sup> Instagram states that “these suggestions are based on posts from accounts like the ones you follow and posts similar to the ones you like or save.”<sup>4</sup> Like Explore, posts presented with this feature carry a short explanation of why they were suggested.

### **How we recorded Instagram’s recommendations**

To generate these recommendations, we worked with volunteers supported by Restless Development who set up fresh Instagram profiles linked to fresh Gmail addresses, and followed a list of accounts selected based on a particular theme, such as wellness. A total of six different lists of accounts were used throughout our evidence gathering phase, and these are detailed in the sections below alongside the suggested posts they generated. The lists were composed of 73 different accounts in total. Of those accounts, 63 are linked to the spread of misinformation or hate and have over 9.9 million followers, making it likely that our findings apply to millions of Instagram users.

Volunteers were set the aim of logging into their new Instagram accounts daily, both in the mobile app and in their browsers. Instagram’s newest “Suggested Posts” feature does not trigger for new accounts without a history of interacting with posts, so our volunteers were asked to scroll through their feeds and the app’s “Explore” section and “like” posts at random to develop the required data to generate suggested content.

Volunteers were instructed to screenshot content containing misinformation that was recommended to their new Instagram accounts, recording details such as the date it was recommended and the account that had posted it. Volunteers were trained by our staff to identify common forms of misinformation and directed to draw on online resources such as fact-checkers. Volunteers recorded recommendations over 64 days, from 14 September to 16 November 2020.

# Phase 1: Recommendations from Instagram’s “Explore” feature

In the first phase of our research beginning on 14 September 2020, volunteers created new accounts and followed a list of ten accounts from one of the following categories:

- Anti-vaxx content creators
- Wellness influencers linked to anti-vaxx
- Health authorities

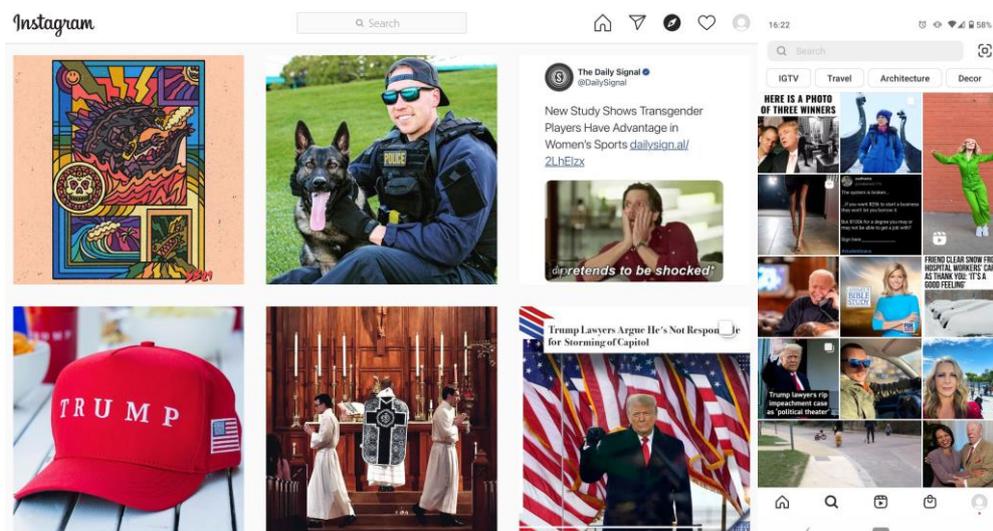
Our volunteers discovered that following a list of just ten accounts failed to trigger Instagram’s newest “Suggested Posts” feature at the bottom of their news feeds. This led to the development of longer and more varied lists for use in Phase 2 of our research which succeeded in triggering recommendations using this feature. As such, all recommendations from Phase 1 were taken from algorithmically generated recommendations on Instagram’s “Explore” feature.

## How Instagram’s Explore feature works

Instagram’s “Explore” feature is described by the platform as “where people go when they want to see photos and videos related to their interests from accounts they may not already follow.” According to Instagram, more than half of its users use Explore every month. Since June 2019, Instagram has inserted adverts between the recommended posts that users encounter when scrolling through Explore.<sup>5</sup>

Users can access Explore either through a dedicated button in Instagram’s web interface symbolised by a compass, or by pressing the search button in the Instagram app, symbolised by a magnifying glass.

The below examples of the Explore page as it appears in the browser and in the Instagram app are taken from an account following a range of far-right profiles.



# List 1: Anti-vaxxers

In this phase, four volunteers were instructed to follow a list of accounts that post original anti-vaccine content or have given a platform to leading anti-vaxxers. All of them were selected because they are relatively professional accounts, some of which have even been given verified status by Instagram.

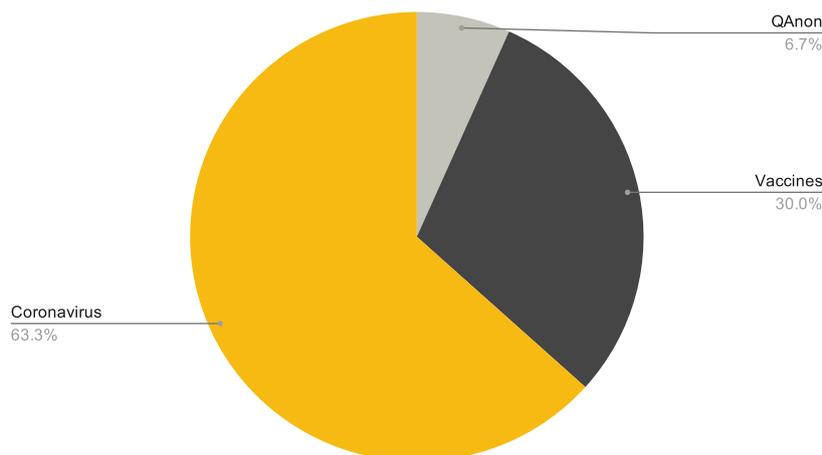
At the time of our research, these accounts had 6.5 million followers on Instagram, making it likely that our findings apply to millions of Instagram users.

1. [Robert F. Kennedy Jr.](#) (verified)
2. [Children's Health Defense](#)
3. [London Real](#) (verified)
4. [Weston A. Price Foundation](#)
5. [NVIC.org](#)
6. [Physicians for Informed Consent](#)
7. [Highwire with Del Bigtree](#)
8. [Dr Rashid A Buttar](#)
9. [Patrick Bet-David](#)
10. [Valuetainment](#)

## Recommendations recorded by our volunteers

Volunteers following the above list of accounts recorded a total of 30 recommendations from Instagram’s algorithm that contained misinformation or hatred between 22 September and 17 October, all from Instagram’s Explore page.

Of those 30 recommendations, 19 primarily contained misinformation about Covid, while 9 concerned misinformation about vaccines and two promoted elements of the QAnon conspiracy theory. Two of the posts volunteers identified contained antisemitic imagery.

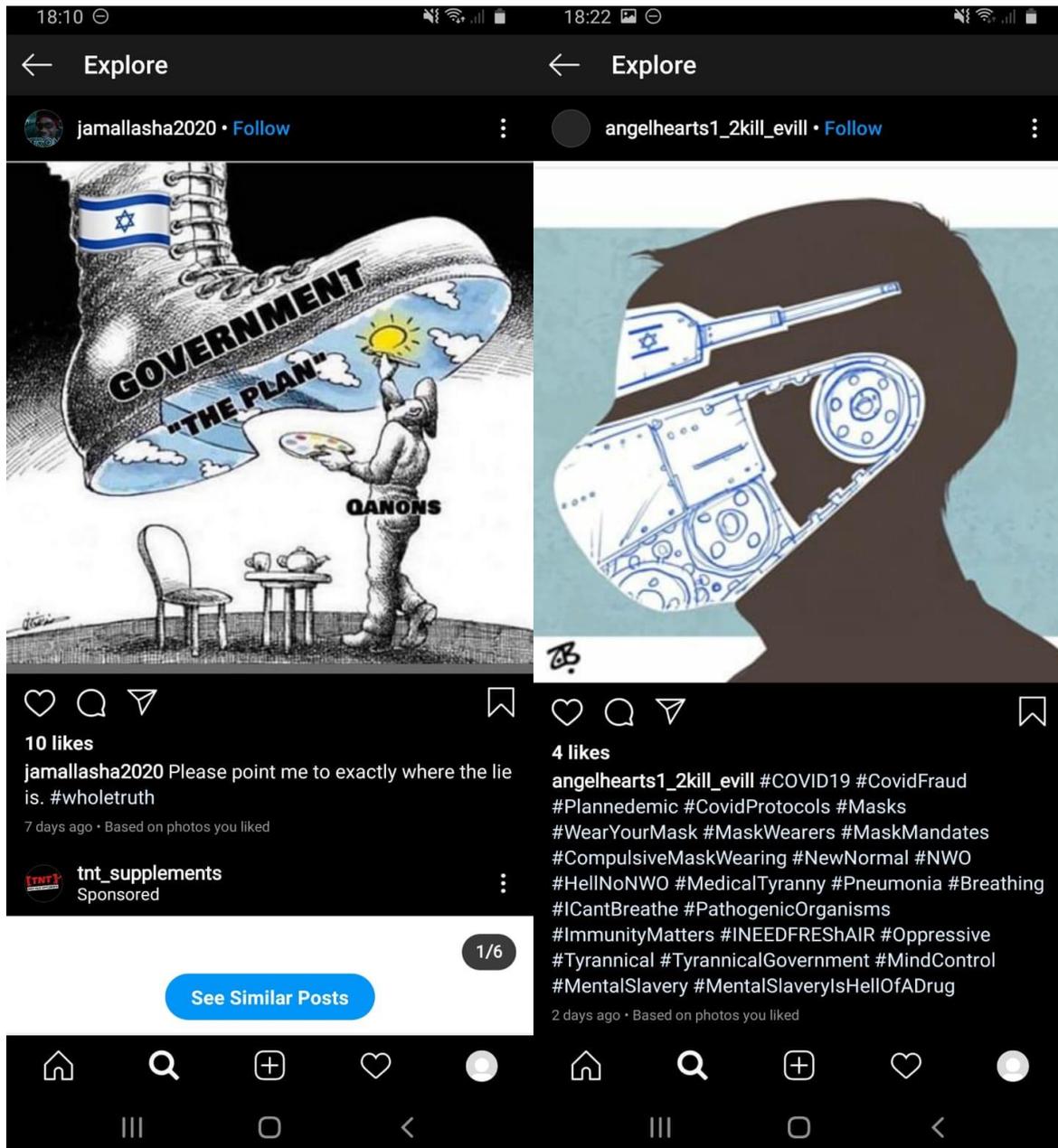


These findings from our volunteers highlighted four key issues that are described in further detail on the following pages:

- Recommendations of antisemitic content
- Recommendations of Covid misinformation to followers of anti-vaxxers
- Recommendations of QAnon conspiracy content to followers of anti-vaxxers
- Adverts displayed next to recommendations for anti-vaccine misinformation

