



The Moderation War Is Coming to Spotify, Substack, and Clubhouse

Alex Kantrowitz

January 14, 2021

Glenn Greenwald was pissed. The *Columbia Journalism Review* had just asked whether Substack should remove the writer Andrew Sullivan from its service. And having recently joined the email newsletter platform himself, Greenwald attacked.

“It was only a matter of time before people started demanding Substack be censored,” he said, taking it a step further than the *CJR*.

Last October, Greenwald left *The Intercept*, a publication he founded, claiming the publication’s editors, who previously hadn’t touched his work, “censored” him ahead of the 2020 election. So he moved to Substack, which advertises itself as a home for independent writing. Seeing a reporter point out Substack’s lack of moderation set him off.

CJR reporter Clio Chang pushed Substack to take a stance on Sullivan for a simple reason. Sullivan had previously published excerpts from *The Bell Curve*, a 1994 book that attempted to link IQ to race. Chang asked Substack’s founders whether his presence could cause other writers

to shy away from the platform. Its paid newsletters, she noted, were already very white and male at the top.

“Often, adherence to neutrality only enforces existing power structures,” Chang wrote as she considered Substack’s hands-off content moderation policies.

The *CJR* article’s stance, and Greenwald’s fiery pushback, felt familiar to those who’ve watched similar arguments regarding Facebook, Twitter, and Google in recent years. Some have criticized these platforms for leaving up too much objectionable content. Others have called the platforms censors, arguing they take down too much. The difference with Substack, of course, is the medium. It wasn’t Big Tech. It was an email platform.

Though the big social networks have been the main focus in the debate over how — and whether — online platforms should moderate user content, the fight is migrating toward smaller platforms devoid of rollicking social feeds. They include email providers like Substack, podcast platforms like Spotify (hello Joe Rogan), and nascent startups like Clubhouse.

These smaller services are coming under scrutiny now that the big platforms have warmed to aggressive moderation, culminating with Twitter, Facebook, YouTube’s suspension of President Donald Trump following the Capitol riot. The battle will only heat up now that Amazon pushed “free speech” social network Parler off the internet. Attention will move to smaller, mainstream services still figuring out their policies, as the precedent they set today could determine how they handle content moving forward.

“It’s inevitable that it’s going to follow the same pattern,” Greenwald told *OneZero*.

Substack, which declined an interview request, is unlikely to ban Sullivan, who is sixth on its leaderboard for paid political newsletters. But the dustup over the *CJR* article demonstrates the fight over these smaller platforms' souls is now in full effect. How it shakes out may well determine the future of these emerging venues on the internet.

The big shift

Substack, Spotify, and Clubhouse's current perspective on content moderation mirror how Twitter, Facebook, and Google once viewed the practice. Twitter executives initially called themselves "the free speech wing of the free speech party." Facebook insisted it had no business touching political content. YouTube allowed Alex Jones and other wingnuts to build misinformation empires on its service. Now, Substack CEO Chris Best — reflecting the smaller platforms' attitude on moderation — told *CJR* that if you're looking for him to take an "editorial position" you should find another service.

After initially resisting aggressive content moderation (aside from no-brainers like child porn), the bigger platforms have slowly relented. "They are agreeing that they need to moderate more aggressively and in more ways than they used to," Evelyn Douek, a Harvard Law School lecturer who studies content moderation, told *OneZero*. And if past is prologue, their path to the current state is worth revisiting.

As unelected businesspeople, the big platform executives once felt queasy about deciding what people could say on their services. They also, pragmatically, wanted nothing to do with the political fights that could erupt after every decision. While they had power over people's ability to speak, they preferred not to use it. Life was easier that way.

By the mid-2010s, however, their perspective began to shift. High-profile harassment, including awful attacks against Robin Williams' daughter, Zelda, plagued Twitter. The trolling caused its core users to leave, according to a post by then-CEO Dick Costolo on an internal forum in 2015. "We suck at dealing with abuse," Costolo said. He promised to do better, then left the company.

As Twitter transitioned from Costolo to new CEO Jack Dorsey in late 2015, the company was in economic peril. Its growth was anemic, and its stock price was stagnant. Then, less than a year into Dorsey's tenure, a racist, sexist Twitter mob harassed the actor Leslie Jones. "I feel like I'm in a personal hell," she tweeted before temporarily deactivating her account.

