



David Zweig

@davidzweig

1. THREAD: THE TWITTER FILES: HOW TWITTER RIGGED THE COVID DEBATE – By censoring info that was true but inconvenient to U.S. govt. policy – By discrediting doctors and other experts who disagreed – By suppressing ordinary users, including some sharing the CDC’s *own data*

2. So far the Twitter Files have focused on evidence of Twitter’s secret blacklists; how the company functioned as a kind of subsidiary of the FBI; and how execs rewrote the platform’s rules to accommodate their own political desires.

3. What we have yet to cover is Covid. This reporting, for The Free Press, [@thefp](#), is one piece of that important story.

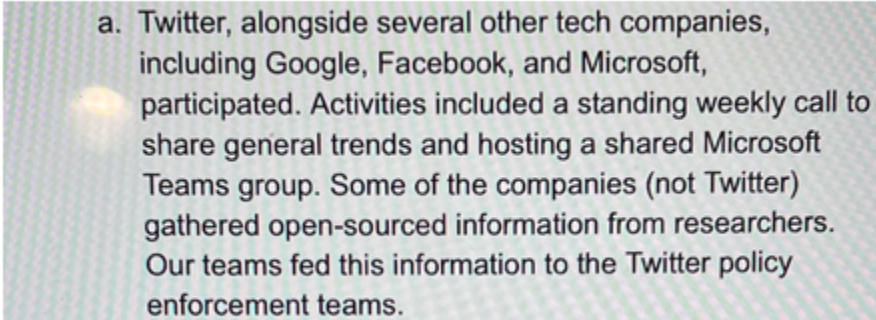
4. The United States government pressured Twitter and other social media platforms to elevate certain content and suppress other content about Covid-19.

5. Internal files at Twitter that I viewed while on assignment for [@thefp](#) showed that both the Trump and Biden administrations directly pressed Twitter executives to moderate the platform’s pandemic content according to their wishes.

6. At the onset of the pandemic, according to meeting notes, the Trump admin was especially concerned about panic buying. They came looking for “help from the tech companies to combat misinformation” about “runs on grocery stores.” But . . . there were runs on grocery stores.

- The Trump White House, specifically Michael Kratsios, led the Trump Administration’s calls for help from the tech companies to combat misinformation. Areas of focus included conspiracies around 5G cell towers, runs on grocery stores, and misinformation that could stoke panic buying and behaviors. Coverage [here](#). Documents [here](#) and [here](#).

7. It wasn't just Twitter. The meetings with the Trump White House were also attended by Google, Facebook, Microsoft and others.

A screenshot of a document snippet with a yellow highlight. The text describes a meeting with the Trump White House involving several tech companies.

a. Twitter, alongside several other tech companies, including Google, Facebook, and Microsoft, participated. Activities included a standing weekly call to share general trends and hosting a shared Microsoft Teams group. Some of the companies (not Twitter) gathered open-sourced information from researchers. Our teams fed this information to the Twitter policy enforcement teams.

8. When the Biden admin took over, one of their first meeting requests with Twitter executives was on Covid. The focus was on “anti-vaxxer accounts.” Especially Alex Berenson:

A screenshot of a document snippet with a mouse cursor. The text describes a meeting request from the Biden White House regarding COVID misinformation.