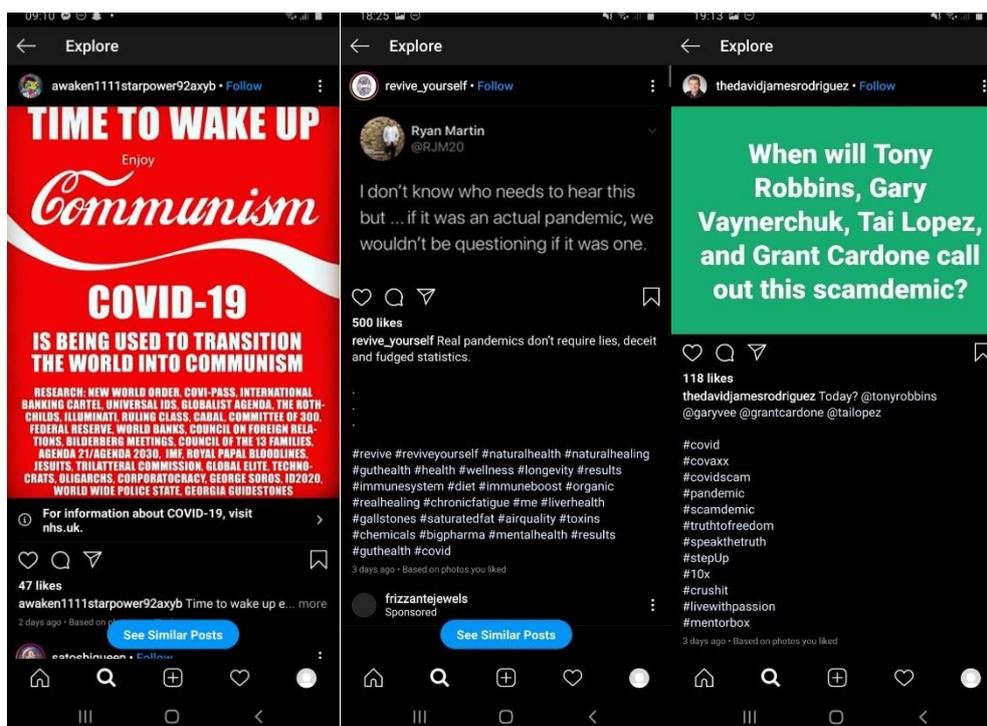
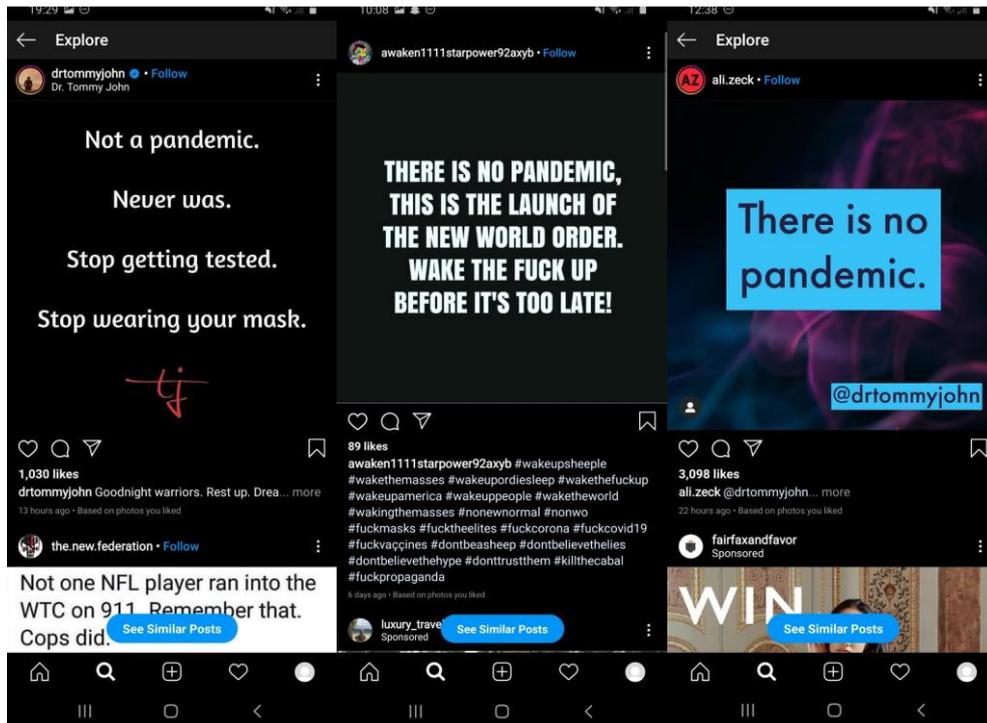
### Followers of anti-vaccine accounts are being recommended antisemitic content

Volunteers following this list received two recommendations from Instagram for posts promoting antisemitic conspiracy theories. The first uses the Israeli flag to imply that Jews control the government, while the second, from a different account, uses the flag to link Israel to mask wearing, a “Plannedemic” and the “NWO”, a supposed “New World Order” directing global events.



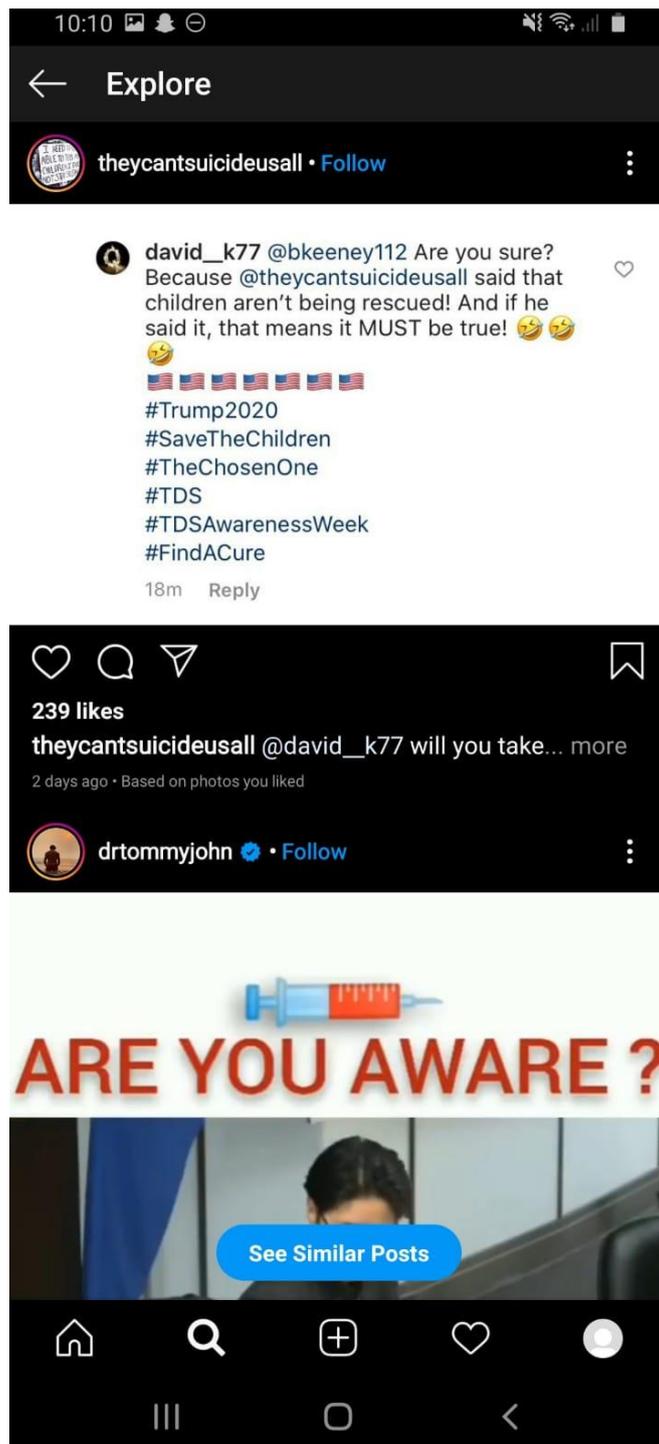
## Followers of anti-vaccine accounts are being recommended Covid misinformation

Volunteers following this list were also recommended a range of Covid misinformation. Numerous posts claimed that “there is no pandemic”, while other examples on people to “stop getting tested” and “stop wearing your mask”. Others mixed Covid denial with grand conspiracy theories, with one post claiming “there is no pandemic, this is the launch of the New World Order”.



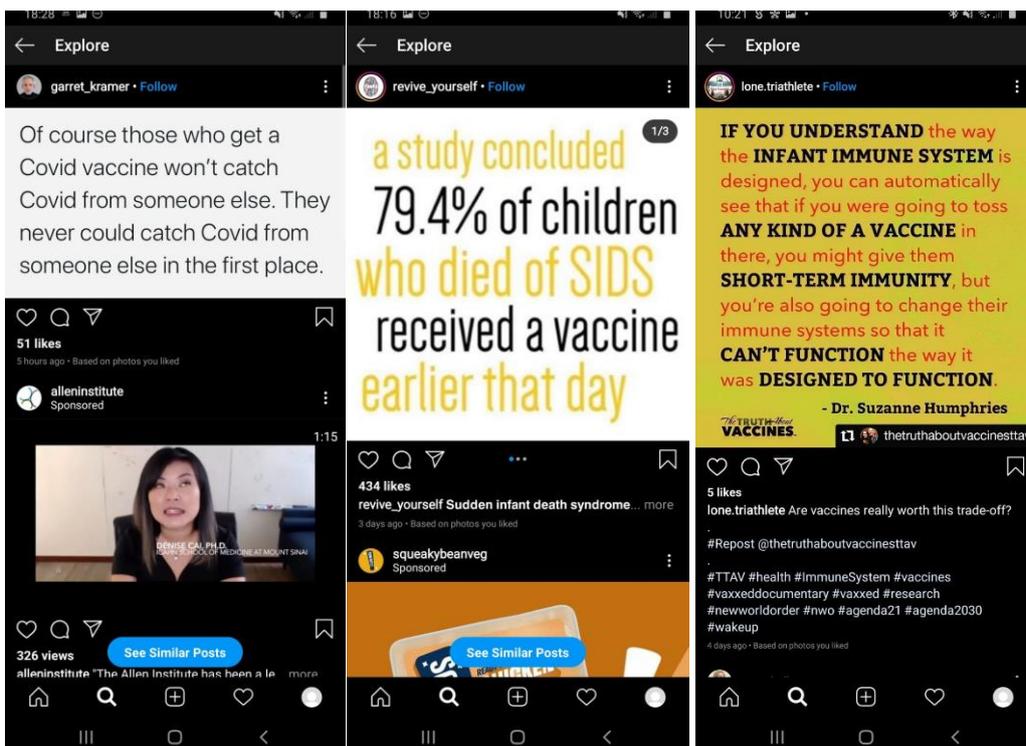
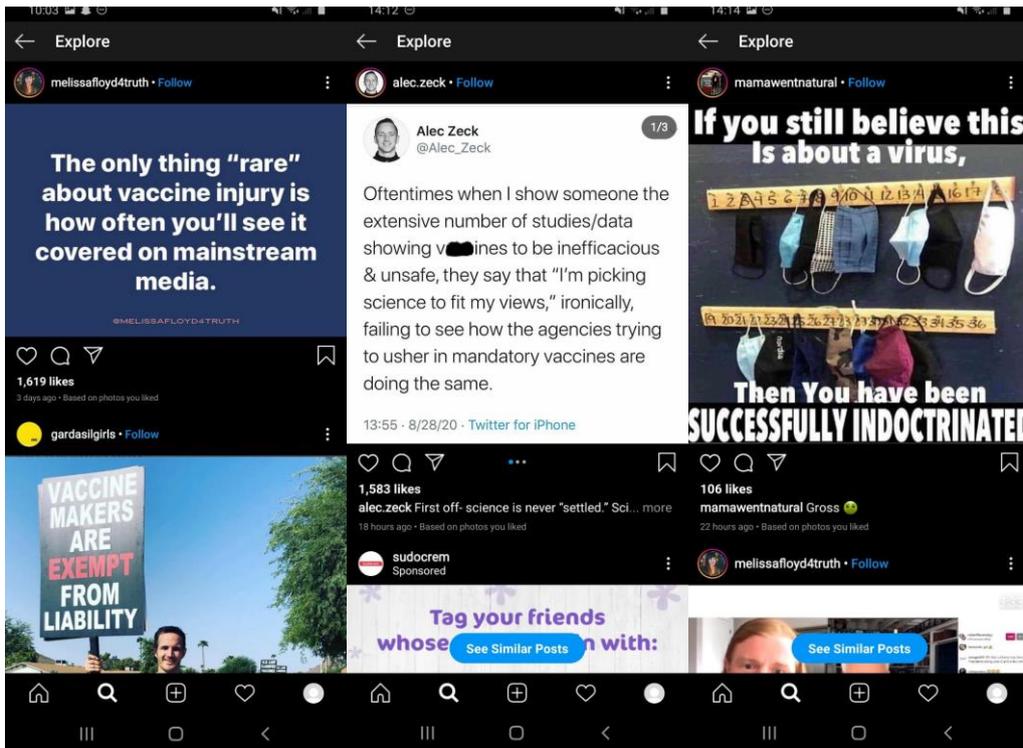
### Followers of anti-vaccine accounts are being recommended QAnon conspiracy content

Volunteers found one example of Instagram recommending a post promoting the QAnon conspiracy theory. This recommendation was made on 22 September 2020, a month after Facebook said that it was “taking action against Facebook Pages, Groups and Instagram accounts tied to offline anarchist groups that support violent acts amidst protests, US-based militia organizations and QAnon.”<sup>6</sup>



## Followers of anti-vaccine accounts are being recommended further anti-vaccine content presented next to adverts

Volunteers found numerous examples of recommended posts containing anti-vaccine and Covid misinformation from accounts they were not following. A number of these posts were presented next to adverts, highlighting Instagram’s use of algorithmic recommendations to increase the time users spend on the platform and boost ad revenues.