For Twitter, losing Jones as she starred in a blockbuster *Ghostbusters* film threatened to further degrade its business. It could turn public opinion against the company even more, and some advertisers along with it. If Twitter did nothing, it also risked the ire of the entire entertainment industry. "Very rarely do celebrities engage in these kind of conversations directly with Twitter executives," a former Twitter employee with knowledge of high-level discussions told me. The individual asked to remain anonymous due to confidentiality agreements. "Very large agencies like CAA or WME say we're going to tell all of our people to not do anything more on Twitter, to switch all of their activities over to Instagram, if you don't take some sort of action."

After the Jones episode, Twitter effectively killed its hands off approach. It banned Milo Yiannopoulos, the troll it saw as a ringleader of the harassment. and Dorsey started saying that fighting hate was his top priority.

Twitter took its stance on harassment amid the 2016 election. And then, after news broke that a Russian campaign to undermine the election thrived on their services, Twitter, Facebook, and

YouTube started removing posts that threatened elections' integrity, adding a new type of content they'd moderate enthusiastically along with targeted harassment.

By the time Covid hit, the big platforms' reticence to remove and label posts had largely evaporated. And when the extent of the plague became clear, they set assertive policies. Facebook prohibited posts that contradicted the WHO's guidelines. Twitter and YouTube acted similarly.

Last week, both Twitter and Facebook suspended Donald Trump's account after a mob he sent to stop Congress from certifying the election breached the Capitol and disrupted the counting. When these companies banned Trump — citing his capacity to inspire more violence — it was not a one-off event, but part of a continuum tracing back years. The days of “we're just a platform” were over.

A consequential war over moderation

As the larger platforms set their terms, those for and against their rules broke into two sides: One for unfettered speech, another for thoughtful moderation.

Republicans, somewhat naturally, joined the anti-moderation side. A significant swath of the Republican party is actively trying to undermine a democratic election, it regularly denies the science on the pandemic, and its leader sometimes winks at hate. The Republicans' ideological allies are anti-moderation absolutists who would rather let these discussions breathe than take them down.

The pro-moderation side sees a danger in leaving rule-violating posts up. They've seen people killed due to online speech, and fear more on the way. They also generally agree that the platforms have a responsibility to protect democracy. This side has attracted liberal activists.

These sides have fought a long war over the bigger platforms' policies and enforcement — with the president of the United States and Congress getting involved — and now the battle is spilling over into the smaller platforms. This time, everything is happening faster, as users and activists familiar with the contours of this conflict are pushing the platforms to set and enforce the rules as they'd like.

“I think we have all become smarter about the downsides of communication at scale,” Siri Srinivas, an investor at Draper Associates, told *OneZero*. As an early user of Clubhouse, Srinivas has seen the trade-offs firsthand. Clubhouse — a live, audio-only conversation platform — has already endured a lifetime's worth of moderation controversy.

“We discovered a lot of the issues with the bigger platforms about five years after they became too large,” Srinivas said. “We're using the same vocabulary to talk about Clubhouse.”

Over the summer, while still in beta, Clubhouse experienced its first high-profile flare-up. After *New York Times* reporter Taylor Lorenz criticized ex-Away CEO Steph Korey for views she expressed about the press, some Clubhouse users attacked Lorenz viciously in a Clubhouse room asking whether journalists have too much power. The attack was bad enough that Clubhouse co-founder Paul Davison spoke with Lorenz about what happened, and what Clubhouse could do to improve.

Clubhouse, which didn't respond to a request for comment, has since added user controls like a block button and an option to report abusive behavior. But the reporting functionality — which includes options to report content under categories like “discrimination or hateful conduct” — doesn't seem to be doing much. Lorenz has kept a running list of disturbing content on Clubhouse. An explicitly anti-LGBT conversation in December, for instance, featured one speaker who said, “I'm bringing down you fucking faggots.”

Clubhouse's tentative moderation approach might reflect the ideological reticence that Facebook, Twitter, and YouTube displayed in their early days. It may also reflect a unique business problem. As smaller platforms take off and fill with content, the cost to moderate can be overwhelming. "That's why you see so much emphasis on automation," the former Twitter employee said. "Taken to its logical conclusion, you could have a full federal jobs program moderating content on a platform like Facebook or Twitter."

