



A face mask is seen in front of the New York Stock Exchange on May 26, 2020. (Johannes Eisele via Getty Images)

How Twitter Rigged the Covid Debate

The platform suppressed true information from doctors and public-health experts that was at odds with U.S. government policy.

By David Zweig
Monday, Dec 26, 2022



*By the time reporter David Zweig got to the 10th floor conference room at Twitter's headquarters on Market Street in San Francisco, the story of the platform's suppression of true information was already international news. **Matt Taibbi**, **Michael***

Subscribe

Shneenberger, Leighton Wooanouse, Abigail Snrier, Lee Rang and I naa

revealed evidence of hidden blacklists of Twitter users; the way Twitter acted as a kind of FBI subsidiary; and how company executives rewrote the platform's policies on the fly to accommodate political bias and pressure.

What we had yet to crack was the story of Covid.

David has spent three years reporting on Covid—specifically the underlying science, or lack thereof, behind many of our nation's policies. For years he had noticed and criticized a bias not only in the mainstream media's

coverage of the pandemic, but also in the way it was presented on platforms like Twitter.

We couldn't think of anyone better to tackle this story. — BW



I had always thought a primary job of the press was to be skeptical of power—especially the power of the government. But during the Covid-19 pandemic, I and so many others found that the legacy media had shown itself to largely operate as a messaging platform for our public health institutions. Those institutions operated in near total lockstep, in part by purging internal dissidents and discrediting outside experts.

Twitter became an essential alternative. It was a place where those with public health expertise and perspectives at odds with official policy could air their views—and where curious citizens could find such information.

This often included other countries' responses to Covid that differed dramatically from our own.

But it quickly became clear that Twitter also seemed to promote content that reinforced the establishment narrative, and to suppress views and even scientific evidence that ran to the contrary.

Was I imagining things? Was the pattern I and others witnessed proof of

purposeful intent? An algorithm gone rogue? Or something else? In other words: When it came to Covid, and the information shared on a service used by hundreds of millions of people, what exactly was being amplified? And what was being banned or censored?

So when *The Free Press* asked if I would go to Twitter to peek behind the curtain, I took the first flight out of New York.

Here's what I found.



The United States government pressured Twitter to elevate certain content and suppress other content about Covid-19 and the pandemic. Internal emails that I viewed at Twitter showed that both the Trump and

Biden administrations directly pressed Twitter executives to moderate the platform's content according to their wishes.

At the onset of the pandemic, the Trump administration was especially concerned about panic buying, and sought "help from the tech companies to combat misinformation," according to emails sent by Twitter employees in the wake of meetings with the White House. One area of so-called misinformation: "runs on grocery stores." The trouble is that it wasn't misinformation: There actually were runs on goods.

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

When the Biden administration took over, its agenda for the American people can be summed up as: Be very afraid of Covid and do exactly what we say to stay safe.

In July 2021, then-U.S. Surgeon General Vivek Murthy released a 22-page advisory concerning what the World Health Organization referred to as an "infodemic" and called on social media platforms to do more to shut

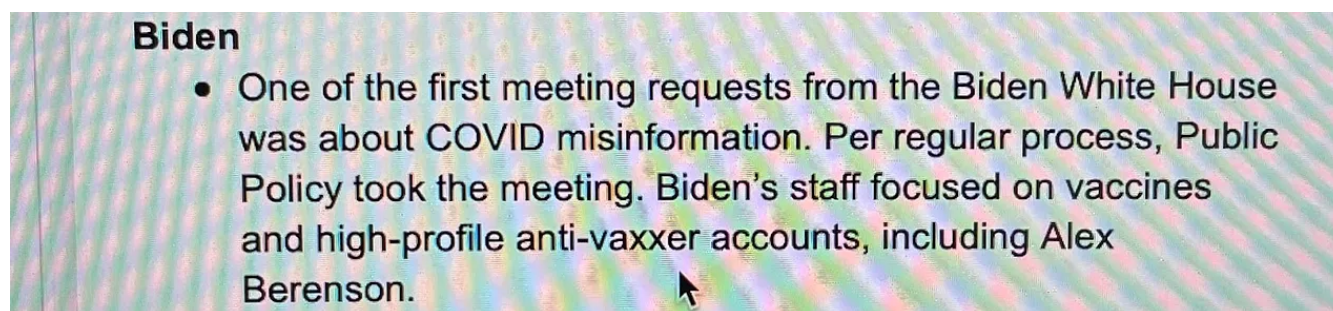
an memo, and called on social media platforms to do more to shut down “misinformation.”

“We are asking them to step up,” Murthy said. “We can’t wait longer for them to take aggressive action.”

That’s the message the White House had already taken directly to Twitter executives in private channels. One of the Biden administration’s first meeting requests was about Covid, with a focus on “anti-vaxxer

accounts,” according to a meeting summary by Lauren Culbertson, Twitter’s Head of U.S. Public Policy.

They were especially concerned about Alex Berenson, a journalist skeptical of lockdowns and mRNA vaccines, who had hundreds of thousands of followers on the platform:

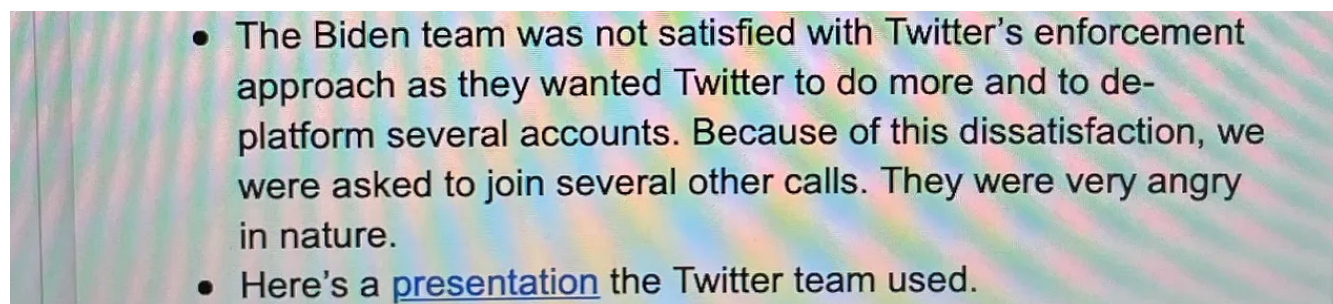


By the summer of 2021, the day after Murthy’s memo, Biden announced publicly that social media companies were “killing people” by allowing misinformation about vaccines. Just hours later, Twitter [locked Berenson](#) out of his account, and then permanently suspended him the next month. Berenson sued Twitter. He ultimately settled with the company, and is now back on the platform. As part of the lawsuit, Twitter was compelled to provide certain [internal communications](#). They revealed that the White House had directly met with Twitter employees and pressured them to take action on Berenson.