## List 2: Wellness influencers linked to anti-vaxx

Four volunteers were instructed to follow a list of accounts comprised of wellness influencers and “alternative health” entrepreneurs who have expressed anti-vaccine views or are linked to prominent anti-vaxxers. The aim of using this list was to investigate what recommendations Instagram serves to people who follow wellness accounts that occasionally post relatively “soft” anti-vaccine content.

CCDH’s previous reports on the anti-vaccine movement, *The Anti-Vaxx Industry* and *The Anti-Vaxx Playbook*, demonstrate that the online anti-vaccine movement has significant influence amongst wellness influencers and “alternative health” entrepreneurs.<sup>7</sup>

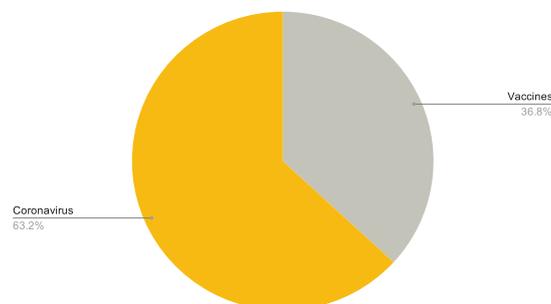
Together these accounts currently have 1.6 million followers on Instagram, making it likely that our findings apply to hundreds of thousands of Instagram users.

1. [David Wolfe](#) (verified)
2. [Zach Bush](#)
3. [Ben Lynch](#)
4. [Global Healing](#)
5. [Wellness Force](#)
6. [ANH International](#)
7. [Bruce H. Lipton, Ph.D.](#) (verified)
8. [Dr. Joseph Mercola](#) (verified)
9. [Erin Elizabeth](#)
10. [Organic Consumers Association](#) (verified)

### Recommendations recorded by our volunteers

Volunteers following the above list of accounts recorded a total of 19 recommendations from Instagram’s algorithm that contained misinformation or hatred between 17 September and 12 October, all from Instagram’s Explore page.

Of those 19 recommendations, 12 primarily contained misinformation about Covid, while 7 concerned misinformation about vaccines.

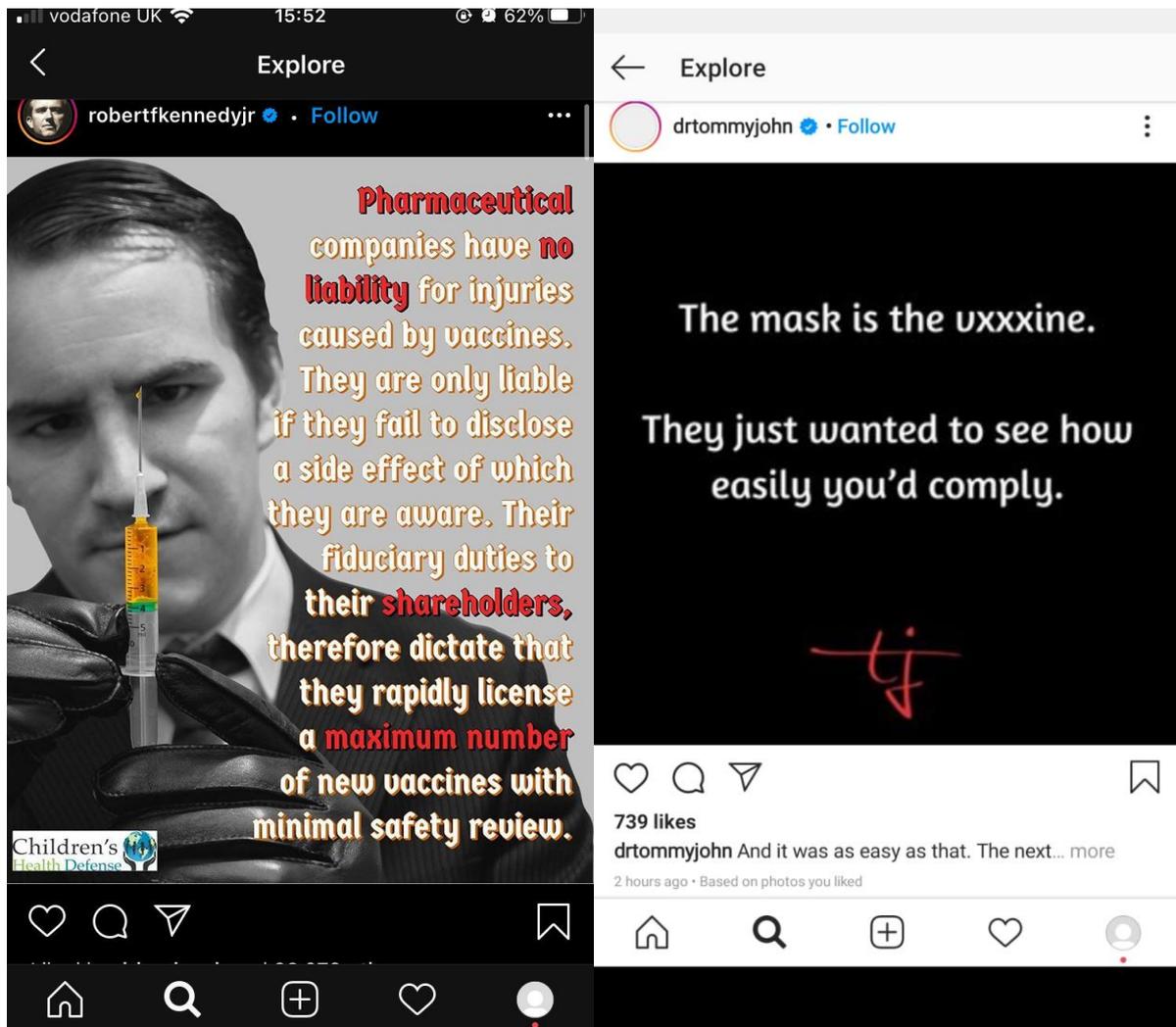


These findings from our volunteers highlighted three key issues that are described in further detail on the following pages:

- Recommendations of harder anti-vaccine content from “verified” anti-vaxxers
- Recommendations for posts by anti-vaxxers repeatedly flagged to Instagram
- Recommendations for conspiracies about Covid and other subjects

### Followers of wellness influencers linked to anti-vaxx are being recommended harder anti-vaccine content from accounts that Instagram has granted “verified” status

Many of the recommended posts received by volunteers following this list contained clear anti-vaccine misinformation. Some of them were from accounts with “verified” status, which may give users the false impression that Instagram has endorsed the poster or judged the content of their posts to be accurate.

