Social Media & the January 6th Attack on the U.S. Capitol

Summary of Investigative Findings

Table of Contents

I. Overview

In the months before the election, President Trump, his allies, other Republican officials, and media personalities across the political right relentlessly attacked the integrity of the electoral process. Large numbers of voters were primed to question the result despite a total lack of evidence of fraud. Years of declining trust in institutions—including the media, government officials, and political leaders—meant no refutation of the "big lie" could loosen its grip on their imaginations.

This is part of a larger trend. For years, observers have warned about the increasing intensity and frequency of Republican politicians' dangerous and incendiary rhetoric. Calls to "take our country back" and warnings of "second amendment remedies" are indicative of a Republican Party that is increasingly willing to use violence to pursue its political ends and increasingly tolerant of extremism within its ranks. Far too often and for far too long, political leaders have capitalized on this escalation instead of confronting it. President Trump's behavior was the culmination of this trend, not its origin.

In recent years, this dynamic has played out on the Internet and especially over social media—though its roots are older than that industry. As an increasing number of Americans receive news and information online, observers have questioned whether social media platforms have independently contributed to the inflammation of political discourse. Whether or not that is true, social media companies own and profit from the services they provide to users. Regardless of their legal liability, they have an ethical obligation to prevent those services from being used to commit crimes, orchestrate violence, or otherwise contribute to offline harm.

This is true whether or not the attention-seeking, algorithmically-driven business model at the core of the social media industry is driving polarization and radicalization. In fact, the Select Committee's investigation, supplemented by written expert testimony, suggests that shoddy content moderation and opaque, inconsistent policies were a larger contributor to January 6th than the—admittedly not insignificant—challenges posed by recommendation algorithms. As one scholar told the Committee, these algorithms are "just one factor in a broader set of social, economic, and technical issues and incentives baked into the platform[s]."[1]

As the debate over social media's political significance continues, the online environment has grown more complex. Mainstream platforms like Facebook, Twitter, YouTube, and Reddit have been joined by smaller alternatives serving as havens for extremism and violent speech. When Reddit quarantined r/The_Donald, an online community supportive of President Trump and infamous for violent threats and hate speech, users migrated to TheDonald.win as an alternative forum free from Reddit's community standards. Other services such as Gab, Parler, and web forums like 4chan and 8kun similarly attracted users who posted rhetoric as well as audiovisual

[1] See interview of Becca Lewis by the Select Committee.

content that was unacceptable elsewhere. They have become places where neo-Nazis and white supremacists digitally mingle with militia members and far-right activists, as well as with curious Internet users – especially young people.

In these darker albeit easily accessible corners of the internet, extremists mobilized in the weeks before January 6th. It was on TheDonald.win, for example, where users discussed constructing the gallows that stood ominously in front of the Capitol the day of the attack.[2] On TheDonald.win and other alt-tech and fringe platforms, users shared maps of the Capitol tunnel system, exchanged tips about what type of restraints (e.g., zip ties) would most effectively detain Members of Congress, and discussed bringing weapons to the Capitol to prevent the certification of the election. Some called for mass lynchings of Democrats and "RINOs" (Republicans In Name Only). These conversations were not relegated to alternative platforms, however.

The Select Committee has also collected evidence from a high-ranking employee serving on Twitter's Safety Policy Team on January 6th; she said that she was deeply concerned about the content that was being posted on Twitter, including real-time posts about the movement of the crowd and breach points of the Capitol on the day. This, combined with the findings of experts and our in-house social media analysts, demonstrate that this sort of mobilization and celebration of the siege was not just occurring on fringe platforms.[3]

Indeed, while extremists mobilized on alternative and fringe platforms, false claims of election fraud and violent, angry rhetoric spread like wildfire across larger mainstream platforms. Some of this was stoked directly by the President himself, who tweeted on December 19th that there would be a "big protest" in Washington on January 6th and that his supporters should "be there, will be wild." Evidence provided to the Select Committee shows that before this tweet, January 6th was not a major target for protesters or violent actors. In the wake of the tweet, it became the primary target, the "last stand" for Trump and his supporters. Within Twitter's Safety Policy Team, for example, employees immediately noticed an escalatory shift in the tenor of content on the platform. Meanwhile, Discord shut down a pro-Trump server within hours of the tweet because of coordinated planning that began as an immediate response.

On Facebook, users coordinated to spread false claims of election fraud over the platform. Internal Facebook research describes how the "Stop the Steal" movement was propagated by a small core of individuals coordinating to send thousands of invitations to Stop the Steal Facebook groups each day while strategically evading enforcement of platform policies. Many of these users did so using multiple accounts, a violation of the platform's terms of service. These groups were rife with incitement to violence, threats, hate speech, and misinformation about the election. But because Facebook had no explicit policy against election denial and its systems for

---

[2] "Gallows or guillotines? The chilling debate on TheDonald.win before the Capitol siege," *Washington Post*, April 15, 2021, <https://www.washingtonpost.com/technology/2021/04/15/thedonald-capitol-attack-advance-democracy/>.
[3] Anika Collier Navaroli Deposition Continuance, *Transcript Forthcoming*

detecting violent rhetoric were unreliable, it took down relatively few of these groups before January 6th. Despite the knowledge that these groups had ties to violent actors, employee recommendations that the company take the problem more seriously were ignored or outright rejected.

Much of the content shared on Twitter, Facebook, Reddit, and other sites came from YouTube, the popular video hosting platform. YouTube claims that a series of changes to its algorithm in 2019 reduced the rate at which the platform recommends radical content to users. However, the platform's tolerance of "borderline content" and its late application of a policy against election denial—which was also not applied retroactively—allowed it to serve as a repository for false claims of election fraud. Whether or not YouTube's algorithm promoted these videos, they were deployed across the rest of the internet.

Major platforms' lax enforcement against violent rhetoric, hate speech, and the big lie stemmed from longstanding fear of scrutiny from elected officials and government regulators. Many of these voices called for stronger platform action and greater corporate responsibility; but on the right side of the spectrum, critics made largely baseless accusations that platform integrity efforts were designed to somehow suppress or censor conservative political speech.

Because of these accusations, platforms used a soft touch. Twitter, fearing political blowback and reveling in its status as the President's favorite app, was hesitant to write and enforce policies against violent rhetoric that would disproportionately affect conservative users. Employees warned senior Twitter leadership that the President was making coded appeals to extremists, driving an uptick in violent incitement on the platform; Twitter leadership hesitated to act until *after* the attack on the Capitol.

For at least two years before the election, Facebook executives intervened to make exceptions to their own policies for prominent right-wing partisans on their platform, including President Trump. The company was caught flat-footed by the surge of anger in Facebook Groups dedicated to denying the election's outcome. Executives resisted pleas to take the problem more seriously. As with Twitter, they changed course only after it was too late.

None of these platforms, mainstream or otherwise, are free from extremist content. As experts noted in a written statement submitted to the Committee, "Extremist content can be found in all corners of the web: on message forums, social networking platforms, streaming services, live chats of video games, static websites, and encrypted communication applications."[4] These applications provide a diverse suite of technical means for cultivating communication and organization, which political actors and extremists used to in distinct ways to spread disinformation and violent rhetoric related to the 2020 presidential election. They also provide

[4] Expert Statement of Heather J. Williams and Alexandra Evans, RAND Corporation, "Extremist Use of Online Spaces," Submitted to the Select Committee on April 25, 2022.

bridges between one another. For example, the long-term boosting of TheDonald.win on a pro-Trump subreddit – which Reddit administrators allowed to continue for months while the subreddit was on probation for violent content – allowed the more extreme forum to attract a greater following when it eventually launched in 2020.

The Select Committee's investigation found that these platforms were leveraged in varying ways by violent extremists—with varying degrees of success—in the run-up to January 6th. Some of these platforms took steps to curtail the propagation of extremist content prior to the attack on the Capitol, but in most cases the most dramatic actions were taken after January 6th despite clear warning signs that stretched across platforms. Other platforms covered in this investigation, such as Gab, 8kun, and TheDonald.win, showed no serious appetite for content moderation, which allowed truly extreme forces to hijack the sites with little hope for curtailing them. As documents provided to the Select Committee shows, these extreme users got out ahead of even the administrators of far-right platforms. For example, Parler was sending internal emails to the FBI warning about the possibility of danger on the 6th. Moderation logs from TheDonald.win show a futile attempt to remove the most violent content, even while other moderators boosted it.

The sum of this is that alt-tech, fringe, and mainstream platforms were exploited in tandem by right-wing activists to bring American democracy to the brink of ruin. These platforms enabled the mobilization of extremists on smaller sites and whipped up conservative grievance on larger, more mainstream ones. Seeing the fury of their base, Republican leaders abdicated their responsibility to the Republic by casting doubt on the election's outcome and, in some cases, plotting to overturn it. The sheer scale of Republican post-election rage paralyzed decisionmakers at Twitter and Facebook, who feared political reprisals if they took strong action. The Select Committee has collected both documentary and testimonial evidence to confirm this, as detailed in the following report.