The summary of meetings by Culbertson, emailed to colleagues in December 2022, adds new evidence of the White House’s pressure campaign, and illustrates how it tried to directly influence what content was allowed on Twitter.

Culbertson wrote that the Biden team was “very angry” that Twitter had not been more aggressive in deplatforming multiple accounts. They

not seem more aggressive in deprioritizing multiple accounts. They wanted Twitter to do more.



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

But Twitter did suppress views—and not just those of journalists like Berenson. Many medical and public health professionals who expressed perspectives or even cited findings from accredited academic journals that conflicted with official positions were also targeted. As a result, legitimate findings and questions about our Covid policies and their consequences went missing.

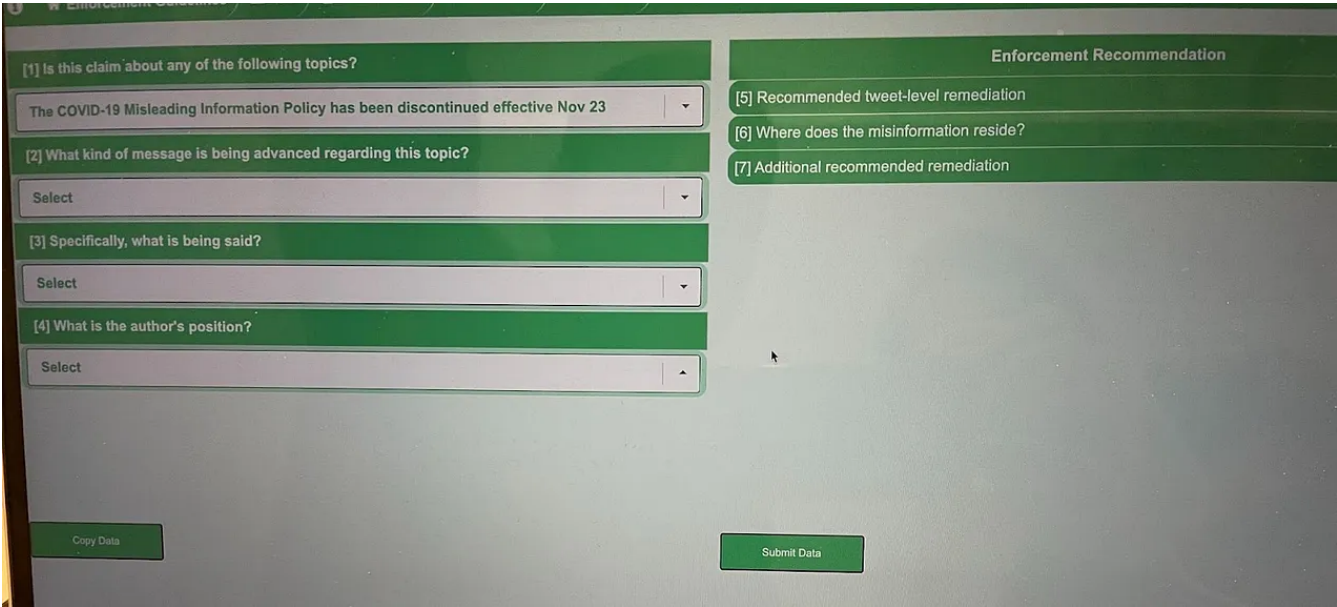
There were three serious problems with Twitter's process.

First: Much of the content moderation on Covid, to say nothing of other contentious subjects, was conducted by bots trained on machine learning and AI. I spent hours discussing the systems with an engineer and with an executive who had been at the company for more than a year before Musk's takeover. They explained the process in basic terms: Initially, the bots were fed information to train them on what to look for—but their searches would become more refined over time both as they scanned the platform and as they were manually updated with additional chosen inputs. At least that was the premise. Though impressive in their engineering, the bots would prove too crude for such nuanced work. When you drag a digital trawler across a social media platform, you're not just catching cheap fish, you're going to snag dolphins along the way.

Second: Contractors operating in places like the Philippines were also moderating content. They were given decision trees to aid in their 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 notion that remote workers, sitting in distant cube farms, were going to police medical information to this granular degree is absurd on its face.

Embedded below is an example template—deactivated after Musk’s arrival—of the decision tree tool that contractors used. The contractor would run through a series of questions, each with a drop down menu, ultimately guiding them to a predetermined conclusion.