Clubhouse might also be siding with the anti-moderation absolutists, allowing borderline conversations to go on — as a strategy. Such an approach would appeal to people who think big tech moderation has gone too far, or those seeking a safe space to say taboo things online. "I check the profiles of people who harassed me and they're all still there hosting big rooms," Lorenz told *OneZero*. Another Clubhouse user called it "audio 4Chan."

"Moderation has always been a top priority for Clubhouse, as it should be for any social platform," a Clubhouse spokesperson told *OneZero*. "Clubhouse has introduced many new moderation features throughout the year, including enhancements to blocking, real-time reporting and investigation of rooms, shared block lists, and more, and will continue to build new tools to detect abuse and to empower moderators as Clubhouse grows."

Either way, Clubhouse's spotty enforcement will only tee up more battles, and more confusion among users. "Having clear principles enunciated in advance, and telling your users what you're doing and how you're thinking about these important issues, benefits us all," said Douek, the Harvard lecturer.

Spotify, which declined to comment for this article, clearly enunciated its principles after signing Joe Rogan to an exclusive podcasting deal earlier this year worth more than \$100 million. The

deal was so pivotal for Spotify's podcast effort that some said it underpaid. But signing Rogan also stuck Spotify with its most significant content moderation crisis to date.

When Spotify started paying Rogan, some of its employees asked the company's management to remove the popular podcaster's episodes on which he expresses controversial statements about the trans community. Spotify is a platform, open for anyone to upload to its service as long as they follow its loose guidelines, which ban content that is "offensive, abusive, defamatory, pornographic, threatening, obscene, or advocates or incites violence." Employees viewed the money going to Rogan as an endorsement, however, and demanded more discretion.

Spotify leadership, with an opportunity to set the tone, chose not to touch the Rogan podcasts. "The fact that we aren't changing our position doesn't mean we aren't listening," Spotify CEO Daniel Ek told employees. "It just means we made a different judgment call."

Rogan watched Ek's decision and saw an opportunity to stick it to Spotify's employees. He soon brought Alex Jones on his show, hosting a man who infamously said the Newtown school shooting was a hoax. "It's very important for Rogan to demonstrate to the people in Spotify trying to control his content — and clearly there is movement inside Spotify to do that — to push back against that," said Greenwald, a recent Rogan guest. "Put on the person who's probably most horrifying to you, just to show you that you have no power to control me."

Ek's decision to side with Rogan didn't end debates within Spotify about whether Rogan, or other controversial hosts, should have a home at the company. There's even an internal Slack channel called #ethics-club within Spotify, with more than 500 members. Public pressure, it's worth noting, has influenced Spotify's content decisions in the past, including when it removed R. Kelly from playlists in 2018 after sexual assault allegations. Spotify backtracked on the move after an outcry.

“I think they’ll be able to maintain their current line because there is just so much broadly offensive or wrong podcast and musical content out there,” Will Duffield, a policy analyst at the Cato Institute, told *OneZero*. “And it paid \$100 million [to Rogan], so that’s hard to walk back.”

For Substack, “we’re just a platform” is a hard line to toe since the company does provide health care stipends, cash advances, and legal help to some of its writers (I use Substack to publish my *Big Technology* newsletter and have worked with the lawyers). Through this support, Substack helps pick some winners and losers, in some way, on its service.

If (and perhaps when) a writer who Substack supports becomes subject to a public pressure campaign, the platform could be in for a complicated ordeal. But the company appears unconcerned. “While we take a hands-off approach with who may use the platform,” it said in a recent blog post. “We will continue to take an active approach in helping and promoting promising writers.”

Substack would do well to have some infrastructure built for the eventuality of a content moderation crisis. Currently, the company has no policy team, and its founders run its content moderation efforts, according to Best. It shows. In light of recent events at the Capitol, I emailed Best to ask how he felt about a top-performing Substack newsletter that questioned the election results. Rather than engage, Best fell back on a corporatism. “As a rule we don’t comment on moderation decisions,” he said.

At, least not yet. “They’re going to have to prepare now,” Greenwald told *OneZero* of Substack. “To resist the onslaught that absolutely will be coming in their direction.”