Recent events demonstrate that nothing about America's stormy political climate or the role of social media within it has fundamentally changed since January 6th. Following the lawful FBI search of President' Trump's residence at Mar-a-Lago, both mainstream platforms and the sites where extremists plotted to assault the Capitol were again aboil with violent speech. Just days later, an armed man threatened the FBI building in Cincinnati, Ohio; reporting soon confirmed he was present at the Capitol riot. Until the incentives for violent, extreme, and even apocalyptic rhetoric are diminished, the threat of political violence will persist. The Select Committee's investigation shows that social media platforms were a key driver of this exact sort of rhetoric prior to the attack on the Capitol. An evaluation of the platforms' shortcomings in responding to these threats is an essential part of examining the ongoing challenges posed by violent far-right extremism and its attempts to crush American democracy.

Social media's relevance to the January 6th attack and the sociopolitical conditions that enabled it extends far beyond bad code or algorithmic manipulation. The Internet has become an easily searchable library of dangerous disinformation and extremist propaganda conveyed as factual

news. This library sits within a broader media environment wrought with anti-democratic discourse and fearmongering as well as encouraging the widespread distrust of government and its institutions. As long as this library exists, and all information within it is treated as equally valuable, people will continue to use it to inform their views about the world.

The report that follows analyzes the Select Committee's findings by each platform, focusing on highlights of testimony, documents, expert statements, and original analysis by Committee staff.

II. Key Findings

This bulleted list of key findings should server as an overarching roadmap for the report that follows. We have organized the report in a largely platform-by-platform analysis, but these investigative findings, to the extent they fit into the structure of the report, are present in the analysis of all the relevant platforms.

- **Social media companies were prepared for threats to the voting process but failed to anticipate post-election violence.** Facebook and Twitter both relaxed key policies following the election which might have helped mitigate the spread of violent incitement in advance of January 6th—suggesting that a defensive posture during election season is inadequate for civic integrity.
- **Social media platforms delayed response to the rise of far-right extremism—and President Trump's incitement of his supporters—helped to facilitate the attack on January 6th.** The Select Committee found that major social media companies all failed to adequately guard against the possibility that their platforms would be utilized by the rising far-right in the lead-up to 2020, as seen by: (1) Twitter's refusal to implement a policy against coded incitement to violence despite multiple warnings from its employees throughout the final months of 2020; (2) Facebook's refusal to adequately police the spread of disinformation or violent content on Stop the Steal groups despite their known nexus to militia groups; (3) Reddit's yearlong quarantine of r/The_Donald, which allowed moderators to freely and consistently promote TheDonald.win as an alternate platform; and (4) YouTube's failure to take significant proactive steps against content related to election disinformation or Stop the Steal.
- **Far-right platforms were the site of violence and grew their user bases as a response to moderation actions by major companies.** For platforms more closely associated with the far-right, ranging from Parler to Gab to TheDonald.win, bare-bones moderation policies made it impossible to even remove the most violent posts in the run-up to January 6th, and it was not clear that site leadership wanted that to happen.
- **Social media platforms did not have significant outreach from law enforcement prior to January 6th.** Speaking with multiple platforms, the Select Committee did not find significant evidence that law enforcement was warning these companies about the possibility for conflict on January 6th. Many companies did not recall proactive outreach from law enforcement whatsoever, as compared to a "war-room"-style crisis communication system that was present on Election Day.
- **Recommendation algorithms are only part of the problem**. Inadequate policy creation and implementation play an outsized role in the continued presence and proliferation of harmful and even violent content across social media in the weeks before January 6th.
- **Facebook did not fail to grapple with election delegitimization after the election so much as it did not even try.** Stop the Steal proliferated through Facebook groups and the

company declined to study false claims of election theft even when advised to by senior staff. Though the company removed the initial Stop the Steal group, a coordinated group of users worked to evade Facebook's takedown and grow the movement on the platform. After the election, nearly all of the fastest-growing groups on Facebook were related to Stop the Steal, and the company took action against only a small fraction of these groups. Despite warnings, Facebook leadership declined to take the problem seriously because it would "only create momentum and expectation for action."

- **Fear of reprisal and accusations of censorship from the political right compromised policy, process, and decision-making.** This was especially true at Facebook. In one instance, senior leadership intervened personally to prevent right-wing publishers from having their content demoted after receiving too many strikes from independent fact-checkers. After the election, they debated whether they should change their fact-checking policy on former world leaders to accommodate President Trump.
- **Donald Trump's willingness to encourage violence online was a challenge to social media companies, large and small.** Both Facebook and Twitter faced significant headwinds in taking aggressive action against problematic content by President Trump and his supporters, partly out of fear that they would be classified as overly partisan. Twitter in particular "relished" its place as the President's platform of choice and put in place policies that elevated him above all other Twitter users. The result was an unwillingness to take aggressive steps to clean up Twitter of inciteful content prior to January 6th and a disorganized, panicked attempt to permanently ban the President after the attack had concluded. For smaller platforms, such as Discord and Twitch, the President's behavior on January 6th prompted changes in how they look at the potential for removing speech by influential accounts or world leaders that could incite violence.
- **Key decisions at Twitter were bungled by incompetence and poor judgement.** Twitter's Safety Policy team, responsible for policies related to violence and incitement, warned for months that the risk of election-related violence was high and rising. They were ignored and retaliated against.
- **Twitter was also "paralyzed" by fear of political reprisals.** President Trump's account was protected from policy enforcement in unique ways—not even other world leaders received the same insulation from accountability. Likewise, Twitter avoided important policy calls around violent incitement because they would disproportionately affect the President's supporters.
- **Twitter failed to take actions that could have prevented the spread of incitement to violence after the election.** An insider account of a former Twitter employee showed how leadership rejected policies in the immediate aftermath of the election that would have allowed the company to remove implicit calls to violence, which they saw in response to President Trump's tweets. This policy was proposed as a direct response to President Trump's exhortation for the Proud Boys to "stand back and stand by," and grew out of a fear that he would use the platform to call his supporters into violent conflict.

- **YouTube's policies relevant to election integrity were inadequate to the moment.** The company did not ban election fraud claims until December 9th, and even then did not enforce that policy retroactively. Forms of election delegitimization not related to ballot fraud were also not considered in violation of this policy, meaning a great deal of election denial remained on the platform through January 6th and into the current day.
- **Facebook's emergency "break the glass" measures were regarded as successful in their goal of protecting election integrity.** To its credit, Facebook took steps to slow down the spread of content and pace of connectivity on its platform, to remove larger degrees of hate speech and violent incitement, and to boost higher-quality civic content. These were research-backed, viewpoint neutral proposals which some company leaders hoped could become permanent.
- **Social media companies largely did not receive clear warnings of violence from law enforcement before January 6th.** Many of them described their interactions with law enforcement as mostly focused on preventing foreign interference or efforts to disrupt the voting process. Law enforcement did not seem focused on the possibility that post-election violence would be planned on or inspired by social media.
- **Extremist users on various platforms took their cues directly from President Trump, particularly after his December 19th tweet.** The Select Committee has gathered evidence to show that far-right forums such as TheDonald.win, Parler, and 8kun were telegraphing their plans for violence following President Trump's December 19th tweet that pushed January 6th as a critical day in the Stop the Steal movement. However, this phenomenon was not confined to extremist platforms; on places like Twitter, Facebook, and Discord, supporters of President Trump closely tracked his claims about a stolen election and subsequently his calls to descend on D.C. to protest the Joint Session of Congress on January 6th, 2021.
- **Trump's December 19th tweet was a transformative moment across social media.** In addition to causing a shift towards more explicit targeting of D.C. on a singular date on Twitter, the President's "be there, will be wild" tweet caused major shifts in other platforms. The Select Committee has collected evidence to show that Discord shut down a server, DonaldsArmy.US, just hours after the tweet, because it immediately became a hub for users of TheDonald.win to organize and coordinate their travel to D.C. and strategize how they could bring firearms into the city in response to the President's call.
- **Smaller platforms did not react quickly enough to the threat posed by an organized far-right extremist movement centered around President Trump.** Aside from Facebook and Twitter, other social media companies did not move quickly enough to interrupt networks of extremists who threatened American democracy. The most damaging example is likely Reddit's belated takedown of r/The_Donald; by waiting a year to remove the subreddit from its site, Reddit allowed moderators to freely advertise TheDonald.win, an alternate forum that hosted violent content prior to January 6th.

- **Less developed social media platforms were forced to make entirely new content moderation policies to respond to President Trump's incitement.** In several cases, platforms such as Twitch and Discord altered their policies on incitement or shifted resources to focus on domestic extremist violence to position themselves to better respond to another situation like January 6th—or another leader like Donald Trump. In the case of Twitch, their new policy on incitement was a direct response to President Trump.
- **Fringe platforms use the cover of "free speech" to excuse a dangerous lack of content moderation.** Far-right platforms, including Gab, Parler, 8kun, 4chan, and TheDonald.win, in most cases lack even a remotely adequate content moderation system. In the case of Gab, there was only *one* employee responsible for surveying posts by millions of users on the day of January 6th. These platforms use the language of the First Amendment to justify these minimalist content moderation decisions; however, the Select Committee's investigation has shown that this dynamic prevents even content that the owners of these sites themselves admit should be deleted from being detected ahead of real-world harm, such as on January 6th. And even in these spaces, executives were concerned about the potential for violence; Parler's employees emailed the FBI that they were "worried" about the Joint Session just a few days before the attack.
- **President Trump could soon return to social media—but the risk of violence has not abated.** On Facebook, the President's suspension is due to be reviewed after two years. As Elon Musk attempts to acquire Twitter he has publicly state he intends to undo the suspension of Trump's account. YouTube has set no clear date or benchmark for the reversal of Trump's suspension and continues to monitor for reduced risk of violence.