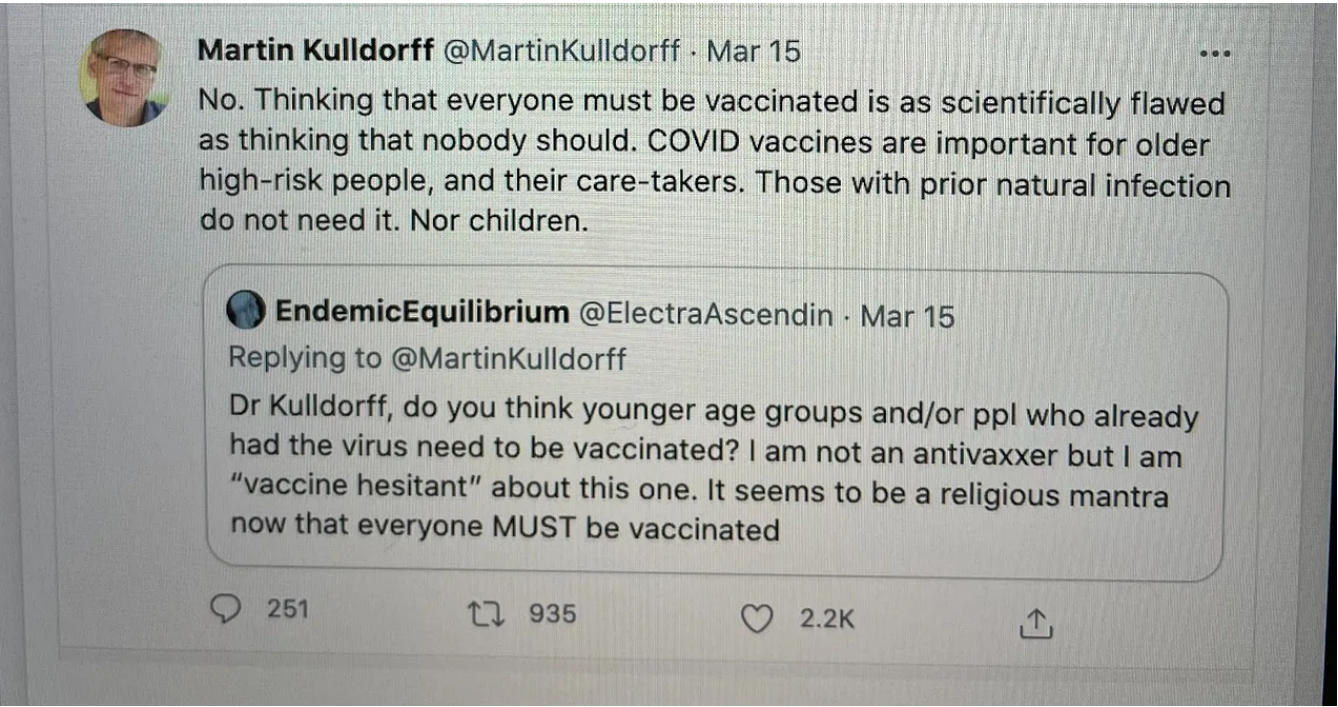
Third: Most importantly, the buck stopped with higher level employees at Twitter. They chose the inputs for the bots and decision trees. They determined suspensions. And as is the case with all people and institutions, there was both individual and collective bias.

At Twitter, Covid-related bias bent heavily toward establishment dogmas. 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.