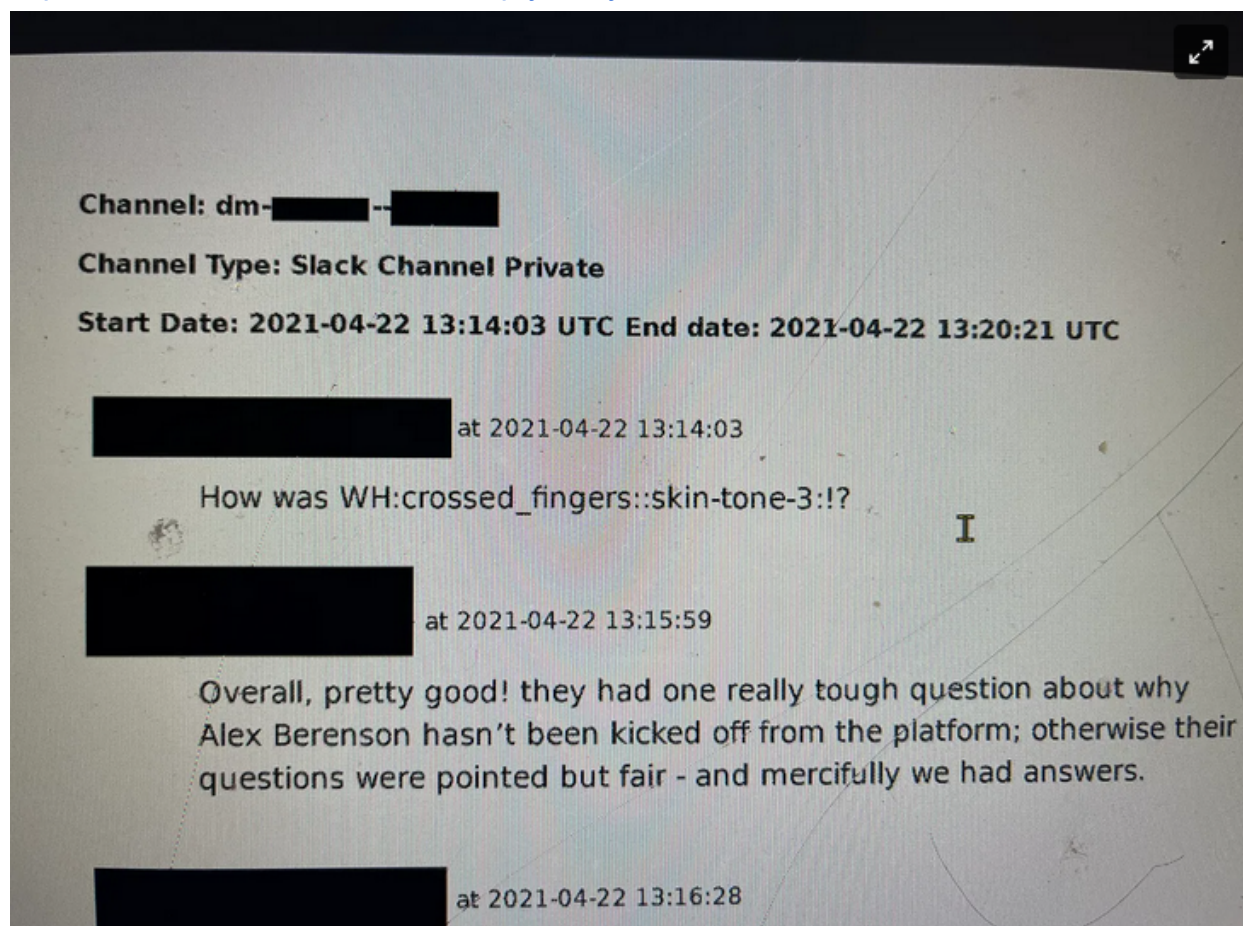
Biden

- One of the first meeting requests from the Biden White House was about COVID misinformation. Per regular process, Public Policy took the meeting. Biden's staff focused on vaccines and high-profile anti-vaxxer accounts, including Alex Berenson.

9. In the summer of 2021, president Biden said social media companies were “killing people” for allowing vaccine misinformation. Berenson was suspended hours after Biden's comments, and kicked off the platform the following month.

10. Berenson sued (and then settled with) Twitter. In the legal process Twitter was compelled to release certain internal communications, which showed direct White House pressure on the company to take action on Berenson.

<https://alexberenson.substack.com/p/jesse-jackson-cant-swim>



11. A December 2022 summary of meetings with the White House by Lauren Culbertson, Twitter's Head of U.S. Public Policy, adds new evidence of the White House's pressure campaign, and cements that it repeatedly attempted to directly influence the platform.

12. Culbertson wrote that the Biden team was "very angry" that Twitter had not been more aggressive in deplatforming multiple accounts. They wanted Twitter to do more.

- The Biden team was not satisfied with Twitter's enforcement approach as they wanted Twitter to do more and to deplatform several accounts. Because of this dissatisfaction, we were asked to join several other calls. They were very angry in nature.
- Here's a [presentation](#) the Twitter team used.

13. Twitter executives did not fully capitulate to the Biden team’s wishes. An extensive review of internal communications at the company revealed employees often debating moderation cases in great detail, and with more care than was shown by the government toward free speech.

14. But Twitter did suppress views—many from doctors and scientific experts—that conflicted with the official positions of the White House. As a result, legitimate findings and questions that would have expanded the public debate went missing.

15. There were three serious problems with Twitter’s process:

First, much of the content moderation was conducted by bots, trained on machine learning and AI – impressive in their engineering, yet still too crude for such nuanced work.

16. Second, contractors, in places like the Philippines, also moderated content. They were given decision trees to aid in the process, but tasking non experts to adjudicate tweets on complex topics like myocarditis and mask efficacy data was destined for a significant error rate

The image shows a screenshot of a decision tree interface used for content moderation. The interface is divided into two main sections: a left column with decision questions and a right column with enforcement recommendations. The left column contains four questions, each with a dropdown menu for the answer. The right column contains three recommendations, each with a dropdown menu for the action. The interface is titled 'Enforcement Recommendation' at the top right. The questions and recommendations are as follows:

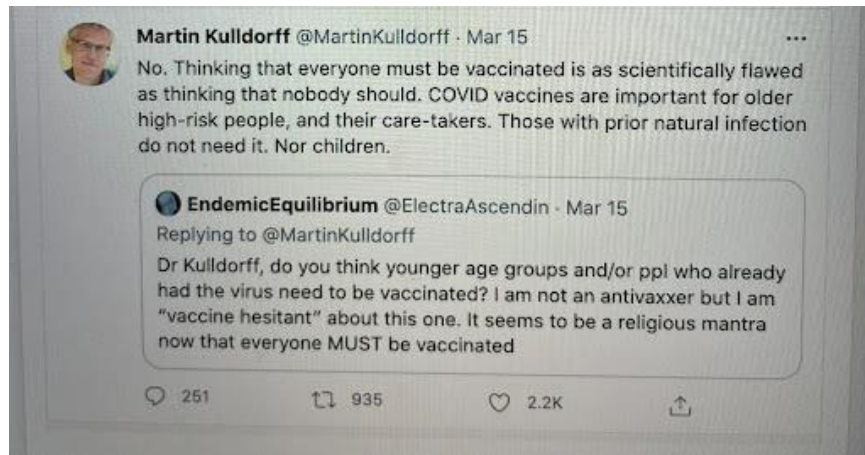
Question	Recommendation
[1] Is this claim about any of the following topics? The COVID-19 Misleading Information Policy has been discontinued effective Nov 23	[5] Recommended tweet-level remediation
[2] What kind of message is being advanced regarding this topic? Select	[6] Where does the misinformation reside?
[3] Specifically, what is being said? Select	[7] Additional recommended remediation
[4] What is the author's position? Select	

17 Third, most importantly, the buck stopped with higher level employees at Twitter who chose the inputs for the bots and decision trees, and subjectively decided escalated cases and suspensions. As it is with all people and institutions, there was individual and collective bias

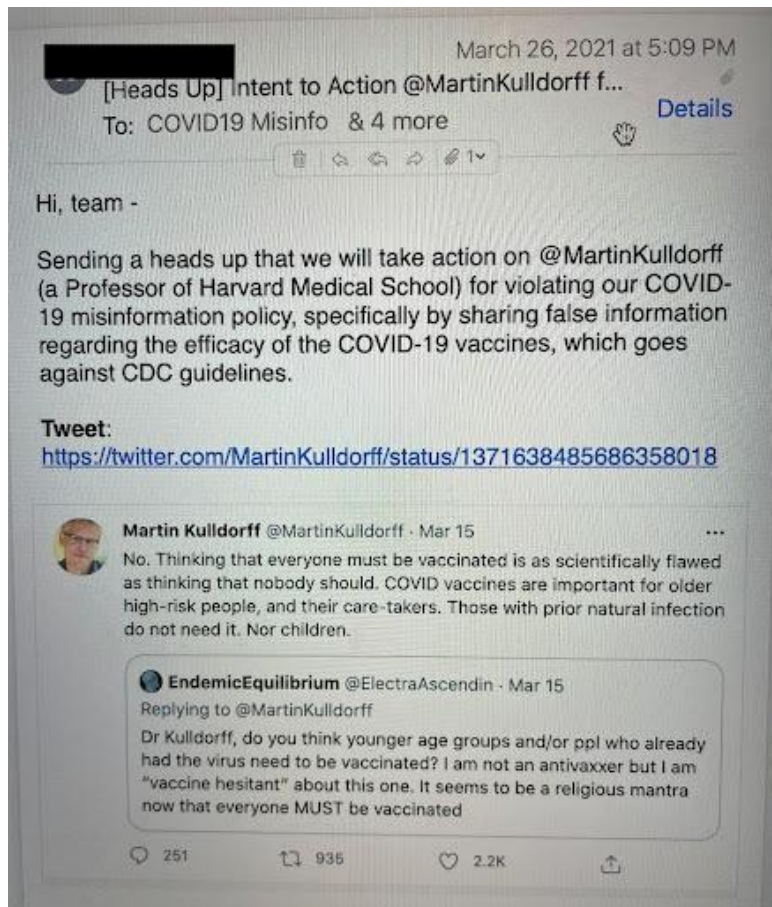
18. With Covid, this bias bent heavily toward establishment dogmas.

19. Inevitably, dissident yet legitimate content was labeled as misinformation, and the accounts of doctors and others were suspended both for tweeting opinions and demonstrably true information.

20. Exhibit A: Dr. Martin Kulldorff, an epidemiologist at Harvard Medical School, tweeted views at odds with US public health authorities and the American left, the political affiliation of nearly the entire staff at Twitter.



21. Internal emails show an "intent to action" by a moderator, saying Kulldorff's tweet violated the company's Covid-19 misinformation policy and claimed he shared "false information."



22. But Kulldorff's statement was an expert's opinion—one which also happened to be in line with vaccine policies in numerous other countries. Yet it was deemed “false information” by Twitter moderators merely because it differed from CDC guidelines.