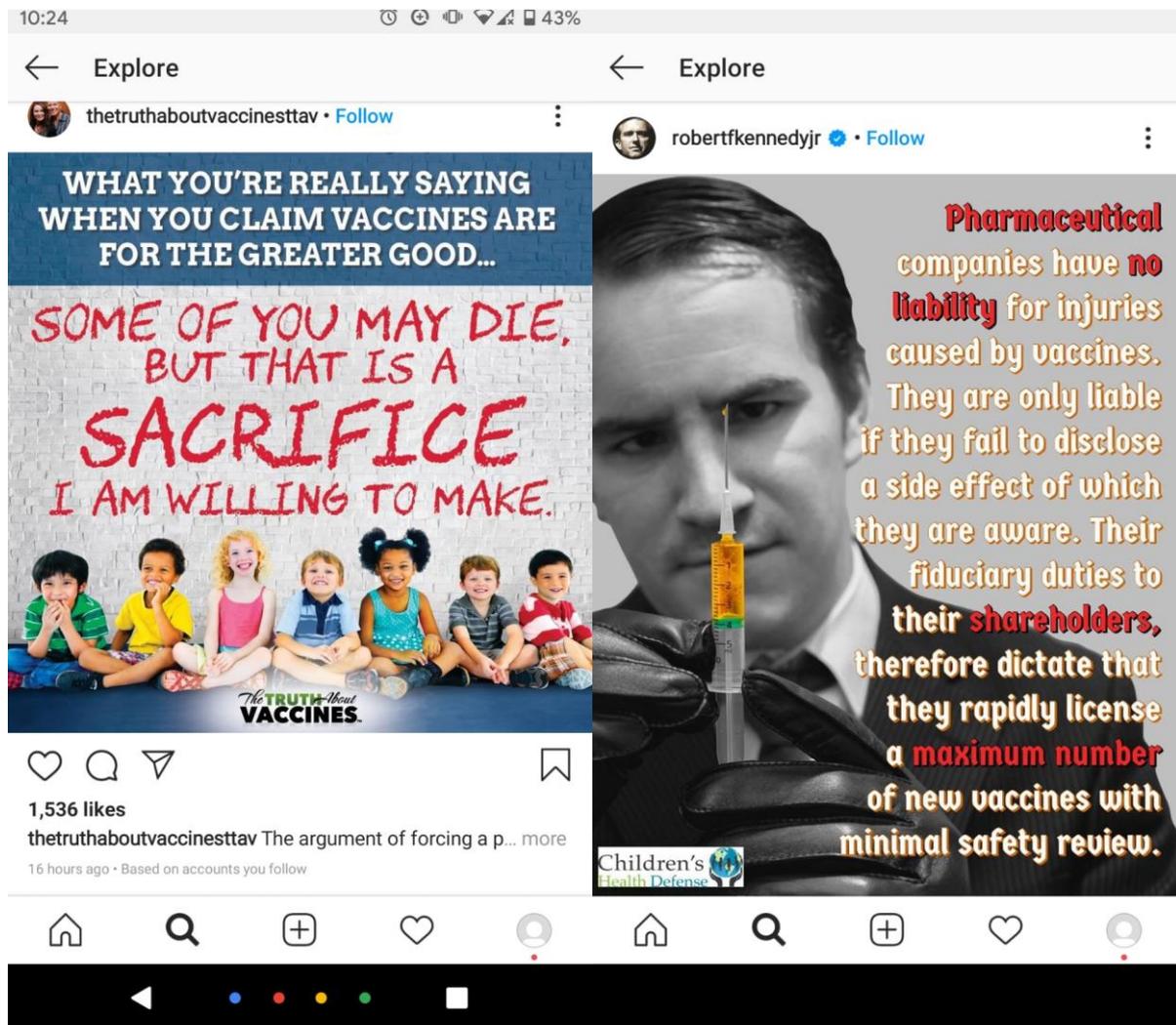


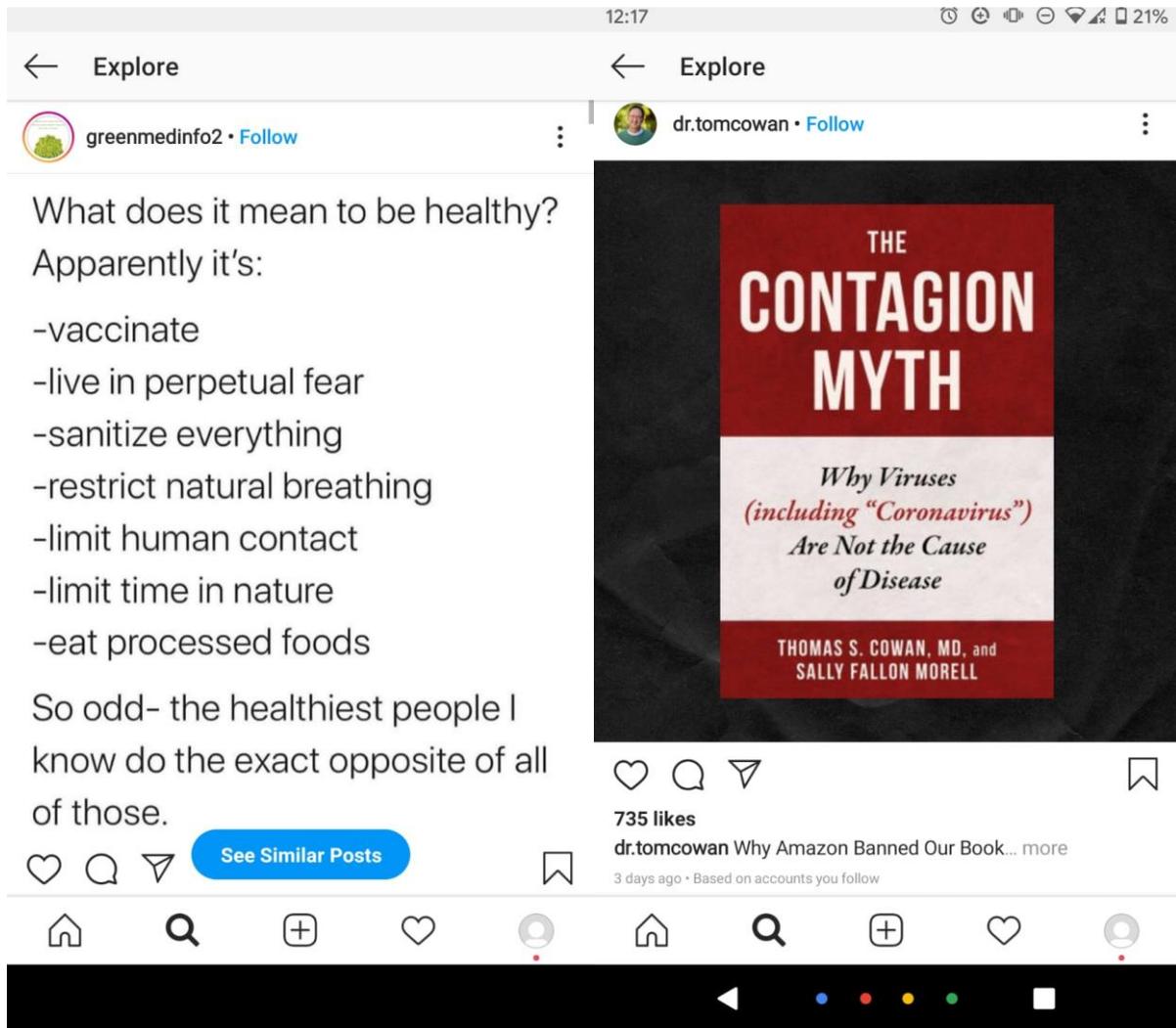
### Followers of wellness influencers linked to anti-vaxx are being recommended harder anti-vaccine content from leading anti-vaxxers already flagged to Instagram

Importantly, volunteers received a number of recommendations for content posted by leading anti-vaxxers who have been repeatedly flagged to Instagram and its parent company Facebook as sources of dangerous misinformation.

More than one volunteer was recommended content from Robert F. Kennedy Jr., while others were recommended posts from “The Truth About Vaccines”, an account run by the anti-vaxx entrepreneurs Ty & Charlene Bollinger. Two more commendations were for posts from “Green Med Info”, an account run by the anti-vaxx entrepreneur Sayer Ji. All of these individuals were identified in CCDH’s report *The Anti-Vaxx Playbook*.<sup>8</sup>

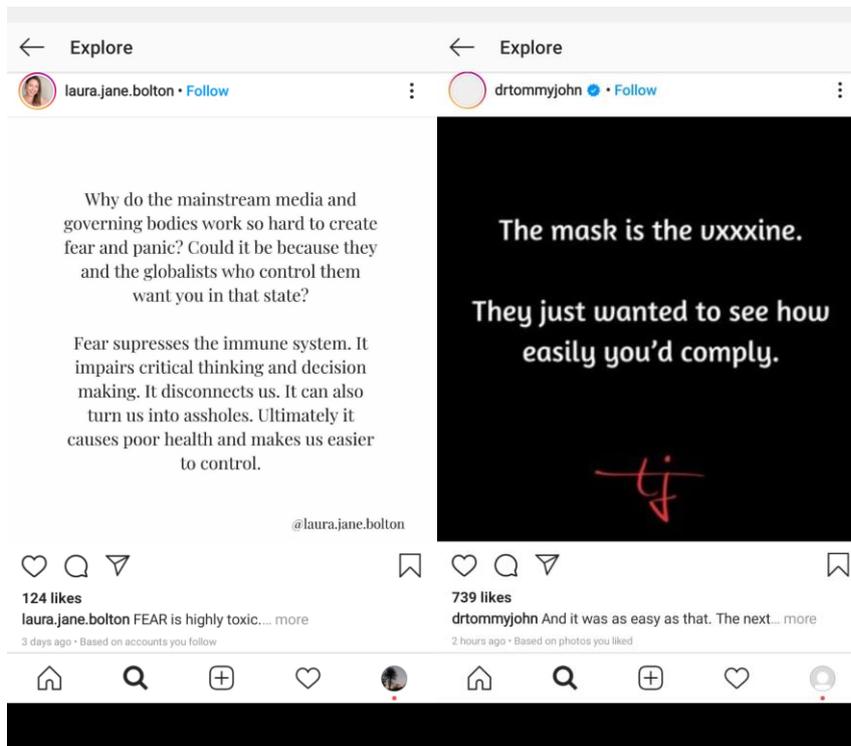
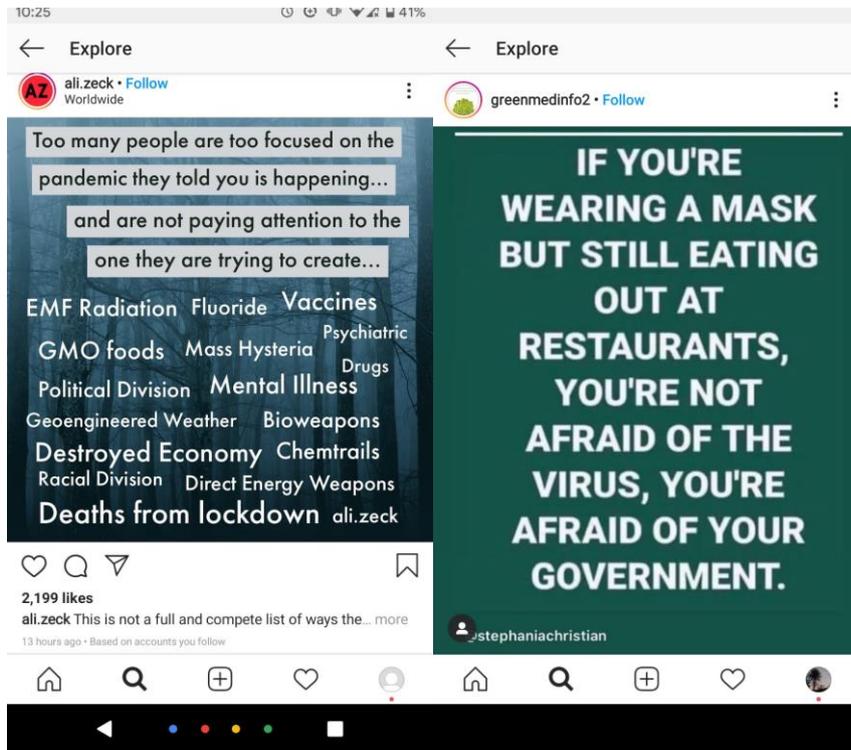
One further post from Thomas Cowan promoted his book claiming that the coronavirus is not a cause of disease. Videos of Cowan’s claims that Covid was linked to 5G mobile network signals spread rapidly during the early stages of the pandemic.<sup>9</sup>





## Followers of wellness influencers linked to anti-vaxx are recommended conspiracy theories about Covid and other subjects

Volunteers following this list recorded a number of posts promoting conspiracy theories that went beyond anti-vaccine misinformation. They included posts linking the Covid pandemic to “EMF radiation” and “chemtrails”, conspiracies about mask wearing, and claims that “globalists” were trying to control people with fear.



## List 3: Health authorities

Five volunteers followed this list of accounts that spread trustworthy information about Covid and vaccines, including national and international health organisations. The aim of this list was to investigate the recommendations that Instagram serves to people following only trustworthy sources of information about Covid and vaccines.

Together these accounts currently have 43.6 million followers on Instagram, making it likely that our findings apply to millions of Instagram users.

1. [nhs.uk](#) (verified)
2. [BBC News](#) (verified)
3. [World Health Organization](#) (verified)
4. [Gates Foundation](#) (verified)
5. [Public Health England](#)
6. [UK government](#) (verified)
7. [Centers for Disease Control](#) (verified)
8. [UNICEF](#) (verified)
9. [Bill Gates](#) (verified)
10. [LSHTM](#)

### **Recommendations recorded by our volunteers**

Volunteers using this list did not find any examples of misinformation recommended by Instagram's algorithm. This suggests that users who strictly avoid following accounts promoting misinformation are unlikely to receive recommendations for misinformation.

However, the second phase of our research investigated the recommendations received by volunteers who follow the above health authorities as well as smaller accounts promoting the QAnon conspiracy, and found that Instagram did recommend misinformation. This suggests that following health authorities is not enough to prevent Instagram's algorithm from recommending misinformation.

# Phase 2: Recommendations from Instagram’s “Suggested Posts” feature

In the second phase of our research beginning on 21 October 2020, volunteers were instructed to follow extended lists of accounts that were formed by adding additional accounts to their existing lists to form the following categories:

- Extended anti-vaxxers
- Health authorities and QAnon conspiracists
- Wellness and far-right accounts

These extended lists successfully triggered recommendations from Instagram’s new “Suggested Posts” feature. It also allowed us to explore how mixed lists of accounts influenced the recommendations generated by Instagram’s algorithm.

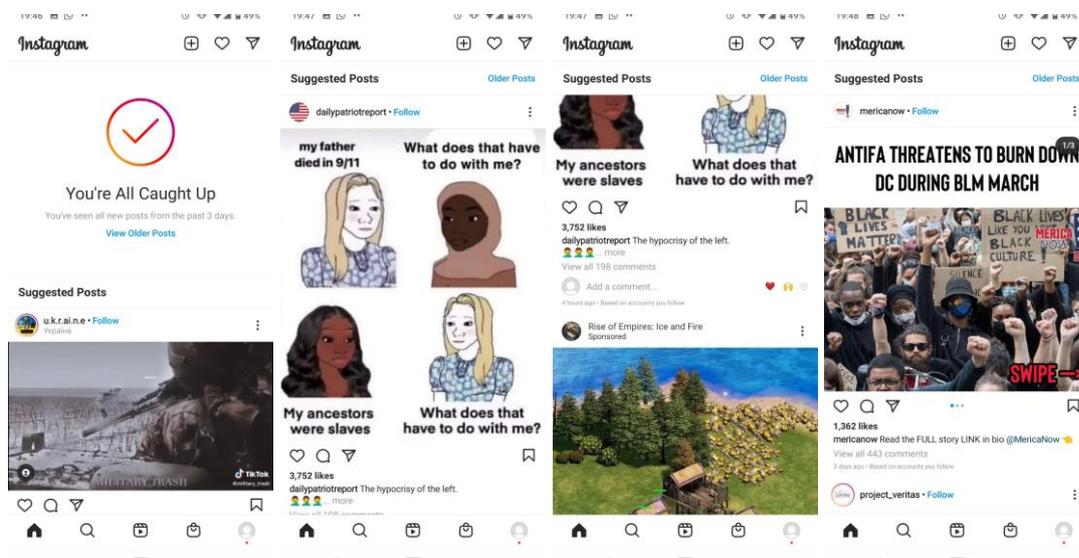
In this phase, volunteers collected 55 recommendations, including 39 Suggested Posts, 11 posts from Explore and five account recommendations. The recommended posts they recorded had attracted 131,142 likes with the help of Instagram’s algorithm.