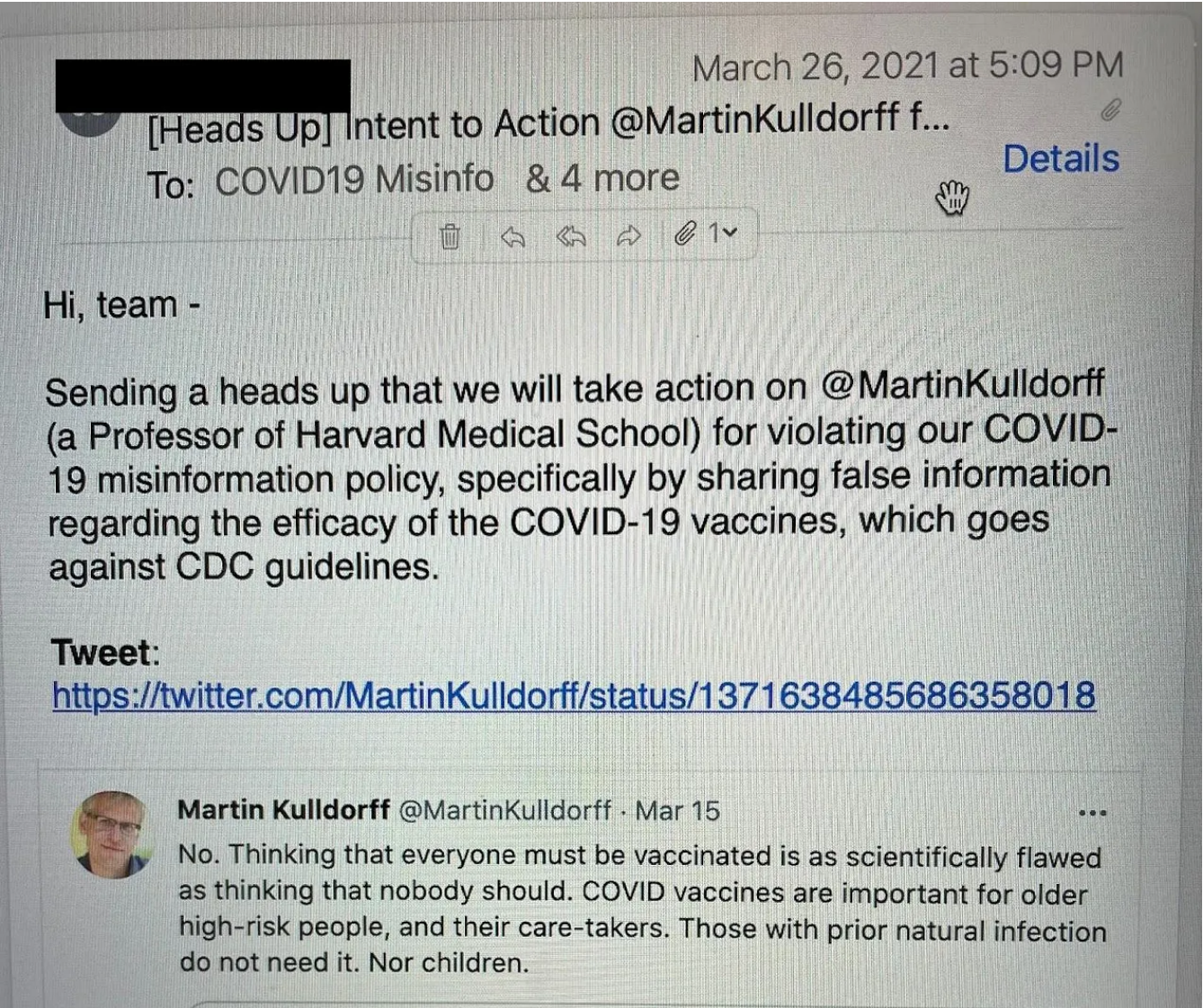


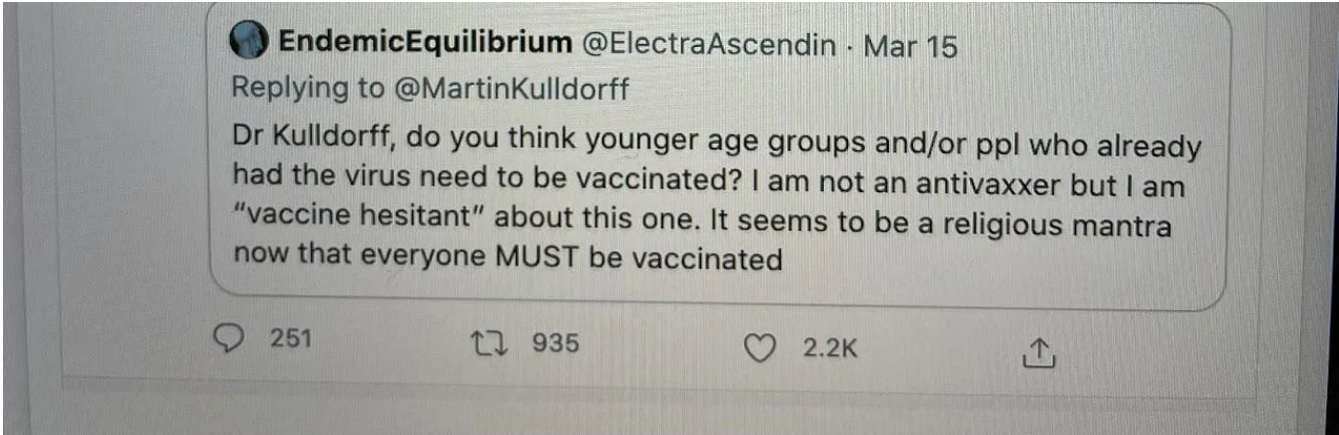
Take, for example, Martin Kulldorff, an epidemiologist at Harvard Medical School. Dr. Kulldorff often tweeted views at odds with U.S. public health authorities and the American left, the political affiliation of nearly the entire staff at Twitter.

Here is one such tweet, from March 15, 2021, regarding vaccination.



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





But Kulldorff’s statement was an expert’s opinion—one that 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. 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 others, a core function of the platform.



View Quote Tweets



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

For example, a tweet by @KelleyKga, a self-proclaimed public health fact checker with more than 18,000 followers, was flagged as “misleading,” and replies and likes disabled, for showing that Covid was not the leading cause of death in children, even though it cited 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 selectedData in this table is from time period:: Year/Month: Dec., 2021 (provisional) through 2022 (provisional and partial)			
UCD - ICD-10 113 Cause List	Deaths	Population	Crude Rate Per 100,000
#Certain conditions originating in the perinatal period (P00-P96)	5,686	Not Applicable	Not Applicable
Data not shown due to 6 month lag to account for delays in death certificate completion for certain causes of death (999)	3,319	Not Applicable	Not Applicable
Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified (R00-R99)	2,619	Not Applicable	Not Applicable
#Congenital malformations, deformations and chromosomal abnormalities (Q00-Q99)	2,548	Not Applicable	Not Applicable
All other diseases (Residual)	1,202	Not Applicable	Not Applicable
#Malignant neoplasms (C00-C97)	925	Not Applicable	Not Applicable
#Accidents (unintentional injuries) (V01-X59,Y85-Y86)	894	Not Applicable	Not Applicable
Major cardiovascular diseases (I00-I78)	543	Not Applicable	Not Applicable
Nontransport accidents (W00-X59,Y86)	506	Not Applicable	Not Applicable
Other and unspecified infectious and parasitic diseases and their sequelae (A00,A05,A20-A36,A42-A44,A48-A49,A54-A79,A81-A82,A85.0-A85.1,A85.8,A86-B04,B06-B09,B25-B49,B55-B99,U07.1)	462	Not Applicable	Not Applicable
Transport accidents (V01-V99,Y85)	388	Not Applicable	Not Applicable
#Assault (homicide) (*U01-*U02,X85-Y09,Y87.1)	383	Not Applicable	Not Applicable
Motor vehicle accidents (V02-V04,V09.0,V09.2,V12-V14,V19.0-V19.2,V19.4-V19.6,V20-V79,V80.3-V80.5,V81.0-V81.1,V82.0-V82.1,V83-V86,V87.0-V87.8,V88.0-V88.8,V89.0,V89.2)	379	Not Applicable	Not Applicable
#Diseases of heart (I00-I09,I11,I13,I20-I51)	348	Not Applicable	Not Applicable
Other heart diseases (I26-I51)	326	Not Applicable	Not Applicable
All other and unspecified malignant neoplasms (C17,C23-C24,C26-C31,C37-C41,C44-C49,C51-C52,C57-C60,C62-C63,C66,C68-C69,C73-C80,C97)	316	Not Applicable	Not Applicable
#COVID-19 (U07.1)	293	Not Applicable	Not Applicable
All other forms of heart disease (I26-I28,I34-I38,I42-I49,I51)	289	Not Applicable	Not Applicable
Malignant neoplasms of meninges, brain and other parts of central nervous system (C70-C72)	281	Not Applicable	Not Applicable
Assault (homicide) by discharge of firearms (*U01.4,X93-X95)	267	Not Applicable	Not Applicable
#Intentional self-harm (suicide) (*U03,X60-X84,Y87.0)	265	Not Applicable	Not Applicable
Malignant neoplasms of lymphoid, hematopoietic and related tissue (C81-C96)	264	Not Applicable	Not Applicable
Leukemia (C91-C95)	237	Not Applicable	Not Applicable

ICD-10 Code	Count	Not Applicable	Not Applicable
Other and unspecified nontransport accidents and their sequelae (W20-W31,W35-W64,W75-W99,X10-X39,X50-X59,Y86)	224	Not Applicable	Not Applicable
Other diseases of respiratory system (J00-J06,J30- J39,J67,J70-J98)	184	Not Applicable	Not Applicable
#Cerebrovascular diseases (I60-I69)	169	Not Applicable	Not Applicable
Intentional self-harm (suicide) by other and unspecified means and their sequelae (*U03,X60-X71,X75-X84,Y87.0)	151	Not Applicable	Not Applicable
#Influenza and pneumonia (J09-J18)	133	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

[View Quote Tweets](#)










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.” Tellingly, the tweet by @KelleyKga that was labeled “misleading” was a reply to a tweet that contained actual misinformation.