23. After Twitter took action, Kulldorff's tweet was slapped with a “Misleading” label and all replies and likes were shut off, throttling the tweet's ability to be seen and shared by many people, the ostensible core function of the platform:



24. In my review of internal files, I found countless instances of tweets labeled as “misleading” or taken down entirely, sometimes triggering account suspensions, simply because they veered from CDC guidance or differed from establishment views.

25. A tweet by @KelleyKga, a self-proclaimed public health fact checker, with 18K followers, was flagged as "Misleading," and replies and likes disabled, even though it displayed the CDC's *own data.*

Kelley K
@KelleyKga

...

Replying to [@greg_travis](#) [@Mssarahmssarah](#) and 11 others

What an excellent example of cherry picking! If you narrow it down to only the specific months you specify, which include the largest Covid wave (seen across the world), AND you ignore all non-disease deaths, AND you ignore cancer, heart disease, SIDS, then Covid is "leading".

Single Year Ages: 18 categories tabulated in this table in from time period: Year/Month/Day, 2021 (provisional) through 2022 (provisional and partial)

ICD-10 Code	Deaths	Population	Crude Rate Per 100,000
U071 (condition originating in the period)	5,585	Not Applicable	Not Applicable
U072 (date not shown due to 6 month lag to account for delays in death certificate completion for certain causes of death)	3,219	Not Applicable	Not Applicable
U073 (symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified)	4,512	Not Applicable	Not Applicable
U074 (congenital malformations, deformations and chromosomal abnormalities)	2,242	Not Applicable	Not Applicable
U075 (all other diseases)	1,272	Not Applicable	Not Applicable
U076 (malignant neoplasms)	323	Not Applicable	Not Applicable
U077 (accidents)	324	Not Applicable	Not Applicable
U078 (major cardiovascular diseases)	243	Not Applicable	Not Applicable
U079 (neurotrauma)	200	Not Applicable	Not Applicable
U080 (other and unspecified infectious and parasitic diseases and their sequelae)	482	Not Applicable	Not Applicable
U081 (trauma)	388	Not Applicable	Not Applicable
U082 (leukemia)	283	Not Applicable	Not Applicable
U083 (other neoplasms)	279	Not Applicable	Not Applicable
U084 (diseases of heart)	242	Not Applicable	Not Applicable
U085 (other heart diseases)	195	Not Applicable	Not Applicable
U086 (all other and unspecified malignant neoplasms)	214	Not Applicable	Not Applicable
U087 (leukemia)	201	Not Applicable	Not Applicable
U088 (all other forms of heart disease)	200	Not Applicable	Not Applicable
U089 (malignant neoplasms of meningis, brain and other parts of central nervous system)	191	Not Applicable	Not Applicable
U090 (leukemia)	207	Not Applicable	Not Applicable
U091 (malignant self-harm)	204	Not Applicable	Not Applicable
U092 (malignant neoplasms of lymphoid, hematopoietic and related tissue)	204	Not Applicable	Not Applicable
U093 (leukemia)	207	Not Applicable	Not Applicable
U094 (other and unspecified neurotrauma accidents and their sequelae)	224	Not Applicable	Not Applicable
U095 (other diseases of respiratory system)	184	Not Applicable	Not Applicable
U096 (cardiovascular diseases)	180	Not Applicable	Not Applicable
U097 (trauma)	174	Not Applicable	Not Applicable
U098 (infections and parasitosis)	120	Not Applicable	Not Applicable
U099 (neoplasms)	117	Not Applicable	Not Applicable

Misleading

Learn why health officials consider COVID-19 vaccines safe for most people.
[Find out more](#)

This Tweet can't be replied to, shared or liked.

10:43 AM · Aug 11, 2022

[View Quote Tweets](#)

🗨️
🔄
❤️
📤

26. Internal records showed that a bot had flagged the tweet, and that it received many “tattles” (what the system amusingly called reports from users). That triggered a manual review by a human who— despite the tweet showing actual CDC data—nevertheless labeled it “Misleading”

27. Tellingly, the tweet by [@KelleyKga](#) that was labeled “Misleading” was a reply to a tweet that contained actual misinformation.

Covid has never been the leading cause of death from disease in children. Yet that tweet remains on the platform, and without a “misleading” label.



Censored Gregory Travis @greg_travis · Aug 11

Replying to @KelleyKga @Mssarahmssarah and 11 others

As do I

Here's some data:



Since December of 2021 COVID has been the leading cause of death from disease in children

1

1

2

1



Kelley K

@KelleyKga

Replying to @greg_travis @Mssarahmssarah and 11 others

What an excellent example of cherry picking! If you narrow it down to only the specific months you specify, which include the largest Covid wave (seen across the world), AND you ignore all non-disease deaths, AND you ignore cancer, heart disease, SIDS, then Covid is "leading".