## How Instagram’s “Suggested Posts” feature works

Instagram announced its “Suggested Posts” feature in August 2020, saying it was “a new way to see content you may like”.<sup>10</sup> Facebook’s Help Centre states that “suggestions are based on posts from accounts like the ones you follow and posts similar to the ones you like or save.”<sup>11</sup>

Users are presented with Suggested Posts after they have seen all the most recent posts from accounts they follow in their news feed. At this point Instagram presents the user with the message “You’re All Caught Up” and begins to show the user Suggested Posts from accounts they don’t follow, along with brief explanations of why they were recommended. Our research shows that Instagram inserts adverts between Suggested Posts to generate extra revenue.

The below examples of the Suggested Posts feature as it appears in in the Instagram app are taken from an account following a range of far-right profiles.



## List 4: Anti-vaxxers extended

Four volunteers followed this list which extended the anti-vaxxers list used in the first phased by adding less established anti-vaccine campaigners as well as alternative health entrepreneurs who play leading roles in the anti-vaccine movement.

Together these accounts currently have 7 million followers on Instagram, making it likely that our findings apply to millions of Instagram users.

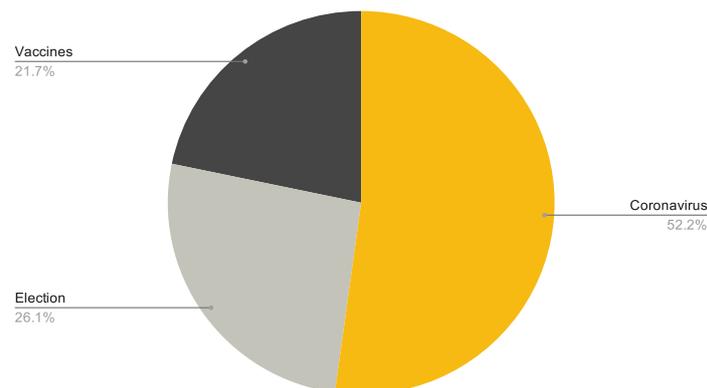
### List 1: Anti-vaxxers plus

11. [Ty & Charlene Bollinger](#)
12. [The Truth About Vaccines](#)
13. [The Truth About Cancer](#)
14. [Vaccines Uncovered](#)
15. [Dr. Shiva Ayyadurai](#)
16. [Rizza Islam](#)
17. [Denise Marie](#)
18. [Circle of Mamas](#)
19. [Erin Elizabeth](#)
20. [Dr. Joseph Mercola](#)

### Recommendations recorded by our volunteers

Volunteers following the above list of accounts recorded a total of 23 recommendations from Instagram’s algorithm that contained misinformation or hatred between 27 October and 16 November, 12 of which were Suggested Posts while the remaining 11 were from the Explore page.

Of those 23 recommendations, 12 primarily contained misinformation about Covid, while six contained misinformation about the US elections and five contained further misinformation about vaccines.



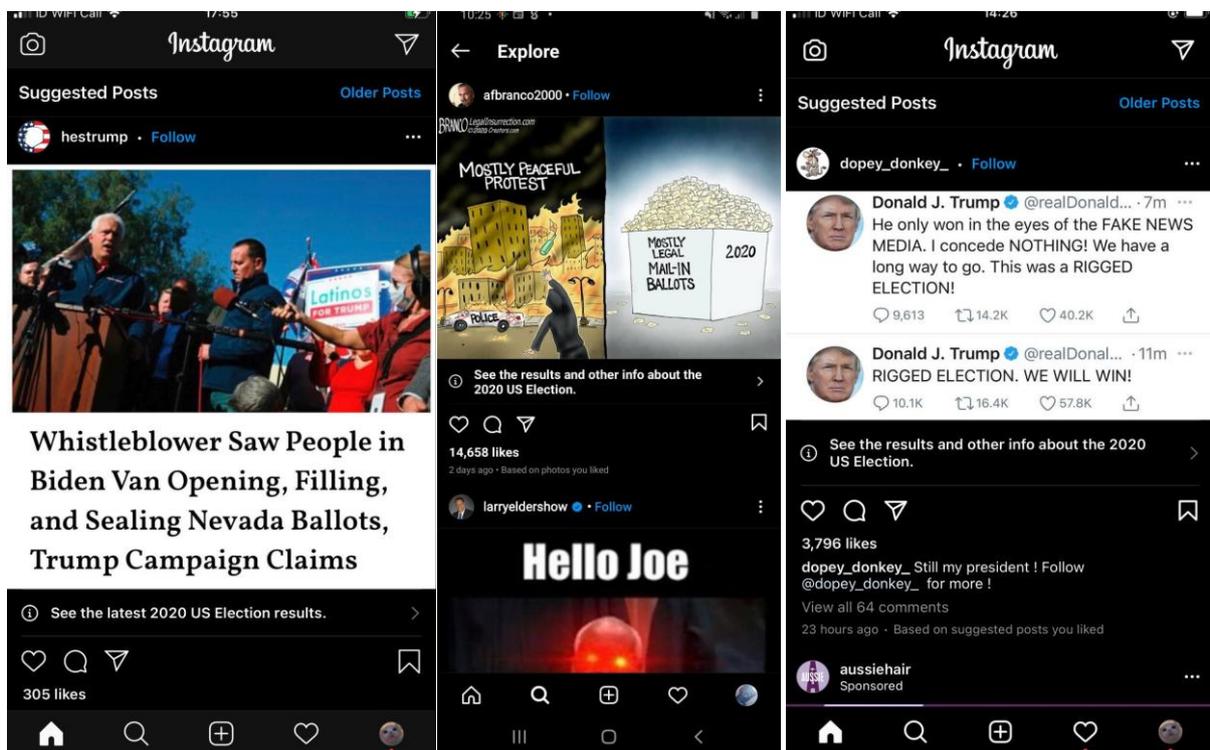
These findings from our volunteers highlighted four key issues that are described in further detail on the following pages:

- Recommendations for US election misinformation
- Recommendations for Covid misinformation
- Recommendations for further anti-vaccine content
- Instagram recommended content it had correctly identified as concerning vaccines or the US elections

## Followers of anti-vaxxers are being recommended misinformation about the US elections that Instagram correctly identified as election-related posts

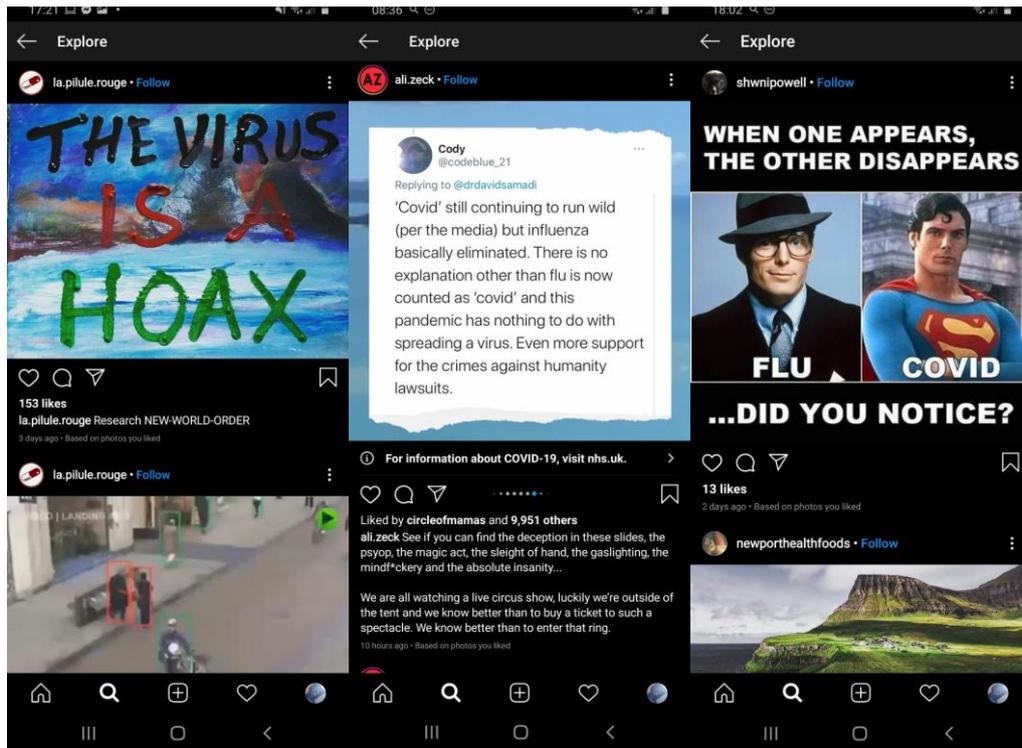
Volunteers recorded Suggested Posts containing misinformation about voting in the US elections. Many of these posts carried information labels about the US elections, showing that Instagram had correctly identified that they concerned the election but had recommended them anyway.

This suggests that Instagram may have been recommending election misinformation to millions of anti-vaxxers on its platform. Research conducted by CCDH has shown how some leading figures in the anti-vaccine movement are have used election misinformation to advance their own agenda.<sup>12</sup>



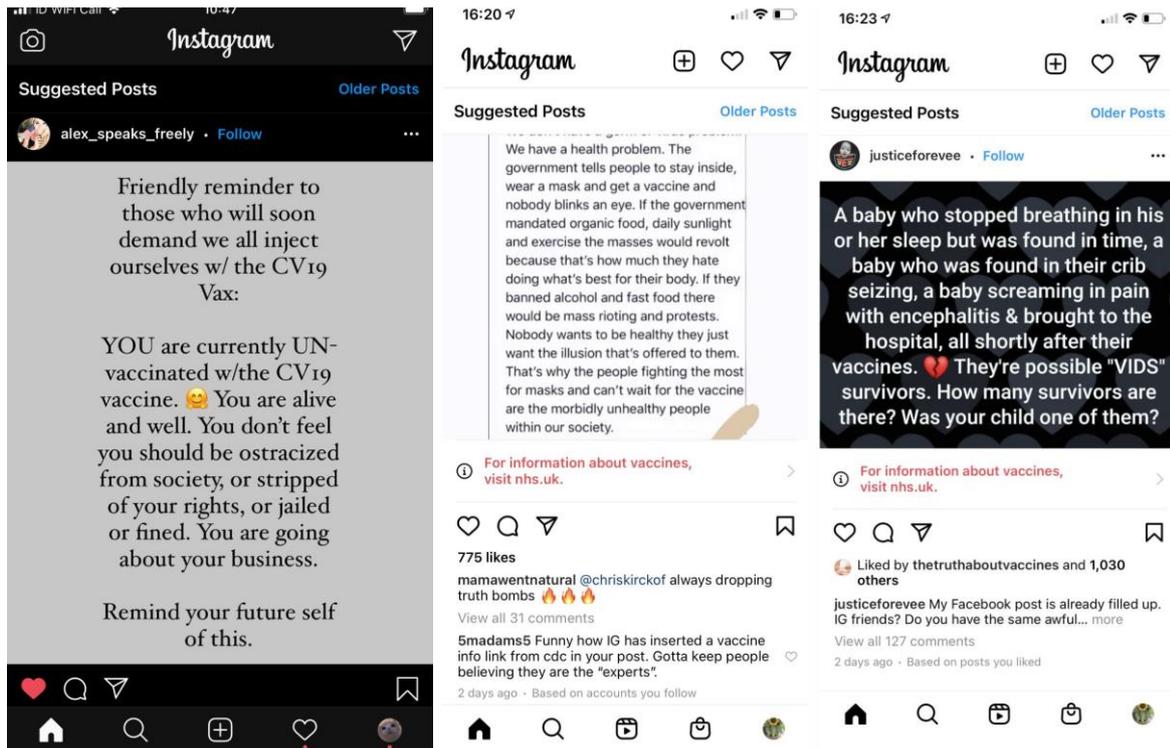
## Followers of anti-vaxxers are being recommended Covid misinformation, some of which Instagram correctly identified as concerning Covid

Volunteers continued to record examples of Covid misinformation recommended to followers of anti-vaccine accounts, a number of which denied that Covid existed. One recommended post carried an information label about Covid, suggesting that Instagram had recommended it despite correctly identifying that its topic was Covid.