III. Twitter

Beginning in 2019, Twitter began to publicly reckon with how to moderate statements from world leaders on its platform. These considerations would inform its approach through the COVID-19 pandemic, protests following the murder of George Floyd, the 2020 elections, and ultimately the January 6, 2021, attack on the Capitol.[5] President Trump's prominent use of the platform was a particular challenge, demanding a special level of attention from company leadership during his Presidency and, especially, during the 2020 election and its aftermath. Twitter's efforts to safeguard the election from activity on its platform largely relied on a combination of labeling tweets containing misinformation and downranking or de-listing certain trending topics or hashtags.

Despite these efforts, offline political tensions in the United States manifested on the service as terms like "civil war" trended and were subsequently de-listed. At the same time, President Trump used the service to communicate directly to extremists who came to Washington on January 6th with the intention to commit violence on his behalf. Twitter failed to prevent their call-and-response from reverberating across its service.

As the election approached, Twitter's Safety Policy team—which is responsible for writing policy guidance for content moderators dealing with violent speech and other issues—struggled for months to get guidance from management and to convince company leadership to take the risk of political violence more seriously. On the eve of January 6th, they warned that violent rhetoric on the platform could lead offline acts of violence. Their concerns were dismissed. Ultimately, Twitter's efforts to safeguard electoral discourse were undercut by poor management of key teams and poor judgement by executives who were concerned about the consequences of angering President Trump's supporters. The swiftness with which the company acted against President Trump's account following January 6th belies its prior recklessness.

In the fall of 2020, Twitter leadership privately feared that President Trump would use their service to incite violence. After the President instructed the Proud Boys to "stand back and stand by" on national television during the first presidential debate, Twitter leadership instructed the Safety Policy team to develop a policy concerning implicit calls for violence. Drawing on a "Black Lives Matter playbook" developed during the protests in summer 2020, the Safety Policy team created a policy against "coded incitement to violence," or CIV.

The Committee deposed a former Twitter employee who worked on the Safety Policy team who gave an inside account of the company's actions leading up to January 6th, some of which contradicted what Twitter represented to the Committee. **The former employee's testimony**

---

[5] "Twitter Policies and Enforcement: A Timeline of Events Surrounding the 2020 Election and the January 6th, 2021, Attack on the Capitol," Prepared for the Select Committee to Investigate the January 6th Attack on the United States Capitol, March 3, 2022.

confirms that Twitter saw President Trump's potential violent incitement of his supporters as a cause for concern even prior to Election Day but chose not to take effective actions to prevent him from using the platform in this way. Moreover, this failure to act was consistent with Twitter's longstanding deferential treatment of President Trump. The former employee told the Committee that the CIV policy was a direct response to the President's "stand back and stand by" comments, since the employee and others at Twitter were worried that the President would use Twitter to incite his supporters to violence.[6]

This policy was finalized on November 4; two days later, Twitter claims was integrated into "post-election protest guidance" meant to instruct content moderation teams on how moderate potentially harmful content in the weeks following the election.[7] The Safety Policy team felt that this document was significantly less thorough and less clear than the policy they initially proposed, and they followed up with their leadership frequently for guidance and clarification.[8]

However, according to the former employee, the inclusion of the CIV policy in the post-election guidance did *not* mean that the policy was able to be enforced, since had not yet been made public or finalized.[9] Instead of folding in the CIV policy into its post-election guidance, Twitter leadership team chose to revert to a less-nuanced restatement of its existing policies on content that wishes harm on others.[10] This made it impossible for the Safety Policy Team to remove posts that were increasingly suggestive of civil war in the weeks after the election.[11] The former employee said that their team never received guidance on the implementation of the policy prior to January 6th, even as they say "individuals becoming increasingly armed and the rhetoric becoming increasingly more violent" during the post-election period.[12]

Twitter's leadership and the Safety Policy team never aligned on how to handle the risk that post-election violence would be incited on the service, and the Safety Policy team complained that leadership was "confused" about the policy's "origin, urgency, and ultimate purpose."[13] This is especially significant because the policy was "escalation only"—tweets flagged as in violation of this policy were referred to the judgement of Trust & Safety leadership, overseen by Twitter Vice President for Trust & Safety, Del Harvey.

Del Harvey personally obstructed the CIV policy. To assist in its development, the Safety Policy team asked Twitter engineers to create a bot which would pull examples of tweets with language to which the policy might apply. It was rare for the team to receive this kind of engineering support, but in this case the effort was made, and the bot collected hundreds of

---

[6] J. Smith Deposition, 15:7-15.
[7] TWITTER00019259; TWITTER00019229.
[8] J. Smith Deposition, 21:1-22:22.
[9] *Id*. at 21:8-22:21
[10] Anika Collier Navaroli Deposition Continuance, *Transcript Forthcoming*
[11] *Id*. at XX.
[12] *Id*. at 22:23-24:13.
[13] Documents on file with Select Committee.

example tweets. Many of the tweets included phrases like "locked and loaded" (associated with discourse around the Kyle Rittenhouse shooting) and "stand back and stand by" (President Trump's directive to the Proud Boys during the first Presidential Debate).[14]

On November 9th, the team met with Del Harvey to discuss their findings. Harvey pushed back on the application of this policy, asserting that the hashtag "locked and loaded" could be a reference to self-defense and should not be the target of content moderation. The team pointed out that only a minority of cases where this hashtag appeared in their analysis could be interpreted as clear references to self-defense, and it was only one of several relevant hashtags, but Harvey was unmoved. The policy was spiked.[15]

After this point, the Safety Policy Team realized that it would be unable to use its already drafted coded incitement policy to lower the temperature on the platform, which was continuing to see content that was suggestive of civil unrest, which the former employee characterized as a gradually amplifying trend that began to accelerate with the President's "stand back and stand by" comment at the debate and continued to gather strength after he failed to concede the election, at which point calls for overthrowing the government became clearer.[16]

Through November and December, the team warned management about the level of coded incitement they were seeing; they wrote that "viral content… needs to be addressed immediately" and predicted that online speech would be "the precipitating factor" to post-election violence. Their manager, however, expressed discomfort about removing content using the language "locked & loaded" as coded incitement. At one point, they told a member of their team that there was no CIV policy, "full stop." The team raised this issue and their manager's general lack of guidance directly with Del Harvey, to no avail.[17]

Poor management of the Safety Policy team was a longstanding problem. Their previous supervisor had left the company in summer 2020—a departure two individuals familiar with the situation said was likely involuntary and due to the "psychologically unsafe work environment" they created. But where the previous manager was hypercritical and thin-skinned, their replacement was disengaged and seemingly uninterested in the policies their team managed. Frequently, they did not even show up to team meetings, and at one point during January 6th itself shrugged their shoulders when asked about whether a post had the potential to incite violent.[18]

In the last months of 2020, the situation grew so bad that at least two team members went to Twitter's Employee Relations department with their concerns. They said that their manager was

---

[14] J. Smith Deposition 20:3-21:7; Interview with J. Johnson.
[15] *Id.*
[16] Anika Collier Navaroli Deposition Continuance at XX.
[17] Deposition with J. Smith at 39:8-10.
[18] Deposition with J. Smith, 125:5; Interview with J. Johnson.

disengaged and that they needed to be retrained on key policies in the Safety Policy team's portfolio. These individuals told Select Committee staff that their concerns were treated dismissively and that, while a later review did recommend the manager be retrained on policy related to how to deal with violent content, to their knowledge no additional training was ever supplied—even after the catastrophic violence of January 6th.[19]

This management turnover and neglect damaged the company's election preparedness. A former member of that team told Select Committee staff that the company had only one Safety Policy project related to the election as late as August 2020; it addressed the potential use of slurs against candidates.[20] When Twitter created an election threat model to explore its preparedness for specific threats, key Safety Policy issues received poor marks because of insufficient processes to guide staff responses. These included policies related to incitement of violence.[21] Team members believed they could be better prepared by setting concrete goals—for example, shorter turnaround time on tough decisions regarding borderline content—but management consistently failed to do so despite their requests.[22]

This was not the only way in which the team felt hamstrung. In September, they complained to management that they did not have access to vital back-end notes on account violations which they needed to do their jobs.[23]

Later, in a meeting on October 21st, counterparts to the Safety Policy team in Asia—who handle the US night shift—also expressed uncertainty about how to operate in the tense weeks before the US election. Afterward, the manager called a member of their team to complain about the colleague who asked the question; the team member felt the phone call was inappropriate and pointed out that their colleague's question was important because neither of their teams had received sufficient guidance.[24]

As Twitter struggled to prepare for potential violence following the election, the President was upping the ante. In the weeks after the election, Twitter staff were concerned by escalating violent rhetoric on the site, much of which was coupled with narratives about election fraud. While Twitter acted against users who spammed hashtags like "#stopthesteal," when a report from the advocacy group Advance Democracy alleged that election fraud narratives dominated the top five tweets, top ten hashtags, and the top links shared on Twitter, a senior manager called the finding "unsurprising."[25]

---

[19] Deposition with J. Smith at 108:25-109:9; Interview with J. Johnson.
[20] Interview with J. Johnson.
[21] TWITTER00020545.
[22] Summary Memo of Interview with J. Johnson.
[23] Files provided to Select Committee, p499.
[24] Deposition with J. Smith; Interview with J. Johnson.
[25] TWITTER00022749; TWITTER000227772.