Single Year Ages 14 categories selected (data in this table is from time period: Year/Month: Dec, 2021 (provisional) through 2022 (provisional and partial))

ICD-10 CM code	Deaths	Population	Crude Rate Per 100,000
U01 (COVID-19)	3,884	Not Applicable	Not Applicable
U02 (COVID-19)	3,219	Not Applicable	Not Applicable
Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified (800-899)	3,219	Not Applicable	Not Applicable
Mental and behavioral disorders, ICD-10-CM (F00-F99)	2,748	Not Applicable	Not Applicable
All other diseases (Residual)	1,202	Not Applicable	Not Applicable
Intentional self-harm (S00-S99)	821	Not Applicable	Not Applicable
Accidents (unintentional injuries) (W00-W99)	804	Not Applicable	Not Applicable
Major cardiovascular diseases (I00-I99)	743	Not Applicable	Not Applicable
Nontransport accidents (E80-E85, E90)	554	Not Applicable	Not Applicable
Other and unspecified infections and parasitic diseases and their sequelae (B00-B99, C00-C99, D00-D99, E00-E09, G00-G09, H00-H99, I00-I99, J00-J99, K00-K99, L00-L99, M00-M99, N00-N99, O00-O99, P00-P99, Q00-Q99, R00-R99, S00-S99, T00-T99, U00-U99, V00-V99, W00-W99, X00-X99, Y00-Y99, Z00-Z99)	462	Not Applicable	Not Applicable
Transport accidents (W00-W99)	388	Not Applicable	Not Applicable
Assault (Intentional self-harm) (T00-T99, Y87.1)	362	Not Applicable	Not Applicable
Motor vehicle accidents (V00-V99, S00-S99, T00-T99, Y87.1, Y87.2, Y87.3, Y87.4, Y87.5, Y87.6, Y87.7, Y87.8, Y87.9, Y87.0, Y87.1, Y87.2, Y87.3, Y87.4, Y87.5, Y87.6, Y87.7, Y87.8, Y87.9, Y87.0)	374	Not Applicable	Not Applicable
Diseases of heart (I00-I99, I10, I11, I12, I13)	348	Not Applicable	Not Applicable
Other heart diseases (I00-I99)	324	Not Applicable	Not Applicable
All other and unspecified malignant neoplasms (C00-C99, D00-D99, E00-E09, G00-G09, H00-H99, I00-I99, J00-J99, K00-K99, L00-L99, M00-M99, N00-N99, O00-O99, P00-P99, Q00-Q99, R00-R99, S00-S99, T00-T99, Y87.1, Y87.2, Y87.3, Y87.4, Y87.5, Y87.6, Y87.7, Y87.8, Y87.9, Y87.0)	324	Not Applicable	Not Applicable
Microbial (ICD-10)	292	Not Applicable	Not Applicable
All other forms of heart disease (I10-I19, I20-I29, I30-I39, I40-I49)	288	Not Applicable	Not Applicable
Malignant neoplasms of esophagus, trachea and other parts of central nervous system (C00-C99)	281	Not Applicable	Not Applicable
Assault (Intentional self-harm) by discharge of firearm (T00-T99, Y87.1)	267	Not Applicable	Not Applicable
Intentional self-harm (Intentional self-harm) (T00-T99, Y87.1)	262	Not Applicable	Not Applicable
Malignant neoplasms of lymphoid, hematopoietic and related tissue (C80-C96)	254	Not Applicable	Not Applicable
Luxations (S10-S19)	237	Not Applicable	Not Applicable
Other and unspecified nontransport accidents and their sequelae (E80-E85, E90, E91, E92, E93, E94, E95, E96, E97, E98, E99, F00-F99, G00-G09, H00-H99, I00-I99, J00-J99, K00-K99, L00-L99, M00-M99, N00-N99, O00-O99, P00-P99, Q00-Q99, R00-R99, S00-S99, T00-T99, Y87.1, Y87.2, Y87.3, Y87.4, Y87.5, Y87.6, Y87.7, Y87.8, Y87.9, Y87.0)	224	Not Applicable	Not Applicable
Other diseases of respiratory system (J00-J99, K00-K99, L00-L99, M00-M99)	224	Not Applicable	Not Applicable
Unintentional self-harm (Intentional self-harm) (T00-T99, Y87.1)	220	Not Applicable	Not Applicable
Intentional self-harm (Intentional self-harm) by other and unspecified means and their sequelae (T00-T99, Y87.2, Y87.3, Y87.4, Y87.5, Y87.6, Y87.7, Y87.8, Y87.9, Y87.0)	211	Not Applicable	Not Applicable
Infections and parasitoses (B00-B99)	110	Not Applicable	Not Applicable
Septicemia (A40-A41)	117	Not Applicable	Not Applicable

Misleading

Learn why health officials consider COVID-19 vaccines safe for most people. [Find out more](#)

This Tweet can't be replied to, shared or liked.