## Followers of anti-vaxxers are being recommended further anti-vaccine misinformation from accounts they do not follow, some of which Instagram had identified as concerning vaccines

Volunteers continued to record Suggested Posts containing vaccine misinformation from accounts they did not follow. They also recorded a number of posts with information labels about vaccines, showing that Instagram’s algorithm had recommended them despite knowing that their topic was vaccines.



## List 5: Far-right and wellness

Four volunteers followed an extended version of the “wellness” list used in the first phase, adding a range of far-right accounts identified in the CCDH report “Hatebook”, plus a number of militia and “accelerationist” groups who aim to bring about civil war, for example groups identifying as part of “Boogaloo” movement. The aim was to study what kind of recommendations Instagram served to someone following accounts promoting far-right ideology and calls for violence.

The following far-right accounts were discovered in the course of research for the joint CCDH and Coalition for a Safer Web report “Hatebook” published in November 2020.<sup>13</sup> These accounts were all using Instagram and Facebook to sell neo-Nazi merchandise such as t-shirts and music.

### List 2: Wellness plus

11. [Asgard Store](#)
12. [Shturm](#)
13. [Voron](#)
14. [Opos Records](#)
15. [Tinnitus Records](#)
16. [Serbon Shop](#)
17. [Stay Brave Streetwear](#)
18. [Iron Youth Distro](#)
19. [Teivaz Store](#)
20. [Walknvt](#)

The following accounts either promote militia groups or an “accelerationist” ideology that aims to bring about civil war. All were identified by original research performed in October 2020, months after Facebook said it had banned militia groups including violent “Boogaloo” groups that harboured an accelerationist ideology.<sup>14</sup>

21. [Slavic Rambo](#)
22. [Minutemen Apparel](#)
23. [Mark The Booginator](#)
24. [Markof Daboog](#)
25. [The Luau Boys](#)
26. [ItsDeezyBaby](#)
27. [Alaskans For Liberty](#)
28. [Boogie Boi Shop](#)
29. [Duncan Lemp](#)
30. [Sic Semper Tyrannis](#)

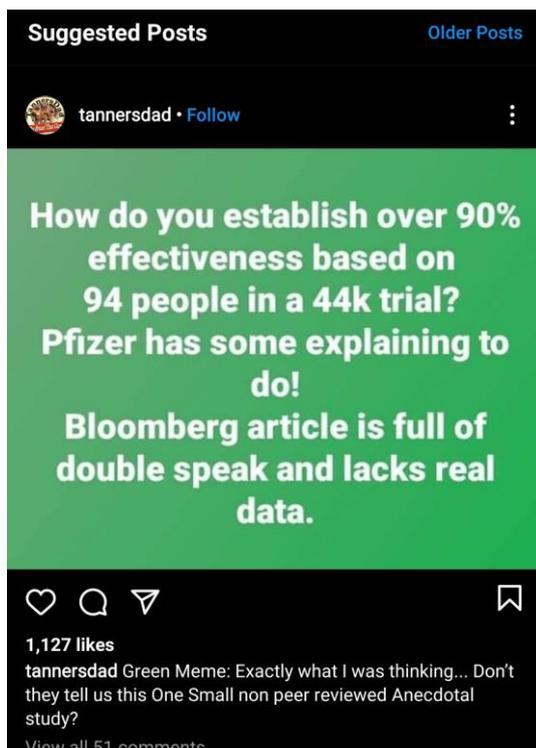
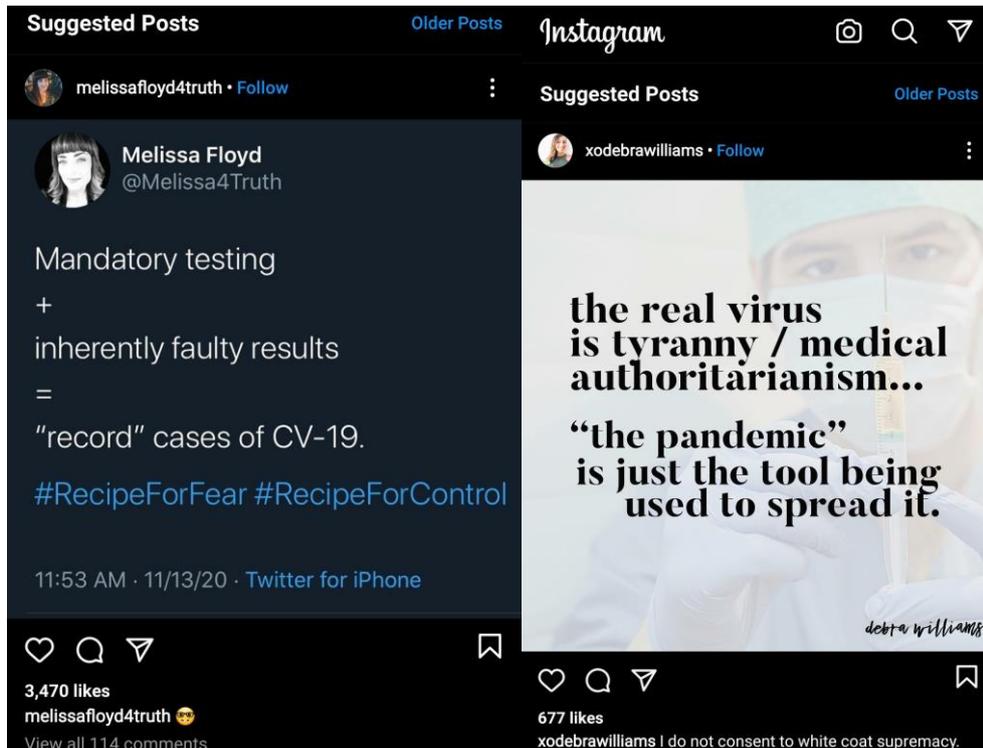
At the time this research was performed, the above Instagram accounts had 44,000 followers. Following the publication of the CCDH and Coalition for a Safer Web report “Hatebook”, all of the far-right accounts except for “Serbon Shop” were removed.

### Recommendations received by our volunteers

Volunteers following the above list of accounts recorded three recommendations, all of which were Suggested Posts. Two contained Covid misinformation while one concerned a Covid vaccine.

## Followers of far-right accounts also received recommendations for Covid and vaccine misinformation

Volunteers following far-right accounts received Suggested Posts claiming that Covid tests are “inherently faulty”, that the pandemic is being used to impose “medical authoritarianism” and that trials of the Pfizer Covid vaccine was approved on false data.



## List 6: QAnon and health authorities

Five volunteers were instructed to follow the “health authorities” list used in the first phase, adding another ten accounts known to promote the QAnon conspiracy theory. The aim was to investigate what recommendations Instagram served to someone following a combination of accounts producing reliable information and accounts spreading conspiratorial misinformation.

All of the following accounts were identified following Facebook’s announcement that it would ban QAnon accounts across all of its platforms.<sup>15</sup>

### List 3: Health authorities plus

11. [Sacha Stone](#)
12. [John Mappin](#)
13. [Great Awakening 2020](#)
14. [Esp Insta](#)
15. [The Storm Doc](#)
16. [Awaken Younify 5D](#)
17. [Maria777](#)
18. [Psyop Survivor](#)
19. [Dark Universe](#)
20. [Perception Shift](#)

The following accounts were initially part of the list but were removed by Instagram during the course of our research, and were replaced with other similar accounts.

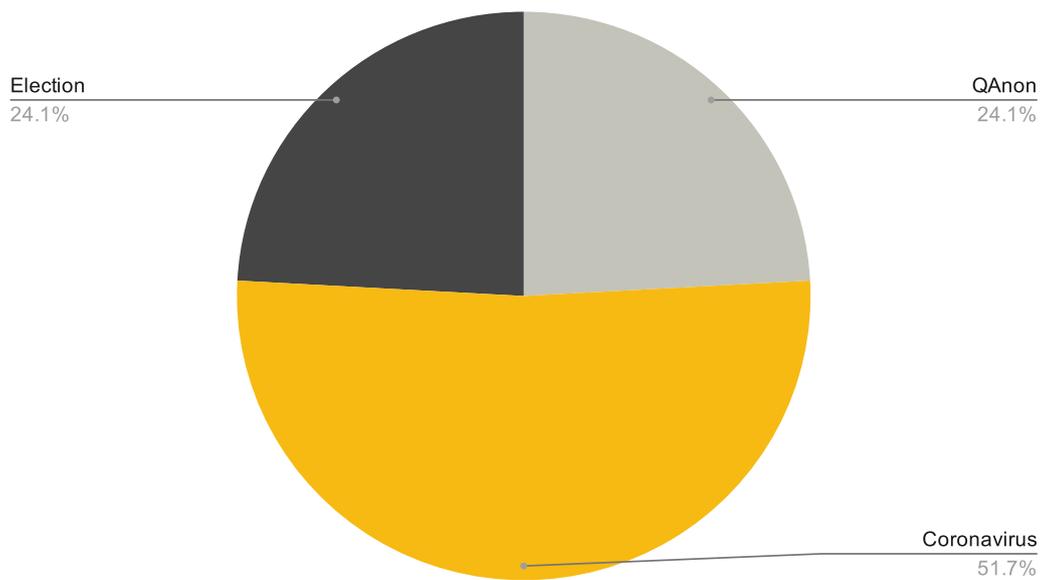
1. [Trump Believer](#)
2. [Crazy Conspiracy Theorist](#)
3. [Veritas](#)
4. [Kurious Keoki](#)
5. [5D Great Awakening](#)

The QAnon accounts on this list have just over 278,000 followers, much less than the 43.6 million followers of the health authorities list it was combined with. It means that each QAnon account on the result list had an average of 46,000 followers, compared to an average of 4.4 million followers for health authority accounts.

### Recommendations recorded by our volunteers

Volunteers following the above list of accounts recorded a total of 29 recommendations from Instagram’s algorithm that contained misinformation or hatred between 23 October and 16 November, 24 of which were Suggested Posts while the remaining five were Suggested Accounts.

Of the 24 Suggested Posts containing misinformation, 15 primarily concerned Covid, while seven concerned the US elections and two promoted QAnon conspiracy theories. All five sets of Suggested Accounts contained recommendations for further accounts spreading QAnon conspiracy theories.



These findings from our volunteers highlighted four key issues that are described in further detail on the following pages:

- Recommendations to follow further QAnon accounts
- Recommendations to follow the backup accounts of QAnon accounts they were already following
- Recommendations for US election misinformation
- Recommendations for Covid misinformation

### **Following a mix of authoritative and misinformation accounts led to recommendations for more misinformation**

Importantly, posts we collected from this list show that following a mix of highly influential health authorities and much less influential QAnon accounts caused Instagram to recommend further misinformation to volunteers.