Then, on December 19th, President Trump tweeted:[26]

This tweet crystallized the threat of violence, and the former employee told the Committee that this was the first time that content on Twitter coalesced around a single event at the Capitol on January 6th, which the employee described as "essentially staking a flag in D.C…for his supporters to come and rally."[27] In a deposition with Select Committee staff, a former member of the Safety Policy team said that the attack would not have occurred with the same magnitude without this tweet, and they testified that there was a distinct shift as users began focusing specifically on the Capitol and the day of January 6th, with users treating the event as a "place to be" and posting specific calls to arms such as "I am locked and loaded and ready to be in D.C. for a civil war on January 6th."[28] The President had issued a call; the response was a violent attempt to prevent the Presidential transition.

The former employee explained to the Committee that Twitter's reluctance to put in place its CIV policy prevented it from responding appropriately to the December 19th tweet, which was not itself an explicit call to violence but caused an avalanche of violent and dangerous responses across the platform. They said that the presence of a CIV would have allowed Twitter to be "more proactive in responding to the responses of this tweet" and "would have allowed Twitter to rein in the extensive calls for violence" that were focusing on the Joint Session following the President's tweet.[29]

On January 5th, Safety Policy staff again raised the issue of coded incitement to violence and asked for guidance on enforcing a policy against it.[30] Unfortunately, their manager was out of the office on this day owing to a death in the family. The supervisor on duty was based in Ireland. As staff warned of the possibility of violence and asked for guidance in this area, that manager

---

[26] Retrieved from https://news.sky.com/story/capitol-riots-how-were-protesters-in-dc-organised-12181496.
[27] J. Smith Deposition, 79:1-23
[28] *Id.* at 80:7-19.
[29] *Id.* at 80:2-16.
[30] A video recording of this meeting is on file with the Select Committee.

was only able to tell them that they could escalate policy violations if violence occurred. She further advised them that she would be unavailable during night hours on Irish time. Because the next supervisor in line was a recent hire, there was no experienced supervisor was on duty the morning of the certification of the Presidential election.[31] The former employee was present at this meeting and told the Committee that she "realized no intervention was coming, and even as hard as I had tried to create one or implement one, there was nothing, and we were at the whims and the mercy of a violent crowd that was locked and loaded."[32]

The Safety Policy team was not the only source of warnings to Twitter's leadership. The weekend before the attack, a representative from the Georgia-based civil rights advocacy nonprofit Fair Fight reported several violent tweets to Twitter targeting the Georgia special election, which ultimately decided control of the US Senate. Among these were tweets from Overstock.com CEO Patrick Byrne, who threatened to "lynch" an election official and claimed to have paid an operative to break into a voting facility to retrieve "samples." These tweets led to individual threats to the physical safety of a specific, named individual on pro-Trump message boards.

Another was from the prominent white nationalist Nick Fuentes, who said during a livestream that Georgians had "no other recourse" than to kill state legislators. Yet another came from Project Veritas, which named a specific advocacy center, resulting in several of the center's employees receiving death threats and being doxed by far-right activists.

Amazingly, Twitter's initial response to Fair Fight said that many of these tweets did not violate its policy against violent threats or were only eligible to be labeled, not removed. Those that were eventually removed remained on the platform until after the Georgia election, on the morning of January 6th.

On January 6th, Twitter management struggled to respond within its policy framework. When the assailants breached the Capitol, the Safety Policy team's regular manager logged on and joined a meeting with two members of the team. They gave two directives: find a rationale to suspend the President's account from the service, and "stop the insurrection." When asked how to fulfill the second objective, the manager shrugged. The team was left to respond to rampant incitement on Twitter under its own initiative, once again without clear instruction.[33]

Twitter has provided the Select Committee with a curated version of events in the days around January 6th, which includes an assertion that there was implementation of the CIV policy after the breach of the Capitol; this was contradicted by the former employee in her sworn

---

[31] Deposition with J. Smith; Interview with J. Johnson.
[32] J. Smith Deposition at 118:8-14.
[33] *Id*. at 123:1-18.

testimony.[34] They said that there still was no finalized coded incitement policy that could be operationalized on that day, and that they were forced to draft a last-minute enforcement guidance with other members of the Safety Policy Team that spelled out how they would ad hoc apply the CIV policy.[35] There was no prior preparation to make that policy clear before the start of the attack itself. The former employee, who had helped to develop this policy in response to concerns about the President using the platform to whip up violence, confirmed that this policy was implemented haphazardly in a matter of hours, while the attack was ongoing, in order to help Twitter get a handle on what was going on. They said that this was the same policy leadership had been resisting for two months and that "it took violence occurring on the ground" in order for the approval to be finalized.[36]

Twitter saw highly detailed posts about the attack on the day of January 6th itself. The former employee, who was one of the few employees monitoring content that day, said that Twitter was faced with a barrage of posts that essentially "live-streamed" the attack, with details that were specific enough to focus on individual breach points and different areas of the Capitol that the crowd had reached.[37] They recalled that, because there was not a CIV policy or special response team in place for the Joint Session—as would have been routine practice for other events with a risk of socio-political unrest—members of the Safety Policy Team were manually taking down violent tweets, including those including "#ExecuteMikePence," using only the Twitter search function.[38] This understaffed, ramshackle made the former employee feel like she was a security guard hovering over the Capitol, trying to defend the building as the crowd tweeted out its progress during the course of the assault.[39]

The former employee also explained that the content on Twitter that day was highly correlated to events on the ground. They noted that President Trump's call to go to the Capitol resulted in an immediate shift in the kinds of posts on the platform, and that users responded to his condemnation of Vice President Pence.[40] They also recalled a tweet by Ali Alexander, which appeared to be supportive of the storming of the Capitol, caused the former employee to ask her boss for clarity on how to apply the CIV policy, and decided on their own that the tweet was grounds for removal.[41]

---

[34] "Twitter Policies and Enforcement: A Timeline of Events Surrounding the 2020 Election and the January 6th, 2021, Attack on the Capitol."
[35] J. Smith Deposition, 81:20-82:3.
[36] *Id.* at 128:4-5.
[37] Anika Collier Navaroli Deposition Continuance, *Transcript Forthcoming*.
[38] *Id.* at XX.
[39] *Id.* at XX.
[40] *Id.* at XX.
[41] *Id.* at XX. In a video posted to Twitter during the January 6th attack, Ali Alexander said "I don't disavow this," while pointing at the crowd surging towards the Capitol. Will Sommer, *Stop the Steal Organizer in Hiding After Blame for Riot*, Daily Beast (Jan. 10, 2021), available at https://www.thedailybeast.com/stop-the-steal-organizer-in-hiding-after-denying-blame-for-riot.

During this period, Twitter senior managers and executive leadership received hundreds if not thousands of pleas from the public to take strong action against President Trump and false claims of election fraud. The Select Committee has reviewed many of these messages, but at least one—from a technology adviser to a US Senator—led to a longer exchange. Received by the head of US Public Policy during the January 6th attack on the Capitol, it alleged that:

> "We're rapidly approaching a stage at which the President's use of your service to incite violence and insurrection is *concretely* and *very directly* producing violence and civil disorder. I would strongly encourage your company to make clear where the red line is—and be ready to enforce it. We've reached a really unprecedented point. Steps that large platforms take in the next 24 hours can have significant effects."[42]

Twitter's head of US Policy responded that they were watching the situation "very closely." The original sender responded that:

> "I am telling you emphatically that you need to put out a statement about where your redline is and be prepared to draw it. Platforms are going to bear a *lot* of responsibility for helping facilitate this. I really hope you do more than watch today. There are reports of shots being fired by these militia groups. You are continuing to allow right-wing voices to stoke this violence. It is not OK."

In the exchange that followed, the original sender also wrote that "It is amazing to me that people like Ron Watkins still have Twitter accounts."

When Twitter's US Policy lead responded, "Who is Ron Watkins," they replied:

> "For real? He and his dad run 8chan/8kun. They are widely believed to have taken over the QAnon conspiracy a few years ago… you should also check out [name redacted by Twitter's counsel] – she's the QAnon- brain-poisoned woman who was shot today for storming the Capitol. Active Twitter user, where she consumed an *enormous* amount of QAnon content but also (particularly recently) Lin Wood content calling for Rod Rosenstein and Democrats to be extra-judicially arrested and tried for treason."

Meanwhile, Twitter staff fumed that their warnings of possible violence had gone unheeded by leadership. They wrote an open letter to their colleagues:

---

[42] TWITTER00022925. Italics in original.