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

2



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 selectedData in this table is from time period:: Year/Month: Dec., 2021 (provisional) through 2022 (provisional and partial)

UCD - ICD-10 113 Cause List	Deaths	Population	Crude Rate Per 100,000
#Certain conditions originating in the perinatal period (P00-P96)	5,686	Not Applicable	Not Applicable
Data not shown due to 6 month lag to account for delays in death certificate completion for certain causes of death (999)	3,319	Not Applicable	Not Applicable
Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified (R00-R99)	2,619	Not Applicable	Not Applicable
#Congenital malformations, deformations and chromosomal abnormalities (Q00-Q99)	2,548	Not Applicable	Not Applicable
All other diseases (Residual)	1,202	Not Applicable	Not Applicable
#Malignant neoplasms (C00-C97)	925	Not Applicable	Not Applicable
#Accidents (unintentional injuries) (V01-X59,Y85-Y86)	894	Not Applicable	Not Applicable
Major cardiovascular diseases (I00-I79)	543	Not Applicable	Not Applicable
Nontransport accidents (W00-X59,Y86)	506	Not Applicable	Not Applicable
Other and unspecified infectious and parasitic diseases and their sequelae (A00,A05,A20-A36,A42-A44,A48-A49,A54-A79,A81-A82,A85.0-A85.1,A85.8,A86-B04,B06-B09,B25-B49,B55-B99,U07.1)	462	Not Applicable	Not Applicable
Transport accidents (V01-V99,Y85)	388	Not Applicable	Not Applicable
#Assault (homicide) (*U01-*U02,X85-Y09,Y87.1)	383	Not Applicable	Not Applicable
Motor vehicle accidents (V02-V04,V09.0,V09.2,V12-V14,V19.0-V19.2,V19.4-V19.6,V20-V79,V80.3-V80.5,V81.0-V81.1,V82.0-V82.1,V83-V86,V87.0-V87.8,V88.0-V88.8,V89.0,V89.2)	379	Not Applicable	Not Applicable
#Diseases of heart (I00-I09,I11,I13,I20-I51)	348	Not Applicable	Not Applicable
Other heart diseases (I26-I51)	326	Not Applicable	Not Applicable
All other and unspecified malignant neoplasms (C17,C23-C24,C26-C31,C37-C41,C44-C49,C51-C52,C57-C60,C62-C63,C66,C68-C69,C73-C80,C97)	316	Not Applicable	Not Applicable
#COVID-19 (U07.1)	293	Not Applicable	Not Applicable
All other forms of heart disease (I26-I28,I34-I38,I42-I49,I51)	289	Not Applicable	Not Applicable
Malignant neoplasms of endocrine, brain and other parts of central nervous system (C20-C23)	281	Not Applicable	Not Applicable

malignant neoplasms of meninges, brain and other parts of central nervous system (C81-C96)	267	Not Applicable	Not Applicable
Assault (homicide) by discharge of firearms (*U01.4,X93-X95)	267	Not Applicable	Not Applicable
#Intentional self-harm (suicide) (*U03,X60-X84,Y87.0)	265	Not Applicable	Not Applicable
Malignant neoplasms of lymphoid, hematopoietic and related tissue (C81-C96)	264	Not Applicable	Not Applicable
Leukemia (C91-C95)	237	Not Applicable	Not Applicable
Other and unspecified nontransport accidents and their sequelae (W20-W31,W35-W64,W75-W99,X10-X39,X50-X59,Y86)	224	Not Applicable	Not Applicable
Other diseases of respiratory system (J00-J06,J30-J39,J67,J70-J98)	184	Not Applicable	Not Applicable
#Cerebrovascular diseases (I60-I69)	169	Not Applicable	Not Applicable
Intentional self-harm (suicide) by other and unspecified means and their sequelae (*U03,X60-X71,X75-X84,Y87.0)	151	Not Applicable	Not Applicable
#Influenza and pneumonia (J09-J18)	133	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

Covid has never been the leading cause of death from disease in children. Yet that tweet not only remains on the platform, it is without any sort of “misleading” label.

Whether by humans or algorithms, content that was contrarian but true, and the people who conveyed that content, were still subject to getting flagged and suppressed.

Sometimes this was done covertly. As [reported](#) earlier by *The Free Press*, Dr. Jay Bhattacharya, a Stanford professor of health policy who argued for focused protection of the vulnerable and an end to lockdowns, was secretly put on a Trends Blacklist.

But many instances were public facing. The [author](#) of the tweet embedded below is a physician who runs the Infectious Disease Ethics Twitter account. The tweet was labeled as “misleading” even though it was referring to the results of a [peer-reviewed study](#) that found an association between the mRNA vaccines and cardiac arrests in young people in Israel.