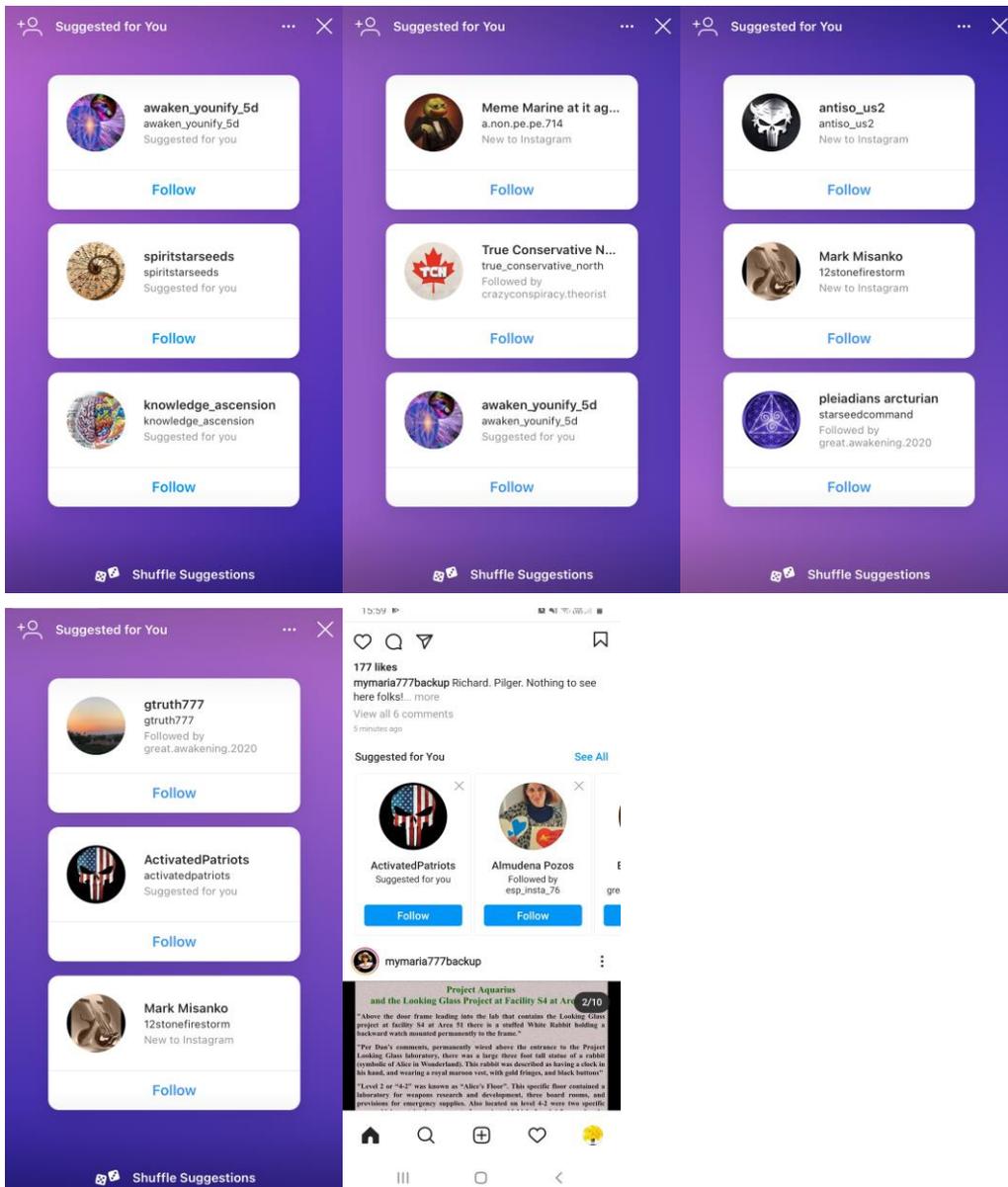
This is a marked contrast with volunteers following health authorities alone in the first phase who received no recommendations for misinformation. It suggests that users who follow trustworthy accounts but begin to follow accounts dealing in misinformation could be subject to a radicalising effect as Instagram presents them with more misinformation and conspiratorial content in the form of Suggested Posts.

## Followers of QAnon accounts received recommendations for further QAnon accounts they do not follow, including backup accounts for accounts they did follow

Volunteers received a number of recommendations for accounts that spread misinformation and conspiracy theories. These included QAnon accounts named “antiso\_us2” and “12stonefirestorm”, as well as a backup for a far-right meme account with the handle “a.non.pe.pe.714”.<sup>16</sup>

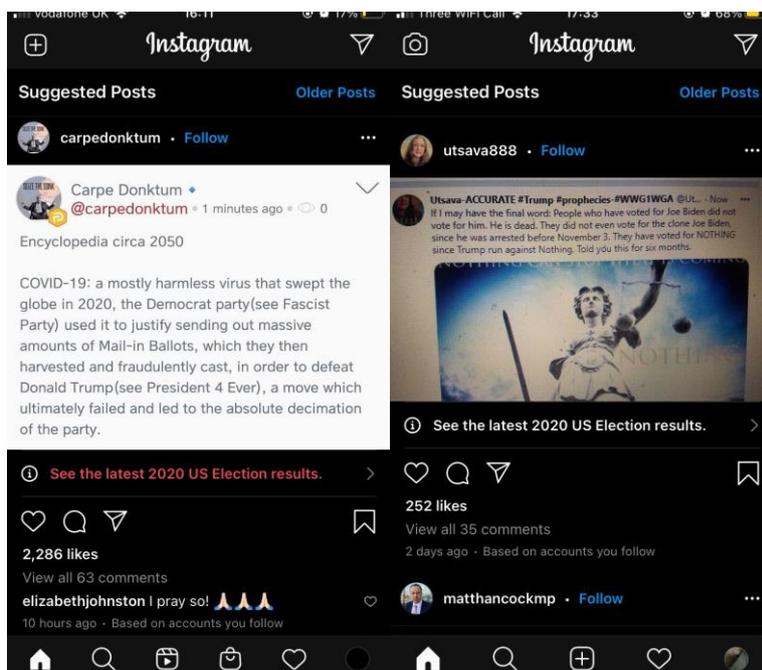
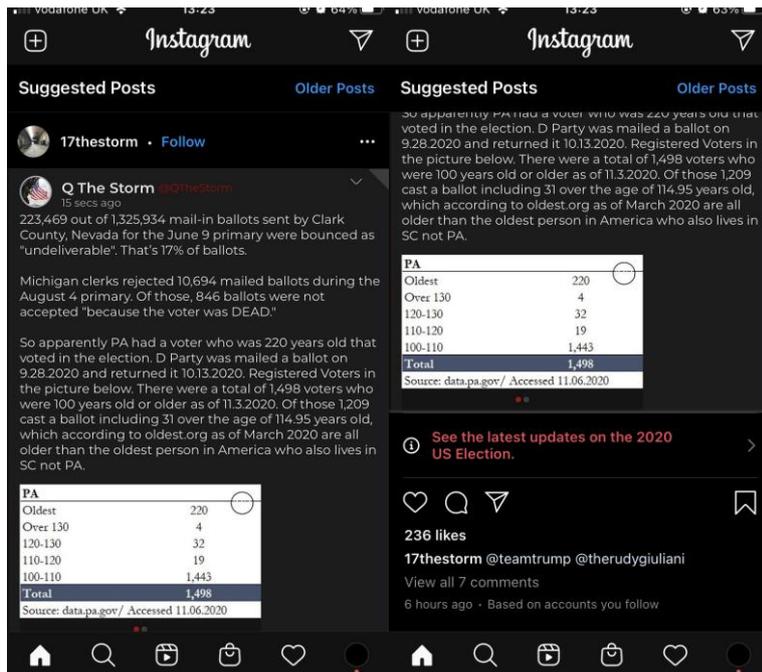
Instagram also recommended “awaken\_younify\_5d”, “spiritstarseeds” and “knowledge\_ascension”, all of which are part of a network of backup accounts linked to “5d\_greatawakening”, an account followed by our volunteers and removed by Instagram during the course of their research.<sup>17</sup>

These recommendations not only suggest that Instagram’s algorithm is helping boost the following of conspiracy and misinformation accounts, but that it is actively directing users to follow backup accounts that share extreme overlap with accounts they already follow, frustrating Instagram’s own attempts to remove misinformation from its platform.



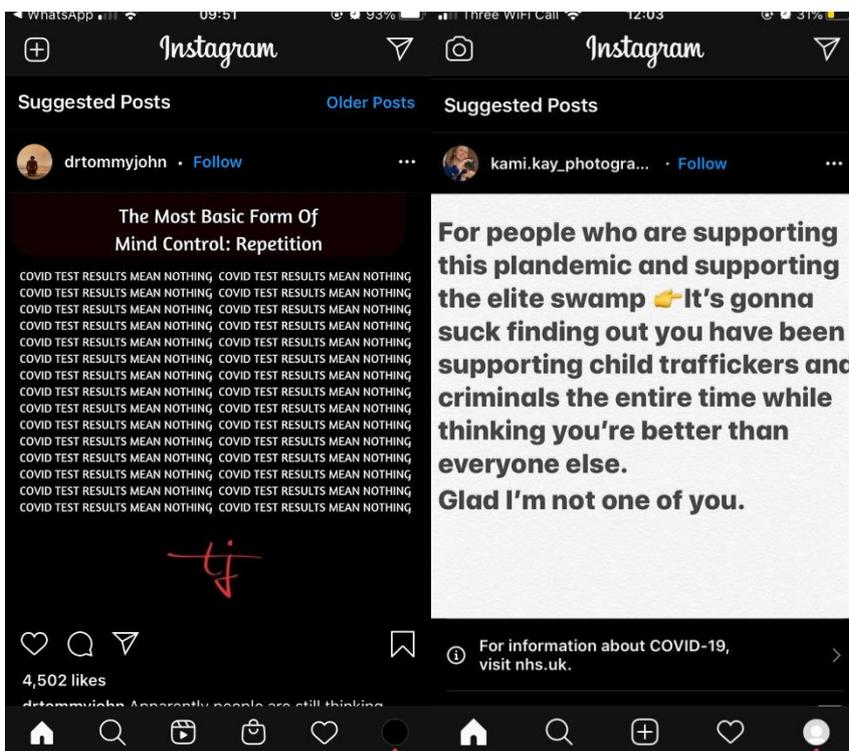
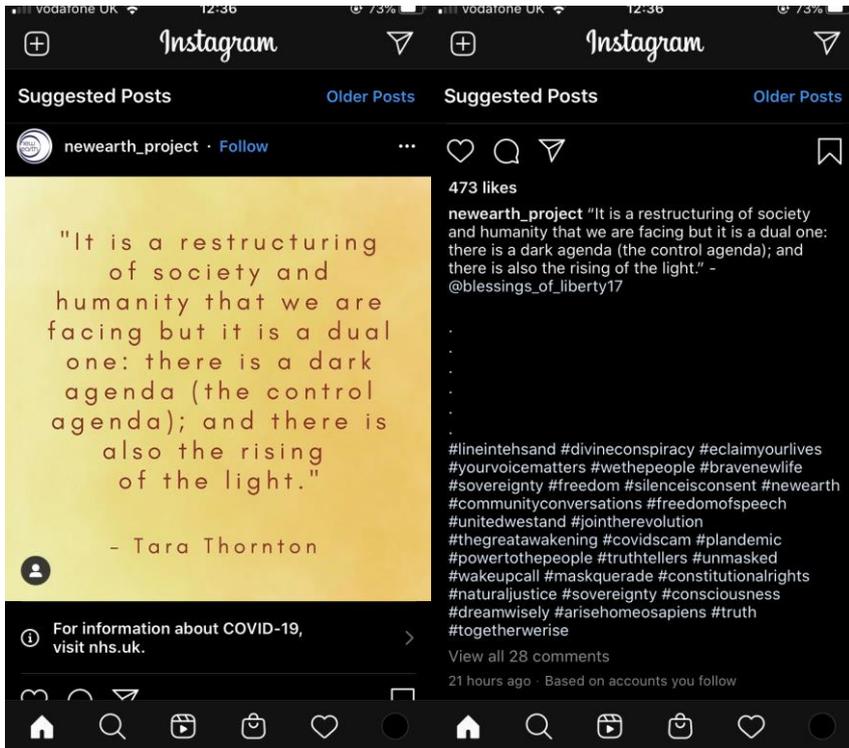
## Followers of QAnon accounts were recommended conspiracy theories that Instagram had identified as concerning the US elections

Volunteers following QAnon accounts received Suggested Posts containing misinformation about the US elections, including the claim that mail-in ballots are “harvested and fraudulently cast”. One Suggested Post claiming that Joe Biden “is dead” appeared above a post by the UK health secretary Matt Hancock MP. A number of these posts carried information labels about the US elections, suggesting that Instagram recommended the posts despite knowing their topic.



## Followers of QAnon accounts are being recommended posts containing Covid misinformation that Instagram had identified as concerning Covid

Volunteers received Suggested Posts containing Covid misinformation, some of which claimed the disease was a “plandemic”. Some of these posts carried automatic information labels about Covid, suggesting that Instagram allowed its algorithm to recommend them despite correctly identifying that they were about Covid.



# Recommendations

This report shows that Instagram extended the use of its algorithm to publish “Suggested Posts” to every user in the middle of a pandemic, without doing the necessary work to ensure it would not promote dangerous misinformation about Covid or vaccines. This is despite Facebook’s promise, made in March 2019, that it “won’t show or recommend content that contains misinformation about vaccinations on Instagram Explore or hashtag pages” first made in March 2019.<sup>18</sup>

## **Instagram must suspend its algorithm**

The pandemic is far from over and vaccination programmes have only just begun. In that context, it is too dangerous for Instagram’s algorithm to continue as it is, publishing misinformation to millions of people that risks undermining vaccination and lengthening the pandemic. It should be suspended until Instagram, and its parent company Facebook, can show that it is no longer publishing misinformation.