- Locked and loaded
- Ready for a civil war
- Need a civil war
- Stop the Steal
- 1776
- Law & Order
- Stop the Steal
- Revolution2021
- 1776
- Freedom or death
- Pledge your lives
- Take back our country
- Coup
- Storm the Capitol
- "Freedom or death"
- Revolution
- civil war
- shoot them
- real patriots
- "antifa hunting"
- this is war
- #1776Again
- Race war now
- Let the revolution begin
- Storm the gates
- We want a total war
- Here comes the big one
- The war is beginning

Staff,

We are disturbed by the day's events in Washington, D.C. and request three actions of Staff:

1. Permanently suspend @realDonaldTrump over his actions on January 6
2. Provide a clear account of the day's decision making process regarding the President's tweets
3. An investigation into the last several years of corporate actions that led to Twitter's role in today's insurrection

For the last four years, we have watched right wing extremists grow on our platform, nurtured by @realDonaldTrump. We have seen Twitter leadership struggle to deal with the violent, hateful rhetoric shared by @realDonaldTrump. We have respected those decisions and had faith in those who make them, knowing that settling policy around hate speech and misinformation is nuanced and difficult. We appreciate stronger measures, like the interstitials recently used on his account and his Jan. 6 timeout.

We do not believe these actions are sufficient.

Trump is no longer a legitimate democratic actor. In a video posted on Twitter, he called the election "fraudulent" and a "landslide" in his favor, while referring to his loyalists as "very special" even as they broke into the Capitol building and clashed with police. At no point did the outgoing President unambiguously denounce their actions.

We must examine Twitter's complicity in what President-Elect Biden has rightly termed insurrection. Those acts jeopardize the wellbeing of the United States, our company, and our employees.

First, we request that @realDonaldTrump's account be suspended permanently, before he can further harm using our platform.

Second, we request more transparent internal communications around Twitter's civic integrity policy and how it applied.

Third, we request an independent investigation into Twitter's role in these events. Despite our efforts to serve the public conversation, as Trump's megaphone, we helped fuel the deadly events of January 6th. We request an investigation into how our public policy decisions led to the amplification of serious anti-democratic threats. We must learn from our mistakes in order to avoid causing future harm.

We play an unprecedented role in civil society and the world's eyes are upon us. Our decisions this week will cement our place in history, for better or worse.

We do not claim to speak for any other Tweeps than those who add their names here.
The Undersigned

Text of letter from Twitter staff on the events of January 6th, 2021.[43]

In the end, only the first of their demands was met. By the end of the week, President Trump would be permanently suspended from the platform. But Twitter ultimately took no transparent steps to account for its previous treatment of his account or Twitter's own contribution to the violence that day.

In the meantime, other teams also responded to the attack on the Capitol. The Product Trust team, for instance, noticed that #CivilWar was trending on the platform. This hashtag was denylisted from trending topics, along with #stormtheCapital, Antifa, #Amerikkka, and other hashtags and phrases.[44] Twitter Services—which oversees frontline content moderation and escalates tweets as necessary to Safety Policy and other teams for policy assessments—updated

[43] TWITTER00000736.
[44] TWITTER00019287.

its enforcement guidance to recommend an "aggressive approach," including proactive filtering and manual review, for tweets with "the propensity to incite offline violence." Accounts which used one of a specific set of phrases and had recently received a "safety label," such as an abuse strike, would be either permanently suspended or filtered from visibility.[45]

This list of phrases included, "locked and loaded," which Del Harvey previously argued should not be treated as incitement because it could refer to self-defense in the home. The team assessed that "stand back and stand by" was being used in ways too nuanced for automatic enforcement and recommended that borderline cases using that phrase be escalated for further review.

Trump's suspension ended the preferential treatment Twitter gave his account for years. The evening of January 6th, a member of the Safety Policy team drafted a memo recommending Trump receive a 12-hour suspension for multiple violations of Twitter's policies. Specifically, three of his tweets were "bounced with a strike" for policy violations including glorification of violence.[46] However, the former employee told us that there was no monitoring by Twitter to determine what would happen when the President's temporary suspension ended, and that the platform was "continuing to fly by the seat of its pants."[47]

On January 8th, Trump—his account now unlocked—tweeted again:

> "The 75,000,000 great American Patriots who voted for me, AMERICA FIRST, and MAKE AMERICA GREAT AGAIN, will have a GIANT VOICE long into the future. They will not be disrespected or treated unfairly in any way, shape or form!!!"

And again:

> "To all of those who have asked, I will not be going to the Inauguration on January 20th."[48]

In response, Yoel Roth, the head of Site Integrity, asked another member of Twitter's Trust & Safety team to write a memo on the possibility of permanently suspending the President's account.[49] In an initial draft, this individual recommended against suspension because the tweet in question did was significantly less violative that previous tweets for which Trump escaped enforcement. Unbeknownst to the author, this draft was provided to Jack Dorsey and Twitter's Chief Legal Officer, Vijaya Gadde. While that first draft was under review, the team met again and produced a second draft incorporating broader context: Trump's tweets were actively

---

[45] TWITTER00020557.
[46] "Twitter Policies and Enforcement: A Timeline of Events Surrounding the 2020 Election and the January 6th, 2021, Attack on the Capitol."
[47] J. Smith Deposition, 140:20-25.
[48] "Twitter Policies and Enforcement: A Timeline of Events Surrounding the 2020 Election and the January 6th, 2021, Attack on the Capitol."
[49] Deposition with J. Smith.

inspiring more violent incitement on the platform, with users targeting January 17th as another day for armed protests at state capitols and in Washington D.C.[50] The former employee explained that they were seeing individuals who were upset that they had not participated in January 6th posting about their desire for another opportunity to join in another, similar event, and that it appeared that this would be an even worse version of the January 6th attack.[51] At the same time, the FBI was warning that state capitols could face significant threats on that day.[52] The team's second draft recommended permanent suspension.

In their deposition, the former employee—who was involved in this stage of the process—emphasized that the Safety Policy Team was seeing "the exact same rhetoric and the exact same language that had led up to January 6th popping underneath" President Trump's tweet about not attending the inauguration, and that they believed that the same sort of attack was likely to recur if President Trump continued to tweet in the same way.[53] The former employee also described Del Harvey's disbelief that another attack was likely, to which she responded, "do you want to have more blood on your hands?"[54]

Refusing to take swift action against the obvious potential for violence spurred on President Trump's Twitter account was, in the mind of the former employee, "absolutely indicative and emblematic of Twitter's hands-off, willfully ignorant approach to the former President's rhetoric on the service and on the platform" and the "wait-and-see approach" that the company's leadership adopted with regard to the former President.[55]

The author of the second memo met later with their manager and Del Harvey; Yoel Roth also attended. Harvey suggested that "minds can differ" on the question of Trump's suspension and suggested Twitter wait until Trump crossed the line again before acting. The memo author objected that this was the rationale that led to the January 6th attack and noted that Harvey made a seemingly opposite decision in a previous case. They left the meeting unsure of Harvey's ultimate opinion on the matter—but later that day, the second draft of the memo was reached the desk of Jack Dorsey, who asked Harvey to make unknown edits to it. Twitter announced the suspension later that day.[56]

The former employee pointed out that Twitter leadership's deferential treatment of President Trump was reflected in their response on January 6th itself. They pointed out that it took hours

---

[50] According to documents provided by Twitter, "#J17" was among the hashtags proactively swept for and enforced on by the company after January 6th. *See* "Twitter Policies and Enforcement: A Timeline of Events Surrounding the 2020 Election and the January 6th, 2021, Attack on the Capitol."
[51] Anika Collier Navaroli Deposition Continuance, *Transcript Forthcoming*.
[52] John Eligon et al., *FBI Urges Police Chiefs Across the U.S. to be on High Alert for Threats*, New York Times (Jan. 13, 2021), available at https://www.nytimes.com/2021/01/13/us/fbi-police-threats-inauguration.html.
[53] J. Smith Deposition, 147:22-148:3.
[54] *Id*. at 148:4-9.
[55] *Id*. at 148:8-24.
[56] "Twitter Policies and Enforcement: A Timeline of Events Surrounding the 2020 Election and the January 6th, 2021, Attack on the Capitol."

after the attack on the Capitol had begun for Twitter to lock President Trump's account. "In other circumstances or situations in which violence was occurring on the ground, if we knew that there would be a leader or some sort of promoter of that, we would've taken it down quickly," said the former employee.[57]

Curiously, not long after the suspension, Twitter's board took interest in a neglected side project on rehabilitation of banned users. Staff took this as a sign that they were exploring ways to revoke the suspension of Trump's account. They noted that the project was rooted in concepts of restorative justice, which requires individuals to recognize the harm they have caused and accept responsibility for it. They felt that this was unlikely in the case of the former President, who made election denial a litmus test for Republican politicians in the months after January 2021.[58] In interviews conducted by the Select Committee, Twitter staff familiar with the company's actions during the election and January 6th described an organization paralyzed by fear of political reprisals. They attributed Harvey's reluctance to enforce the CIV policy to the knowledge that it would overwhelmingly affect the President's supporters and invite accusations of anti-conservative censorship. This fear also explains why the company was so reluctant to act against the President's account until the summer of 2020, though the company may also have "relished" its status as the President's signature app. Trump's account received protections afforded to no other world leader: for instance, attempts to view his account in the company's back-end dashboards triggered a security alert system. Only Jack Dorsey or Vijaya Gadde could approve policy enforcement against the President's account.[59]

The aftermath of the January 6th attack shows a company rapidly backtracking on previous decisions while failing to reflect on institutional failure. On January 15, 2021, Twitter's Safety Policy team requested a seven-day proactive "Sweep" for phrases used in connection with post-election violence and the January 6th attack, such as "Stop the Steal," "America First," "Take back our country," "Civil War 2," and "6MWNE" ("Six Million Were Not Enough," a reference to the Holocaust).[60] The company also banned 70,000 accounts associated with QAnon.[61] But the former employee said that this belated action the fact that this step was reflective of the fact that, "it often took violence or death to occur in order 6 for Twitter to prioritize making policy decisions."[62]

In the weeks after January 6th, Twitter hired outside counsel to conduct a retrospective assessment of the company's actions leading up to that day. Attorneys interviewed members of the Safety Policy team, who never saw the final product of this process. Meanwhile, at least one member of the Safety Policy team asked their manager on seven separate occasions between

---

[57] J. Smith Deposition, 136:6-9.
[58] *Id.* at 63:23-64:5.
[59] *Id.*
[60] TWITTER00019282.
[61] "Twitter Policies and Enforcement: A Timeline of Events Surrounding the 2020 Election and the January 6th, 2021, Attack on the Capitol."
[62] J. Smith Deposition, 140:5-6.