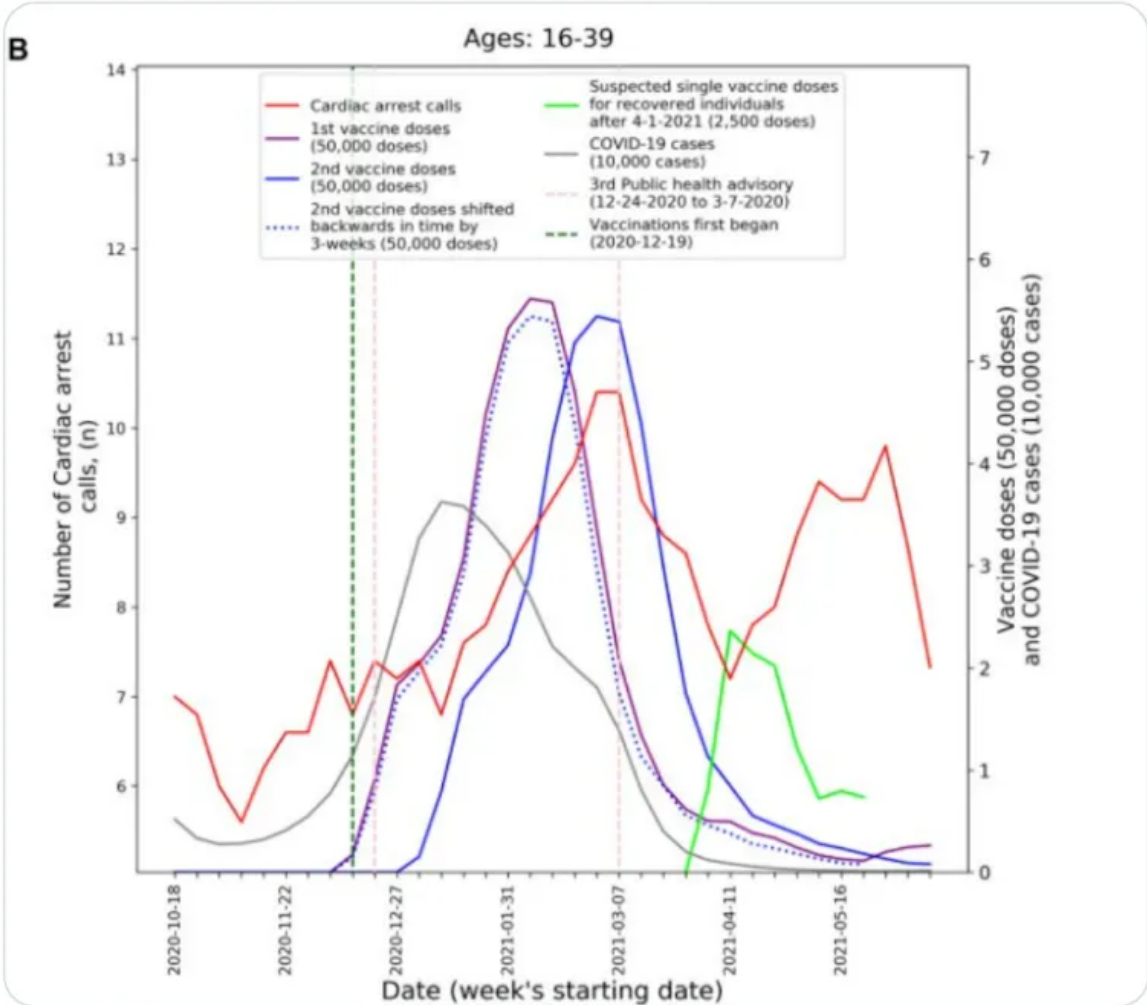
Infectious Disease Ethics
@ID_ethics

...

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/s41598-021-01111-1)



Misleading

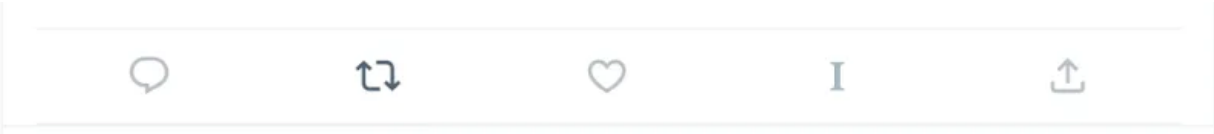
Learn why health officials consider COVID-19 vaccines safe for most people.
[Find out more](#)

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

This Tweet is from a suspended account. [Learn more](#)

11:12 PM · Apr 28, 2022

[View Quote Tweets](#)



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 referencing the results from a peer-reviewed

[study](#) that found a deterioration in sperm concentration and total motile count in sperm donors following mRNA vaccination.



Andrew Bostom, MD, MS
@andrewbostom

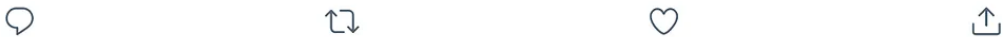
Primary Covid-19 BNT162b mRNA vaccination temporarily impairs semen concentration & total motile count among semen donors, with apparent rebound by ~5mos, but no data on boosting effect. Does boosting yield another decline?, followed by??



onlinelibrary.wiley.com
Covid-19 vaccination BNT162b2 temporarily impairs semen concentration and total motile count am...
Background The development of covid-19 vaccinations represents a notable scientific achievement. Nevertheless, concerns have been raised regarding their possible detrimental impact on male fertilit...

10:52 AM · Jun 19, 2022 · Twitter Web App

16 Retweets 25 Likes



Twitter’s logs revealed that an internal audit, conducted by Twitter after Bostom’s attorney contacted the company, found that only one of Bostom’s five 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.

The one Bostom tweet found to still be in violation of Twitter policy cited data and drew a conclusion that was totally legitimate. The problem was only that it was inconvenient to the public health

establishment’s narrative about the relative 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 recent U.S. pediatric flu (seasonal/ 1 yr. pandemic) and covid-19 deaths

Period	Age range (years)	Deaths
19, 2020	0-17	199 /129*
19, 2021	0-17	592 /385*
mic H1N1 Flu, 2009	0-17	1220
nal Flu, 2012	0-17	1161
nal Flu, 2014	0-17	803
nal Flu, 2017	0-17	526
nal Flu, 2019	0-17	486

*Adjusted for CDC 35% over counting of covid-19 pediatric deaths by death certificate re

www.cdc.gov/nchs/nvss/vsrr/covid_weekly/index.htm Covid-19: sonal flu burdens: https://www.cdc.gov/flu/about/burden/past-seasons.html; ng the Burden of 2009 Pandemic Influenza A (H1N1) in the United States (April 2009–April 2010)* ademic.oup.com/cid/article/52/suppl_1/S75/499147]

www.cdc.gov/mmwr/volumes/70/wr/mm7014e2.htm?s_cid=mm7014e2_x

. Distribution of death certificates with COVID-19 diagnosis* across five mutually exclusive categories defined by p ition of co-occurring diagnoses, by demographic characteristics, setting of death, and manner of death characteristics r Health Statistics, United States, January–December 2020

	No. (row %)
No. of death certificates	182
COVID-19 only	8 (4.4)
COVID-19 and ≥1 chain-of-event condition only	70 (38.5)
COVID-19 and ≥1 significant contributing condition only	18 (9.9)
COVID-19 and ≥1 chain-of-event and ≥1 significant contributing condition	22 (12.1)

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

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

administration of other vaccines, simultaneously; ** 1 case r vaccines; the 3rd occurred in relation to a confirmed influe s case of myocarditis"