Instagram should only reinstate its algorithmic publishing features when they are proven to be safe. This should involve a new assessment of the safety of its recently introduced “Suggested Posts” feature given that Facebook decided to limit its use of such features just two years ago because of their impact on the mental health of users.

## **Fixing Instagram’s algorithm**

Our research suggests a number of concrete steps that Instagram can take to prevent its algorithm from promoting misinformation. In future, the potential of a new feature to promote dangerous misinformation or hatred should be thoroughly assessed before it is introduced.

- **Exclude posts about Covid, vaccines or contentious elections from its algorithm.** Instagram’s own text and image recognition algorithms are correctly identifying posts on these subjects and affixing information labels to them, but its publishing algorithm is amplifying them regardless of whether they contain misinformation. Such posts should be excluded from Instagram’s recommendation algorithm given that it cannot ensure that they do not contain dangerous misinformation.
- **Maintain a blacklist of accounts known to spread misinformation.** Our research shows that Instagram is publishing and amplifying the posts of well-known anti-vaxxers that have already been reported to the platform. It should create a blacklist of accounts whose posts are excluded from Instagram’s algorithms, identifying them with the help of health authorities such as the WHO as well as NGOs like CCDH. Freedom of speech does not mean freedom of reach: persistent spreaders of dangerous misinformation do not have a right to reach millions of people on social media aided by platforms’ recommendation algorithms.
- **Limit the number of suggested posts that a user is exposed to.** Two years ago, Facebook took steps to limit the time users spend in its apps over concerns for their mental health. It should once again introduce clear measures to limit the amount of algorithmically published content its users can consume in a given time period.

## Reduce the impact of misinformation on Instagram

In addition to taking measures to prevent its algorithm from publishing misinformation, our research suggests that Instagram can do much more to reduce the impact and spread of misinformation on its platform.

- **Remove verified status from known anti-vaxxers, and prevent spreaders of hate and misinformation from acquiring verified status.** Regardless of Instagram's intentions, many users will falsely interpret verified status as a sign that Instagram has endorsed a user or vouched for the accuracy of their posts.
- **Introduce warnings on algorithmically published content.** Research shows that social media users have a poor understanding of why certain posts are included in their news feeds.<sup>19</sup> Instagram should warn users that posts it publishes to their feeds have not been checked for their accuracy and may contain misinformation.
- **Promote posts that inoculate users against misinformation.** Research by Jon Roozenbeek & Sander van der Linden shows that "inoculating" users by "pre-emptively exposing, warning, and familiarising people with the strategies used in the production of fake news" makes them more resistant to the misinformation they subsequently encounter online.<sup>20</sup> Instagram should work with experts to design effective "inoculation" content and promote it in users' feeds to make them more resistant to misinformation they might encounter on the platform.
- **Investigate the algorithm's role in promoting "backup accounts".** While following accounts known to spread the QAnon conspiracy theory, our volunteers received recommendations to follow "backup accounts" for them. Instagram should review this feature to ensure it is not helping spreaders of misinformation evade efforts to remove their accounts and content.

## Remedy the damage already done by Instagram's algorithm

Instagram's "Suggested Posts" feature was rolled out in August 2020. Since then, it will have recommended dangerous misinformation to millions of users during a pandemic. Instagram's "Explore" feature has been publishing misinformation to users for even longer, having first introduced algorithmic recommendations in April 2014.<sup>21</sup> Instagram should remedy the damage caused by both features.

- **Display corrective posts to users exposed to misinformation.** Instagram should identify those users who were recommended posts known to contain misinformation or from actors known to spread misinformation, and present them with a greater number of corrective posts. These corrective posts should be designed with input from experts to ensure that they help address the negative social consequences of misinformation such as vaccine hesitancy without inadvertently entrenching their opinions through the "backfire effect".<sup>22</sup>
- **Donate ad revenue from misinformation to health organisations.** Our volunteers identified a number of algorithmically published posts presented alongside adverts, highlighting the feature's role in increasing the time users spend on Instagram and the number of adverts they see. Instagram's parent company Facebook made a commitment in November 2020 not to profit from vaccine misinformation.<sup>23</sup> It should make a commitment to donate revenue from ads displayed next to posts containing misinformation or from known spreaders of misinformation to health organisations fighting the pandemic.

## Appendix: Youth Leadership

Youth Against Misinformation launched in April 2020 in response to the growing Covid-19 misinformation pandemic online. The objectives of the initiative were to fight the Covid-19 misinfodemic, empower young people to balance the narrative in online spaces and support other groups to tackle Covid-19 misinformation online by identifying new trends.

Restless Development strives to unlock youth power everywhere and ensure young people can lead change while growing as leaders themselves. We do this by creating environments where young people can learn, be creative and shape our programmes, as well as ensuring that they build useful skills to better develop their leadership, activism and employment opportunities. Feedback from volunteers involved in this programme has shown that 100% feel empowered to help stop the spread of dangerous Covid-19 misinformation while over 95% agree that their work has had impact in achieving the goals of the programme.

The safety and security of staff and volunteers is of paramount importance to Restless Development; the young volunteers involved in Youth Against Misinformation have been fully trained and supported to identify and report Covid-19 and anti-vaccine misinformation, including setting up new email and social media accounts. All volunteers in the Youth Against Misinformation programme received extensive safeguarding training, including how to protect their online safety, had weekly debriefs with a Restless Development staff member and also had a team leader in their group to support them while undertaking the research to ensure their ongoing health and wellbeing. Volunteers' identities have been kept anonymous throughout the campaign to protect them from online abuse.

For more information on Restless Development's safeguarding policy visit <https://restlessdevelopment.org/safeguarding/>

## Appendix: Misinformation recorded

Screenshots of all 104 pieces of misinformation recorded by our volunteers, along with details of the posts and when they were recorded, are available on our website at [counterhate.com/malgorithm](https://counterhate.com/malgorithm).

**Published 9 March 2021**

**© 2020 Center for Countering Digital Hate Ltd**

- 
- <sup>1</sup> Facebook, 25 November 2019, <https://ai.facebook.com/blog/powered-by-ai-instagrams-explore-recommender-system/>
- <sup>2</sup> Instagram, retrieved 18 January 2021, <https://www.facebook.com/help/instagram/231704903690654>
- <sup>3</sup> TechCrunch, 19 August 2020, <https://techcrunch.com/2020/08/19/instagram-finds-new-ad-space-at-the-end-of-your-feed-with-launch-of-suggested-posts-feature/>
- <sup>4</sup> Instagram, retrieved 19 January 2021, <https://help.instagram.com/381638392275939>
- <sup>5</sup> Instagram, 26 June 2019, <https://business.instagram.com/blog/bringing-ads-to-explore/>
- <sup>6</sup> Facebook, 19 August 2020, <https://about.fb.com/news/2020/08/addressing-movements-and-organizations-tied-to-violence/>
- <sup>7</sup> “The Anti-Vaxx Industry”, Center for Countering Digital Hate, 6 July 2020, <https://www.counterhate.co.uk/anti-vaxx-industry>
- “The Anti-Vaxx Playbook”, Center for Countering Digital Hate, 22 December 2020, <https://www.counterhate.com/playbook>
- <sup>8</sup> “The Anti-Vaxx Playbook”, Center for Countering Digital Hate, 22 December 2020, <https://www.counterhate.com/playbook>
- <sup>9</sup> CBC, 23 March 2020, <https://www.cbc.ca/news/technology/fact-check-viral-video-coronavirus-1.5506595>
- <sup>10</sup> Instagram, 20 August 2018, <https://about.instagram.com/blog/announcements/testing-recommended-posts-in-feed>
- <sup>11</sup> Instagram, retrieved 19 January 2021, <https://help.instagram.com/381638392275939>
- <sup>12</sup> Center for Countering Digital Hate, Twitter, 6 January 2021, <https://twitter.com/CCDHate/status/1346894921467363332>
- CNN, 4 February 2021, <https://edition.cnn.com/2021/02/04/politics/anti-vaxxers-stop-the-steal-invs/index.html>
- “The Anti-Vaxx Industry”, Center for Countering Digital Hate, 6 July 2020, page 17, <https://www.counterhate.co.uk/anti-vaxx-industry>
- <sup>13</sup> “Hatebook: Facebook’s neo-Nazi shopfronts funding far-right extremism”, Center for Countering Digital Hate and Coalition for a Safer Web, 22 November 2020, <https://www.counterhate.com/hatebook>
- <sup>14</sup> BBC News, 1 July 2020, <https://www.bbc.co.uk/news/world-us-canada-53244339>
- BBC News, 20 August 2020, <https://www.bbc.co.uk/news/technology-53849295>
- <sup>15</sup> BBC News, 6 October 2020, <https://www.bbc.co.uk/news/world-us-canada-54443878>
- <sup>16</sup> antiso\_us2, Instagram, 10 January 2021, <https://archive.vn/RiHi3>
- 12stonefirestorm, Instagram, 20 December 2020, <https://archive.vn/81WuQ>
- a.non.pe.pe.714, Instagram, archived 31 January 2021, <https://archive.vn/W54j0>
- oc.anon.714, Instagram, archived 31 January 2021, <https://archive.vn/qBjHA>
- <sup>17</sup> awaken\_younify\_5d, Instagram, archived 31 January 2021, <https://archive.vn/hmbWP>
- spiritstarseeds, Instagram, archived 31 January 2021, <https://archive.vn/MFsc5>
- knowledge\_ascension, Instagram, 28 August 2020, <https://archive.vn/Ovmu0>
- awaken\_younify\_5d, Instagram, 12 October 2020, <https://archive.vn/CRGp4>
- spiritstarseeds, Instagram, 19 September 2020, <https://archive.vn/bhvMX>
- awaken\_younify\_5d, Instagram, 17 October 2020, <https://archive.vn/vxE9N>
- <sup>18</sup> Facebook, 7 March 2019, <https://about.fb.com/news/2019/03/combating-vaccine-misinformation/>
- <sup>19</sup> Pew Research Center, 5 September 2018, <https://www.pewresearch.org/fact-tank/2018/09/05/many-facebook-users-dont-understand-how-the-sites-news-feed-works/>
- <sup>20</sup> Roozenbeek, J., van der Linden, S. Fake news game confers psychological resistance against

---

online misinformation. *Palgrave Commun* 5, 65 (2019). <https://doi.org/10.1057/s41599-019-0279-9>

<sup>21</sup> TechCrunch, 23 April 2014, <https://techcrunch.com/2014/04/23/amateur-beauty-vs-professional-entertainment/>

<sup>22</sup> Nyhan B, Reifler J. Does correcting myths about the flu vaccine work? An experimental evaluation of the effects of corrective information. *Vaccine*. 2015 Jan 9;33(3):459-64. doi: [10.1016/j.vaccine.2014.11.017](https://doi.org/10.1016/j.vaccine.2014.11.017). Epub 2014 Dec 8. PMID: 25499651.

Skurnik, I., Yoon, C., Schwarz, N. (2007). "Myths & Facts" about the flu: Health education campaigns can reduce vaccination intentions. Retrieved from [http://webuser.bus.umich.edu/yoonc/research/Papers/Skurnik\\_Yoon\\_Schwarz\\_2005\\_Myths\\_Facts\\_Flu\\_Health\\_Education\\_Campaigns\\_JAMA.pdf](http://webuser.bus.umich.edu/yoonc/research/Papers/Skurnik_Yoon_Schwarz_2005_Myths_Facts_Flu_Health_Education_Campaigns_JAMA.pdf)

Peter C, Koch T. When Debunking Scientific Myths Fails (and When It Does Not): The Backfire Effect in the Context of Journalistic Coverage and Immediate Judgments as Prevention Strategy. *Science Communication*. 2016;38(1):3-25. doi:[10.1177/1075547015613523](https://doi.org/10.1177/1075547015613523)

<sup>23</sup> UK Government, 8 November 2020, <https://www.gov.uk/government/news/social-media-giants-agree-package-of-measures-with-uk-government-to-tackle-vaccine-disinformation>