February and July 2021 about an internal retrospective of events leading to Donald Trump's suspension. This individual was not aware of any internal retrospective conducted by the end of 2021—by which time many members of the US team who were employed at Twitter on January 6th had left the company.[63] No retrospective document of this kind was provided to the Select Committee despite multiple requests to Twitter's counsel.

In a briefing for the Select Committee with Twitter's current Global Lead for Trust & Safety Policy, Juan Felipe Rincón, the extent of Twitter's efforts to bury this chapter of its history came into stark relief.[64] Rincón was not employed by Twitter during the January 6th attack, having assumed his role in July 2021. Committee staff asked Rincón if he had reviewed any company retrospectives of January 6th in his capacity as Global Lead for Trust & Safety Policy. He replied that he had not because the company's response to January 6th was "controversial," and he would rather focus on improving the policymaking process rather than questioning decisions made during a specific event. Further, Rincón said that he "intentionally tried not to ask questions" about this topic.

When asked if he was aware of any retrospective accounts of Twitter's handling of the events of January 6th, he said that he was not and that he tried not to be, because such a retrospective would be subject to legal scrutiny and available only on a "need-to-know" basis. When asked who, if not the head of Trust & Safety, was need-to-know, he answered, "those involved."[65]

Select Committee staff asked multiple times how, as the Global Lead for Trust & Safety Policy, Mr. Rincón could have failed to take a full inventory on the state of his team and Twitter policy by reviewing the company's response to January 6th; he replied that he and the Select Committee staff must merely have different styles of management. [66]

Individuals familiar with Twitter's Safety Policy work told Select Committee staff they found these answers unusual and troubling.[67] The Select Committee was also told that after ninety days on the job, Mr. Rincón produced an assessment of the company's safety policy comparing it to that of its peers. This assessment found Twitter's policies on violent incitement lacking. Despite this, no substantive work was done to improve the situation in 2021. The CIV policy was not used in other high-profile elections that year, such as the 2021 Brazilian elections. No one at the company seemed interested in pushing for the CIV policy, having been "traumatized" by the events of January 6th.[68]

---

[63] Letter to Candyce Phoenix, May 19, 2022.
[64] Memo on March 23, 2022, Select Committee Briefing with Twitter.
[65] *Id.*
[66] *Id.*
[67] J. Smith Deposition, 112:19-22; Summary Memo of Interview with J. Johnson.
[68] Summary Memo of Interview with J. Johnson.

Since January 6th, Twitter's challenges have not eased, nor has the risk of violence abated. Safety Policy professionals who spoke with the Select Committee believe that the company's lack of transparency fuels the accusations of bias it fears. They said that if the company more proactively explained its decision-making processes and the data they are based on, it might help the public appreciate the nuance of those decisions and help hold the company more accountable for policy inconsistencies like the treatment of Trump's account. As one professional put it, "the prevention of harm is an apolitical goal."

Unfortunately, they feel that the company has not accepted accountability for its role in the violence on January 6th. They do not believe the dynamics within the company have changed. And they warned continued rise of right-wing extremism in the United States means that there is a very real risk of another day like January 6th—and that next time, it will be worse.[69]

Outside experts have attested that verified, high-follower accounts were not held proportionately responsible for spreading election-related disinformation. They told the Select Committee that "social media sites developed policies aimed at accounts that repeatedly spread false and misleading content about the election, but preliminary research suggests that the rules were far more likely to be enforced for 'unverified' accounts while verified, highly-followed accounts continued to spread false and misleading content."[70] This finding mirrors Twitter's treatment of Trump's account, the ultimate high-follower breaker of Twitter's rules, and suggests that reluctance to act against high-visibility users extended beyond the former President.

Twitter's response to violent rhetoric is the most relevant affect it had on January 6th, but the company's larger civic integrity efforts relied heavily on labeling and downranking. In June of 2019, Twitter announced that it would label tweets from world leaders that violate its policies "but are in the public interest" with an "interstitial," or a click-through warning users must bypass before viewing the content.[71] In October of 2020, the company introduced an emergency form of this interstitial for high-profile tweets in violation of its civic integrity policy.[72] According to information provided by Twitter, the company applied this interstitial to 456 tweets between October 27th and November 7th, when the election was called for then-President-Elect Joe Biden. After the election was called, Twitter stopped applying this interstitial.[73] From the information provided by Twitter, it appears these interstitials had a measurable effect on exposure to harmful content—but that effect ceased in the crucial weeks before January 6th.

The speed with which Twitter labels a tweet obviously impacts how many users see the unlabeled (mis)information and how many see the label. For PIIs applied to high-profile

---

[69] Deposition with J. Smith; Summary Memo of Interview with J. Johnson.
[70] See expert testimony from Renee DiResta and Kate Starbird; see also memo by Select Committee staff, "Platform-specific findings," Meghan Conroy & Alex Newhouse, August 31, 2022.
[71] Deposition with J. Smith; Interview with J. Johnson.
[72] "Twitter's Responses to Select Committee Staff Questions of March 16, 2022." Twitter (April 15, 2022).
[73] "Twitter Policies and Enforcement: A Timeline of Events Surrounding the 2020 Election and the January 6th, 2021, Attack on the Capitol."

violations of the civic integrity policy, about 45% of the 456 labeled tweets were treated within an hour of publication, and half the impressions on those tweets occurred after Twitter applied the interstitial. This number rose to more than eighty percent during election week, when staffing resources for civic issues were at their highest; after the election, staff were reassigned to broader enforcement work.[74] In answers to Select Committee questions during a briefing on the company's civic integrity policy, Twitter staff estimates that PIIs prevented more than 304 million impressions on violative content. But at an 80% success rate, this still leaves millions of impressions.

On May 26th, 2020, Twitter applied a fact-checking label to a tweet from the President containing "potentially misleading information" about mail-in ballots.[75] This was the first time Twitter had labeled the President's account. Days later, the President tweeted:

> .... These THUGS are dishonoring the memory of George Floyd, and I won't let that happen. Just spoke to Governor Tim Walz and told him that the Military is with him all the way. Any difficulty and we will assume control but, when the looting starts, the shooting starts. Thank you!

Twitter applied an interstitial to the tweet, which violated its glorification of violence policy. On June 23rd, Twitter applied another interstitial to a tweet from the President's account. It did so again on August 23rd.

Starting at least with the August 23rd tweet but possibly before, Twitter limited the ways in which users could interact with Trump's tweets which had received an interstitial: they could no longer like, reply to, or retweet it, but could retweet with a comment. Twitter's rationale was that this last feature was usually used to signal disagreement, though later analysis found that once the ability to retweet was removed, retweeting with a comment quickly became another way to signal support.[76]

These actions occurred against a backdrop of political tension and high stakes. Twitter was under intense pressure to act due to the ongoing COVID-19 pandemic and nationwide protests against police violence; but it faced other pressures, too. On May 28th, President Trump signed an executive order targeting social media company's protection against legal liability for content moderation decisions.[77] He did so under the pretense of unsubstantiated claims of anti-conservative bias, tweeting that "This will be a Big Day for Social Media and FAIRNESS!"

---

[74] *Id.*; "Twitter Policies and Enforcement: A Timeline of Events Surrounding the 2020 Election and the January 6th, 2021, Attack on the Capitol."
[75] "Twitter labels Trump's tweets with a fact check for the first timer," *Washington Post* (May 27, 2020), https://www.washingtonpost.com/technology/2020/05/26/trump-twitter-label-fact-check/.
[76] TWITTER00000374.
[77] "Trump signs an executive order taking direct aim at social media companies," Tech Crunch (May 28, 2020), https://techcrunch.com/2020/05/28/trump-social-media-executive-order/.