10:43 AM · Aug 11, 2022

28. Whether by humans or algorithms, content that was contrarian but true was still subject to getting flagged or suppressed

This tweet was labeled “Misleading,” even though the owner of this account, [@_euzebiusz_](#), a physician, was referring to the results of a published study

Infectious Disease Ethics
@ID_ethics

Below, the first study I have seen showing that mRNA vaccines in young people (age 16-39) are associated with cardiac arrests (not just myocarditis)

First vaccine dose: ●
Second vaccine dose: ●
Cardiac arrests: ●
Covid cases: ●

[nature.com/articles/s4159...](https://www.nature.com/articles/s4159...)

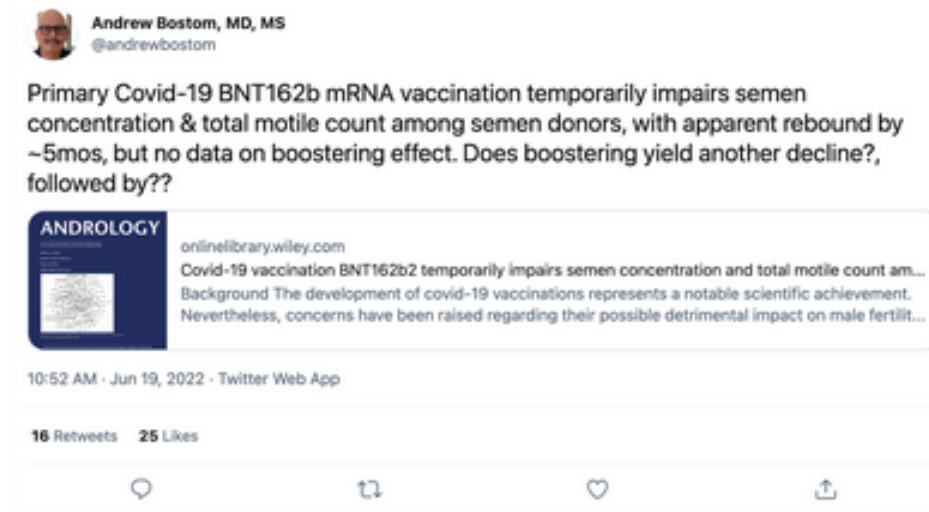
B Ages: 16-39

The graph displays two y-axes: the left axis represents the 'Number of Cardiac arrest calls (x1)' ranging from 6 to 14, and the right axis represents 'Vaccine doses (50,000 doses) and COVID-19 cases (110,000 cases)' ranging from 0 to 7. The x-axis shows the 'Date (week's starting date)' from 2020-10-18 to 2021-05-14. Key events are marked with vertical dashed lines: 'Vaccinations first began' (2020-12-19), '3rd Public health advisory (12-24-2020 to 3-7-2021)' (2021-03-07), and 'Suspected single vaccine doses for recovered individuals after 4-5-2021 (2,500 doses)' (2021-04-11). The legend includes: Cardiac arrest calls (red line), 1st vaccine doses (50,000 doses) (purple line), 2nd vaccine doses (50,000 doses) (blue line), 2nd vaccine doses shifted backwards in time by 3-weeks (50,000 doses) (dotted blue line), Suspected single vaccine doses for recovered individuals after 4-5-2021 (2,500 doses) (green line), COVID-19 cases (110,000 cases) (grey line), and 3rd Public health advisory (12-24-2020 to 3-7-2021) (dotted red line).

⚠ Misleading
Learn why health officials consider COVID-19 vaccines safe for most people.
[Find out more](#)

This Tweet can't be replied to, shared or liked.

29. Andrew Bostom, a Rhode Island physician, was permanently suspended from Twitter after receiving multiple strikes for misinformation. One of his strikes was for a tweet referring to the results from a peer reviewed study on mRNA vaccines.



30. A review of Twitter log files revealed that an internal audit, conducted after Bostom's attorney contacted Twitter, found that only 1 of Bostom's 5 violations were valid.

We have re-reviewed each of the strikes and collaborated with Site Integrity to validate our findings. Of the 5 violations, 4 of them are not invalid and are not in violation, and only this tweet is in violation: 1519678640178577410 of misinformation. As a result, we have now unsuspended the user and responded to their appeal to let them know that we have rescinded their suspension. Also, we have left a pinned note on the account to ensure that before any enforcement action will be taken against the account, it will be subject for review by Site Integrity. We will also remove the annotations on the tweets we have marked as no vio.

31. The one Bostom tweet found to still be in violation cited data that was legitimate but inconvenient to the public health establishment's narrative about the risks of flu versus Covid in children.