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

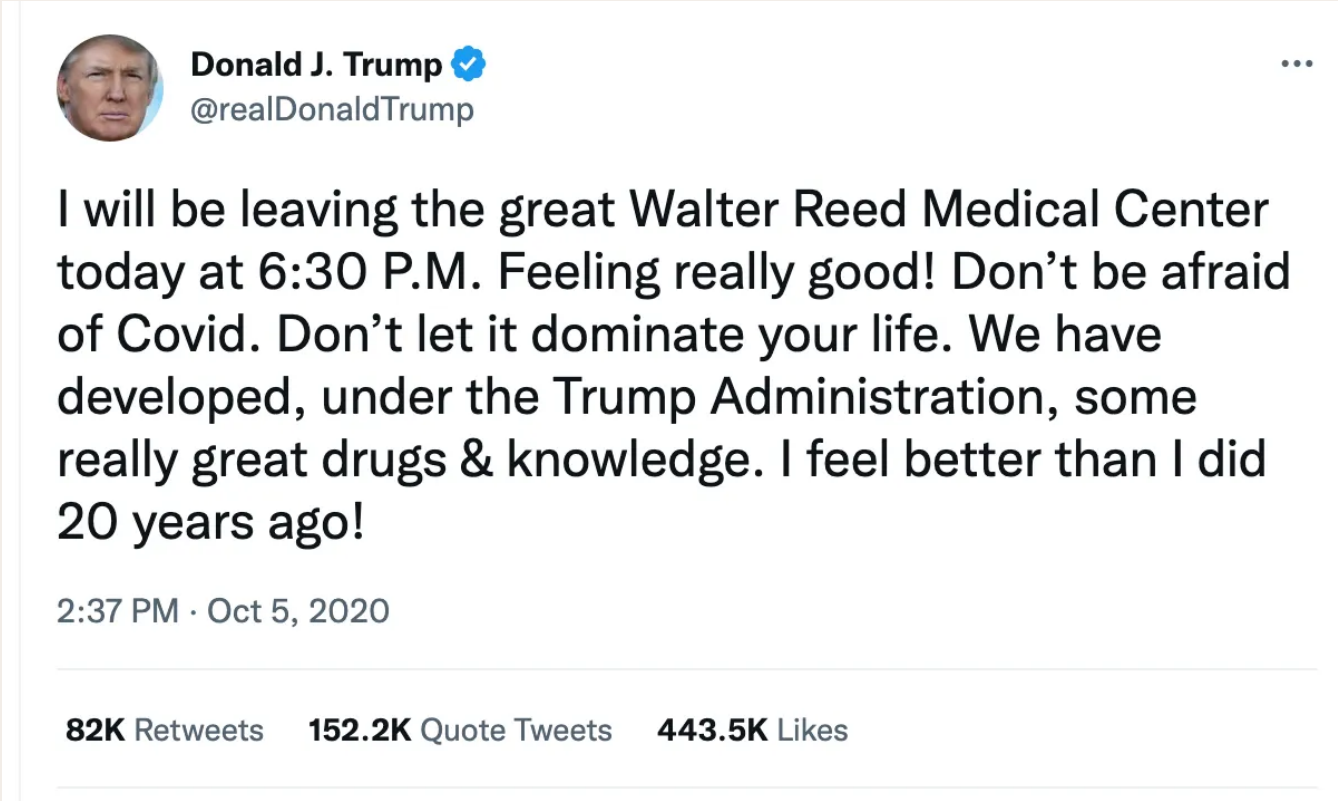
Years Administered	Vaccine-Related Total Hospitalizations	Myocarditis/Myopericarditis Hospitalizations
2010-20	1	
2021-22	7	

7:03 AM - 28 Apr 2022

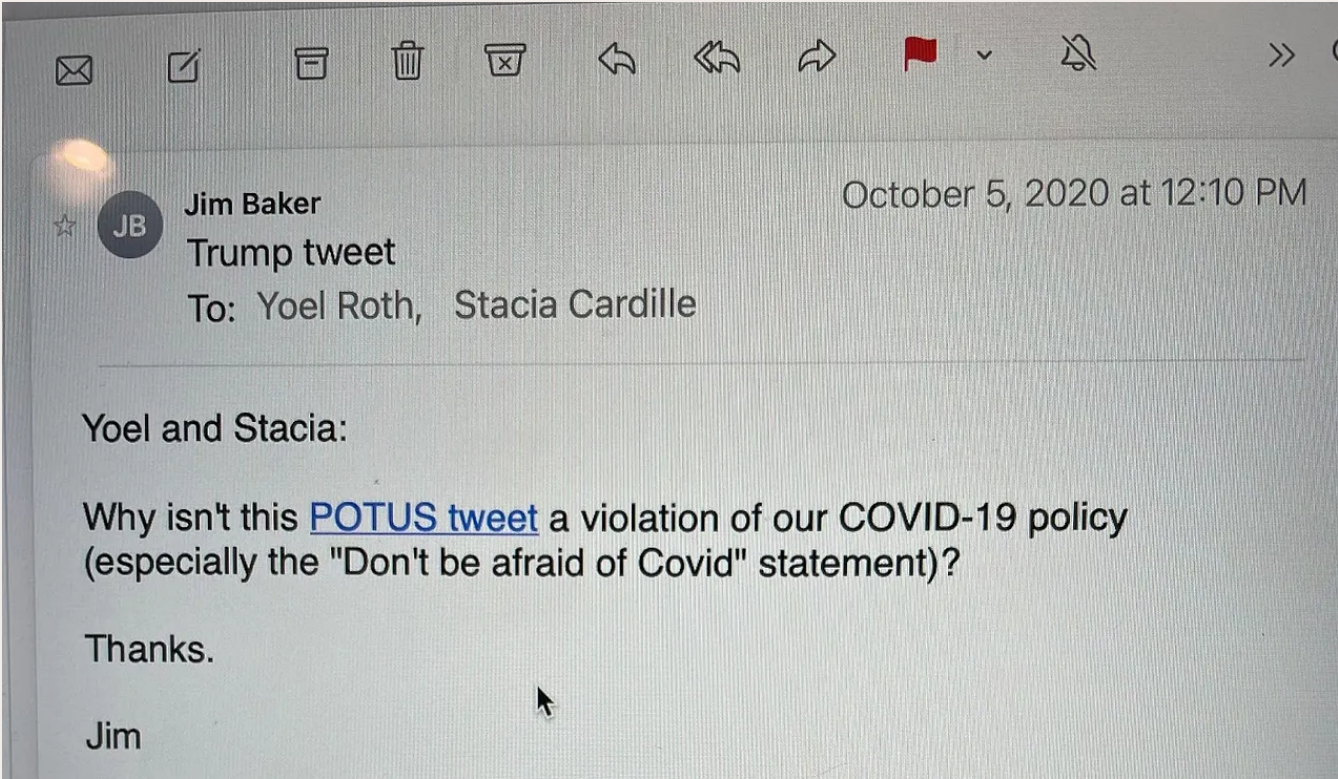
This tweet was flagged not only by a bot but also manually, by a human being—which goes a long way to illuminating both the algorithmic and human bias at Twitter. “It seems grossly unfair,” Bostom told me when I called to share with him my findings. “What’s the remedy? What am I supposed to do?” (His account was restored, along with a number of others, on Christmas Day.)