Twitter used other, less disruptive "annotative" labels in lieu of removing content. These labels did not need to be clicked through and linked to "reliable information on the relevant topic." As with interstitials, labeling speed matters: according to Twitter, the median time to apply these labels was 2 hours and 28 minutes. Between October 27th and November 11th, the majority of all impressions on labeled tweets occurred after the label was applied.[78] As with interstitials, this rate was higher—about seventy percent—on election day, before staff were reassigned to other duties. The click-through rate on these links was between three and five percent—high, perhaps, for online advertisements or survey responses, but still only a small fraction of users exposed to the labeled tweet.[79]

Unlike the emergency interstitials applied to pre-election civic integrity violations, Twitter continued to apply these annotative labels to civic integrity violations, including for election fraud claims after election day. Between November 8th and January 6th, 102 tweets from President Trump's account were labeled for violating the civic integrity policy.[80]

Example of Annotative Tweet Label.

---

[78] "Twitter's Response to Select Committee Staff Questions of March 16, 2022."
[79] *Id.*
[80] Twitter will "bounce" violative tweets for enforcement, including removal; depending on the policy violation, they may or may not receive a strike. Many policies during this period appear to have been "bounce with no strike" policies, which may (along with his political position and notoriety) have allowed Trump to escape a ban before January 2021. According to statements made to Select Committee staff, any enforcement action against Trump's account could only be approved by the CEO or Chief Legal Officer. See also, "Twitter Policies and Enforcement: A Timeline of Events Surrounding the 2020 Election and the January 6th, 2021, Attack on the Capitol."

Like interstitials, deamplification was an effective measure for reducing views of harmful tweets. Twitter estimates that by deamplifying labeled tweets and hashtags in users' feeds, it reduced impressions on this content by 40% during the week of the election and by about 80% by December 14th.[81] The use of "soft interventions" like this to reduce views of violative tweets, hashtags, and trending topics was a key part of Twitter's strategy.

Twitter also "denylisted" certain phrases from its search and trending topics functions, including "Stop the Steal," which Twitter permanently denied from appearing in search results on November 6th.[82]

Twitter also engaged in account-level deamplification: on November 7th, the company announced a "super-spreader" initiative" which demoted the visibility of tweets from accounts which repeatedly spread misleading information violating Twitter's civic integrity policy.[83] More than 2,500 accounts were deamplified under this policy in the first week.

[81] "Twitter's Response to Select Committee Staff Questions of March 16, 2022."
[82] "Twitter Policies and Enforcement: A Timeline of Events Surrounding the 2020 Election and the January 6th, 2021, Attack on the Capitol."
[83] *Id.*

IV. Meta (Facebook)

Facebook began preparing for the 2020 election more than a year in advance.[84] The company was worried that its service would once again be abused to mislead voters. It knew that the AI systems it used to proactively identify misinformation, hate speech, and violent incitement were insufficient. And it feared criticism of its enforcement actions from both sides of the political spectrum.

Among other changes, staff advocated for making "big ranking bets" to "drive down top-line misinfo" in users' newsfeeds. They experimented with promoting content that users reported as useful, important, or interesting. As early as February 2020, there were threats they warned would be difficult to prepare for in time for the election. While the 2016 election inspired Facebook to invest heavily in combating foreign information operations and "coordinated inauthentic behaviors," it had not made similar strides toward combating inauthentic activity from domestic actors (which might include "political parties, PACs, and social activist groups," as well as "dangerous organizations" and "commercially-motivated parties.") A ranking exercise early in the company's efforts found the risk of *domestic* inauthentic behavior being used to suppress votes or spread misinformation was "very high" because the company "was doing nothing to address it" in early 2020. They needed buy-in from leadership; but even if they could gain permission to tackle this problem, they forecasted that the threat would only fall from "very high" to "high."

Facebook's civic integrity team worked diligently to address competing, complex risk areas in advance of the election by developing 63 "break the glass" measures designed to slow the flow of viral, potentially harmful content.[85] These measures were the subject of ongoing internal debate and even resistance, but they eventually won over skeptics of these measures including CEO Mark Zuckerberg and other executives. After the election, their mood mixed triumph and relief.[86] Unlike in 2016, the company had weathered the election without landing in the center of a political firestorm.

This mood was short lived. The early warnings of "domestic inauthentic behavior" proved prescient. In the weeks after the election, a relatively small group of organizers coordinated to overcome the company's content moderation efforts and propagate hundreds of Facebook groups devoted to "Stop the Steal." Despite a "spirit of the policy" decision to remove the first Stop the Steal group, its replacements quickly became the fastest growing on Facebook. When the company removed one, the organizers would quickly reconstitute it with a backup, sometimes using fake accounts to bypass bans and other measures. Calls for violence were rife within these

---

[84] FB-CAP-00005233.
[85] FB-CAP-00009657.
[86] FB-CAP-00013763.

groups. Meanwhile, right-wing media went into overdrive supporting President Trump's "big lie," providing an endless source of outrage to fuel the movement.

Failure to swiftly make judgement calls and fill policy gaps limited Facebook's response to Stop the Steal. The company had no policy against using false claims to delegitimize the election, meaning moderators had to wait for other types of "strikes" to accrue in Stop the Steal groups before acting. When senior managers warned leadership to take Stop the Steal more seriously, the company demurred.

It is worth noting that Facebook's Stop the Steal groups "helped solidify the Stop the Steal movement's offline component. For example, on November 5, Facebook events were scheduled for locations including California; Virginia; Washington, DC; Pennsylvania; and Florida."[87]

Facebook's hesitance stemmed from long-running fears that even-handed policies aimed at objectively improving the quality of information and discourse on the platform would disproportionately impact the right-wing media ecosystem, angering Republican politicians. On many occasions since at least 2018, company leadership bent over backward to make policy exceptions for right-leaning outlets and individuals. More so than any profit-seeking pursuit of greater user engagement, this trend led to the company's failure to control activity on its service that ultimately contributed to the events of January 6th.

Facebook spent much of 2020 refining its policies against dangerous organizations in response to events that presaged the attack on the Capitol. Brian Fishman, Facebook's Director for Counterterrorism and Dangerous Organizations Content Policy, began to worry about an event like January 6th almost a year earlier during the Virginia Civil Discourse League's Lobby Day.[88] As armed individuals gathered at the Virginia state capital in Richmond, the FBI became concerned about "really nasty" online chatter by members of the Boogaloo movement who were present at the protest. Ultimately, the Bureau arrested plotters hoping to incite violence between protesters and police. Fishman said that these events caused "quite a bit" of reflection for him and his team.[89]

The Dangerous Organizations policy became more complex with time.[90] In the beginning, Fishman's work focused on hate groups, terrorist organizations, and organized criminal enterprises. Even expressions of praise and support for these actors was in violation of the policy. Later, it grew to include individuals like Alex Jones who spread hate speech—but support for Jones, or sharing of content from his website InfoWars, was not forbidden because they were

---

[87] "The Long Fuse: Misinformation and the 2020 Election." The Election Integrity Partnership. 2021. Pg. 87.
[88] See "2020 Richmond Lobby Day: Blueprint for a Violent Year," Anti-Defamation League, January 18, 2021, available at https://www.adl.org/blog/2020-richmond-lobby-day-blueprint-for-a-violent-year.
[89] Transcribed Interview with Brian Fishman, 26:24, 27:10.
[90] *Id.* at 110:15.

participants in political discourse.[91] (Public reporting claims that this distinction was a personal intervention by Mark Zuckerberg.)[92]

In February of 2020, the company publicly announced an update to the policy on dangerous organizations. It separated dangerous actors into three tiers. Tier one included violent organizations which target civilians, such as hate, criminal, and terrorist groups. Tier two included "violent non-state actors" which did not primarily target civilians. Tier one and tier two organizations and their members were prohibited from operating on Facebook. For tier one entities, any form of praise or vocal support was prohibited; for tier two, the company disallowed praise for explicitly violent acts.[93]

Actors like those on tiers one and two were banned from Facebook under previous versions of the policy, but tier three was more novel. It included "militarized social movements" and "violence-inducing conspiracy networks," which:

> "…repeatedly engage in violations of our Hate Speech or Dangerous Organizations policies on-or-off the platform or demonstrate strong intent to engage in offline violence in the near future, but have not necessarily engaged in violence to date or advocated for violence against others based on their protected characteristics."[94]

These tier three organizations included anti-government militias and hate groups which had not yet carried out explicit acts of violence, though they might inspire them. These actors were prohibited from having a "presence" or coordinating on Facebook's platforms, but the company did not explicitly prohibit praise or vocal support of them.[95]

Months later, in June 2020, an internal Facebook intelligence report warned of growing danger from QAnon activity on Facebook and other social media platforms. [96] The report noted that in 2019 the FBI identified the conspiracy theory as a potential domestic terrorism threat, and QAnon was becoming more popular within extremist militia groups such as the Oath Keepers. It noted at least a dozen instances where believers caused real-world harm through murder, attempted kidnappings, and other acts of violence. Troublingly, the authors warned that "QAnon believers have clearly indicated that they hope to influence the upcoming US elections; should conspiracy theories or other misinformation be perceived to sway the results, the company will face intense scrutiny." But there was also a countervailing risk: "given the significant number of

---

[91] *Id.* at 19:4.
[92] "Joel Kaplan's Policy Team Sways Big Facebook Decisions Like Alex Jones Ban," Buzzfeed, https://www.buzzfeednews.com/article/ryanmac/mark-zuckerberg-joel-kaplan-facebook-alex-jones.
[93] "Dangerous Individuals and Organizations," Meta Platforms, accessed May 22, 1988, available at https://transparency.fb.com/policies/community-standards/dangerous-individuals-organizations/.
[94] "Dangerous Individuals and Organizations," Meta Platforms, accessed May 22, 1988, available at https://transparency.fb.com/policies/community-standards/dangerous-individuals-organizations/.
[95] *Id.*
[96] FB-CAP-00001597.

politicians, to include the President, who have shared QAnon theories, the company may also face charges of anti-conservative bias based on its policy decisions."