Andrew Bostom, MD, MS
@andrewbostom

Follow

U.S. & local Rhode Island data: Influenza is more lethal than covid-19 in children, while covid-19 vaccination causes >>> serious morbidity than influenza vaccination in children

Comparison of Influenza (10-Years) Versus Covid-19 Myocarditis/Myopericarditis Hospitalizations Among 6 to 17 Year Olds

Years Administered	Vaccine-Related Total Hospitalizations	Myocarditis
2010-20	415*	
2021-22	1790	

Distribution of death certificates with COVID-19 diagnosis

Age Group	No. of death certificates	COVID-19 diagnosis	COVID-19 diagnosis	COVID-19 diagnosis	COVID-19 diagnosis
0-4	10	1	1	1	1
5-9	10	1	1	1	1
10-14	10	1	1	1	1
15-19	10	1	1	1	1
20-24	10	1	1	1	1
25-29	10	1	1	1	1
30-34	10	1	1	1	1
35-39	10	1	1	1	1
40-44	10	1	1	1	1
45-49	10	1	1	1	1
50-54	10	1	1	1	1
55-59	10	1	1	1	1
60-64	10	1	1	1	1
65-69	10	1	1	1	1
70-74	10	1	1	1	1
75-79	10	1	1	1	1
80-84	10	1	1	1	1
85-89	10	1	1	1	1
90-94	10	1	1	1	1
95-99	10	1	1	1	1
100	10	1	1	1	1

7:03 AM - 28 Apr 2022

32. That this tweet was not only flagged by a bot, but its violation manually affirmed by a staff member is telling of both the algorithmic and human bias at play. Bostom's account was suspended for months and was finally restored on Christmas Day.

33. Another example of human bias run amok was the reaction to this tweet by Trump. Many Trump tweets led to extensive internal debates, and this one was no different.

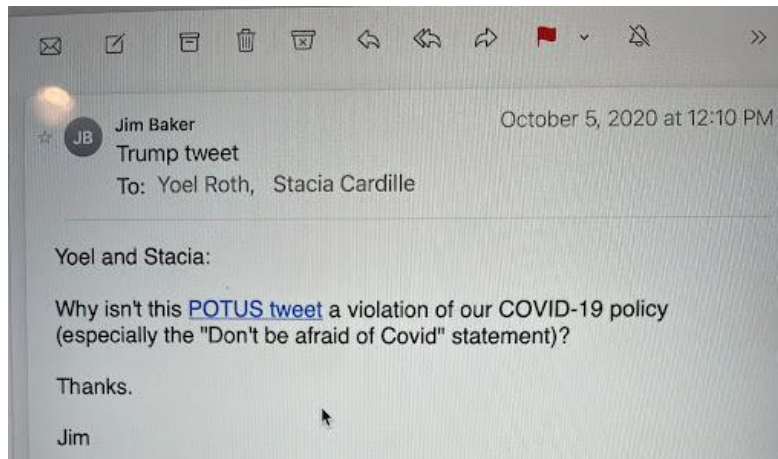
Donald J. Trump
@realDonaldTrump

I will be leaving the great Walter Reed Medical Center today at 6:30 P.M. Feeling really good! Don't be afraid of Covid. Don't let it dominate your life. We have developed, under the Trump Administration, some really great drugs & knowledge. I feel better than I did 20 years ago!