Another example of human bias run amok was the reaction to the below tweet by then-President Trump. Many Trump tweets led to extensive

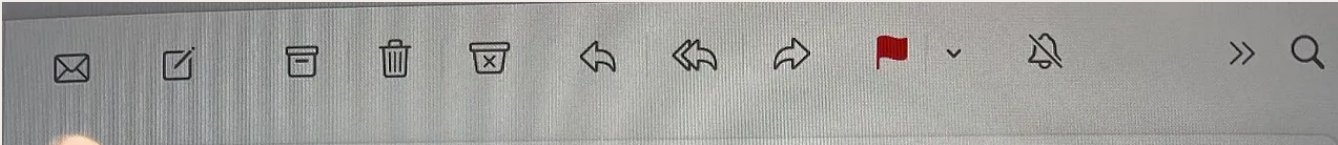
internal debates at the company, and this one was no different.

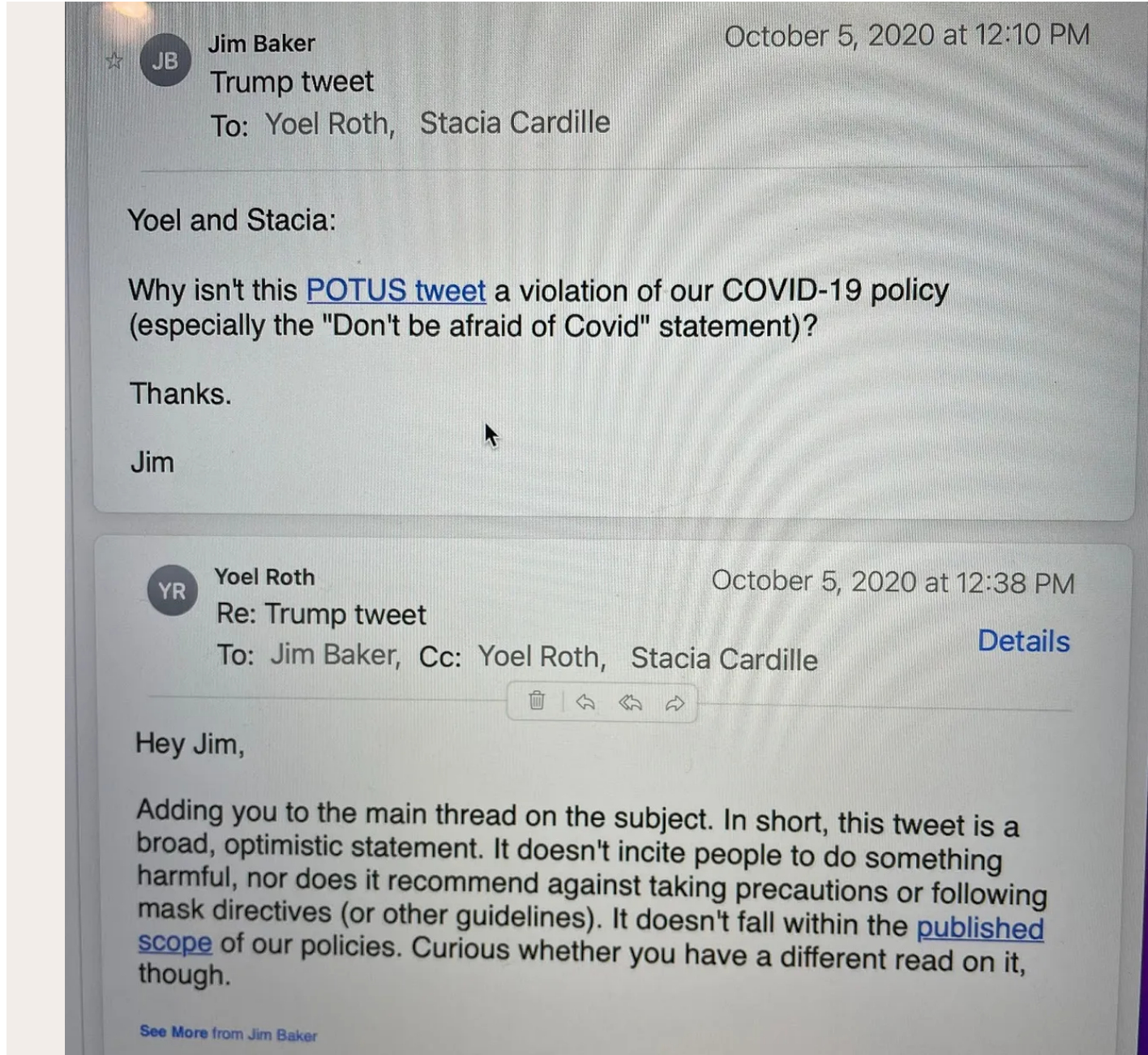


In a surreal exchange, Jim Baker, at the time Twitter’s Deputy General Counsel, asks why telling people to not be afraid wasn’t a violation of Twitter’s Covid-19 misinformation policy.



In his reply, Yoel Roth, Twitter’s former head of Trust & Safety, had to explain that optimism wasn’t misinformation.





Remember @KelleyKga with the CDC data tweet? Twitter’s response to her in an exchange about why her tweet was labeled as “misleading” is clarifying:

“We will prioritize review and labeling of content that could lead to increased exposure or transmission.”