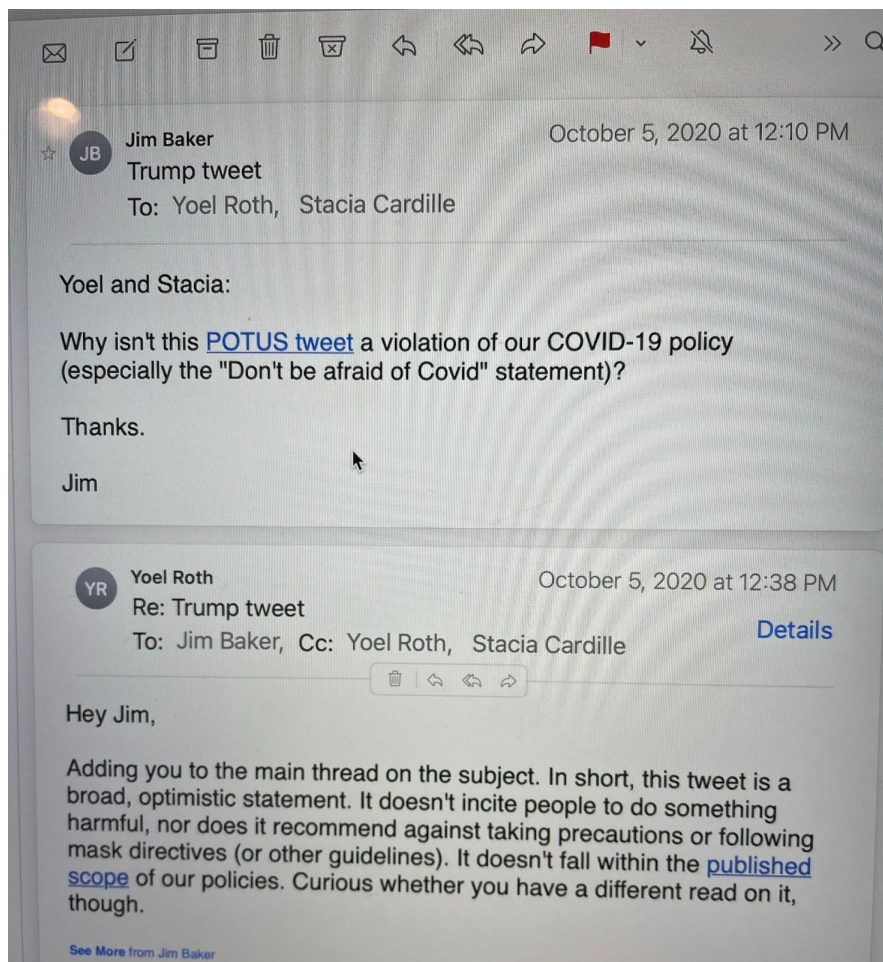
2:37 PM - Oct 5, 2020

82K Retweets 152.2K Quote Tweets 443.5K Likes

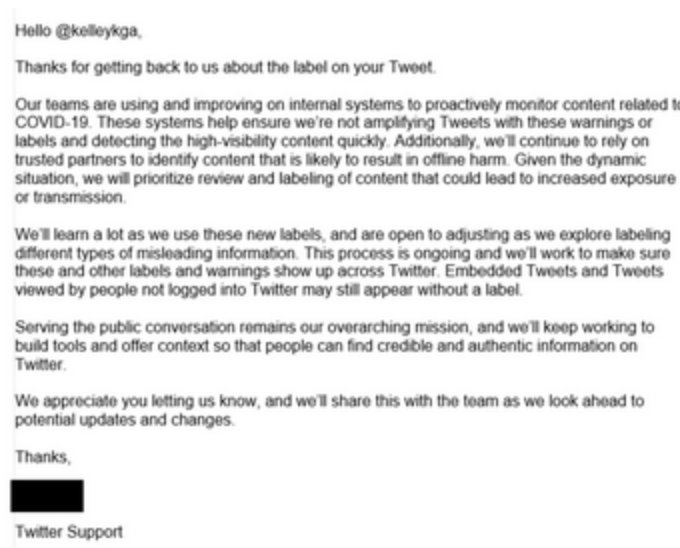
34. In a surreal exchange, Jim Baker, at the time Twitter's Deputy General Counsel, asks why telling people to not be afraid wasn't a violation of Twitter's Covid-19 misinformation policy.



35. Yoel Roth, Twitter's former head of Trust & Safety, had to explain that optimism wasn't misinformation.



36. Remember @KelleyKga with the CDC data tweet? Twitter’s response to her is clarifying: “we will prioritize review and labeling of content that could lead to increased exposure or transmission.”



Hello @kelleykga,

Thanks for getting back to us about the label on your Tweet.


Our teams are using and improving on internal systems to proactively monitor content related to COVID-19. These systems help ensure we're not amplifying Tweets with these warnings or labels and detecting the high-visibility content quickly. Additionally, we'll continue to rely on trusted partners to identify content that is likely to result in offline harm. Given the dynamic situation, we will prioritize review and labeling of content that could lead to increased exposure or transmission.

We'll learn a lot as we use these new labels, and are open to adjusting as we explore labeling different types of misleading information. This process is ongoing and we'll work to make sure these and other labels and warnings show up across Twitter. Embedded Tweets and Tweets viewed by people not logged into Twitter may still appear without a label.

Serving the public conversation remains our overarching mission, and we'll keep working to build tools and offer context so that people can find credible and authentic information on Twitter.

We appreciate you letting us know, and we'll share this with the team as we look ahead to potential updates and changes.

Thanks,



Twitter Support

37. Twitter made a decision, via the political leanings of senior staff, and govt pressure, that the public health authorities’ approach to the pandemic – prioritizing mitigation over other concerns – was “The Science” . . .

38. Information that challenged that view, such as showing harms of vaccines, or that could be perceived as downplaying the risks of Covid, especially to children, was subject to moderation, and even suppression. No matter whether such views were correct or adopted abroad.

39. What might this pandemic and its aftermath have looked like if there had been a more open debate on Twitter and other social media platforms—not to mention the mainstream press—about the origins of Covid, about lockdowns, about the true risks of Covid in kids, and much more?

40. Thanks to @ShellenbergerMD, @lwoodhouse, @lhfang and the team @thefp for their help reporting this story.

41. An expanded version of this thread is available now @thefp!
<https://www.thefp.com/p/how-twitter-rigged-the-covid-debate>