On August 19, 2020, Facebook decided the first risk outweighed the second and applied the dangerous organizations policy to groups, pages, and Instagram accounts linked to QAnon, as well as to "offline anarchist groups that support violent acts amidst protests" and "US-based militia organizations."[97] The company removed thousands of groups, pages, accounts, and ads tied to QAnon and various militia groups and acted to reduce the reach and distribution of remaining accounts and hashtags on the platform.[98] This number swelled well into the tens of thousands through the end of 2020 and the beginning of 2021.

The Dangerous Organizations Policy is "actor-based"—it is not designed to affect individual pieces of content but rather to prevent entire organizations or movements from mobilizing on the platform at all. While actor-based policies they are less reactive than waiting for an organization to post something which violates policy, they are also more likely to generate false positives. [99] For this reason, Facebook has historically been hesitant to deploy them aggressively.

On October 19, 2020, the Tech Transparency Project published the result of its investigation into the ways in which Facebook allowed militia groups to not only exist, but actively recruit on the platform. Indeed, the study identified ads on the platform, and those ads reached tens of thousands of users: "the investigation identified more than 50 Facebook pages and groups dedicated to militia organizations, including some associated with the so-called Three Percenters, an anti-government extremist movement." The investigation found that "activity is continuing despite Facebook's announced action against 'militarized social movements'" in August 2020.[100] Some of the key takeaways from the investigation include:

- "Facebook for years allowed militia groups to run recruitment ads on the platform and profited from the activity. Some of the ads reached tens of thousands of users.
- As recently as October [2020], Facebook hosted an ad encouraging militias to attend a 'freedom march" in cities across the country just days before the election.
- At least 53 Facebook militia pages and groups [were] still active on the platform. Some of them even [had] the word 'militia' in their name.
- Facebook's recommendation algorithm [was] still pointing users who visit militia pages to other militia pages, potentially accelerating radicalization.
- Members of 'patriot' and pro-Trump Facebook pages [had] posted explicit threats to kill public officials and racial justice protesters."[101]

---

[97] FB-CAP-00000262.
[98] FB-CAP-00000262.
[99] Transcribed Interview with Brian Fishman, 15:3, 20:10..
[100] Tech Transparency Project, "Facebook Ran Recruitment Ads for Militia Groups" (October 19, 2020), available at https://www.techtransparencyproject.org/articles/facebook-ran-recruitment-ads-militia-groups
[101] *Id.*

Militia ads were not the only concerning ads that Facebook allowed on its platform. Between August 15, 2020, and Election Day, the Trump campaign ran over 750 unique ads across Facebook and Instagram, some of which questioned the integrity of the voting process.[102]

These echoed the sentiments of Trump's Facebook posts, some of which included[103]:

- "Big problems and discrepancies with Mail In Ballots all over the USA. Must have final total on November 3rd" (October 26, 2020; 156,569 likes; 22,318 comments; 8,299 shares)
- "A 3 day extension for Pennsylvania is a disaster for our Nation, and for Pennsylvania itself. The Democrats are trying to steal this Election. We have to get out and VOTE in even larger numbers. The Great Red Wave is coming!!!" (October 29, 2020; 144,209 likes; 14,947 comments; 11,174 shares)
- "The Election should end on November 3rd., not weeks later!" (October 30, 2020; 606,513 likes; 51,386 comments; 22,382 shares)
- "The Supreme Court decision on voting in Pennsylvania is a VERY dangerous one. It will allow rampant and unchecked cheating and will undermine our entire systems of laws. It will also induce violence in the streets. Something must be done!" (November 3, 2020; 108,521 likes; 13,136 comments; 6,921 shares)

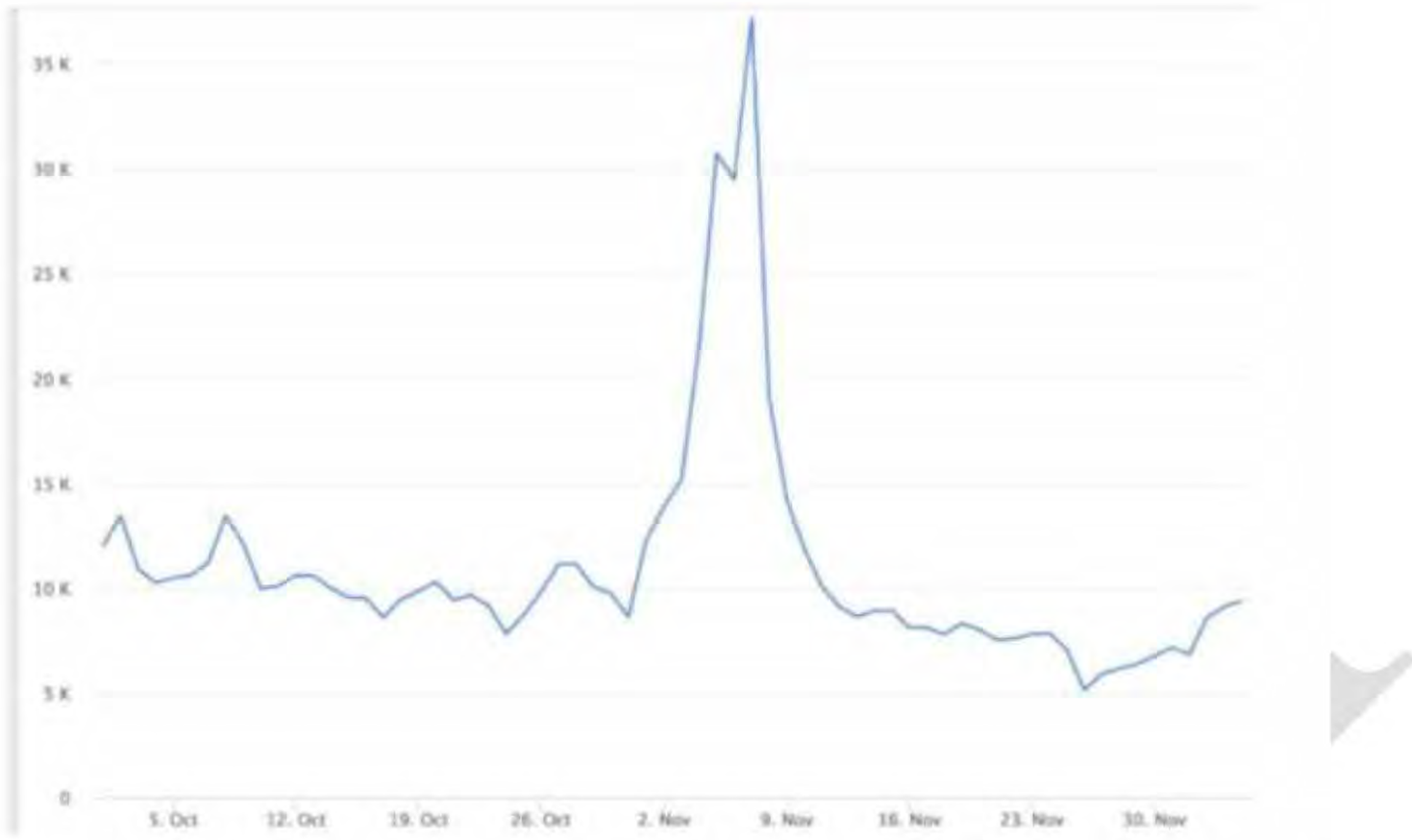
After the election, Facebook rolled back key protections and dismantled its civic integrity team while watching "Stop the Steal" grow. As Fishman's team worked to update and enforce the dangerous organizations policy, a different team, led by Samidh Chakrabarti under the management of Facebook Vice President for Integrity Guy Rosen, developed a set of "break the glass" measures designed to slow the spread of viral misinformation, hate speech, incitement to violence, and other threats to the election process on Facebook. Many of these measures came at the expense of slowing growth, a source of friction between Chakrabarti's team and other parts of the organization. These measures were debated robustly within the company, and while some of them were delayed or diluted, the company did ultimately deploy a robust set of protective measures before the election.

But the larger threat proved to be to the transition, not the election. In the days after the voting stopped, Facebook saw a significant spike in violence and incitement on the platform—by some counts, it rose as high as 45%.[104] As rates misinformation also rose significantly due to false claims of voter fraud, the company rolled out a second suite of "break glass" measures.[105] These included the use of a "News Ecosystem Quality" (NEQ) score to demote content from

---

[102] Dr. Jennifer Stromer-Galley, Statement for the record, submitted to the Select Committee on May 16, 2022.
[103] *Id.*
[104] "Facebook Has A Metric For 'Violence And Incitement Trends.' It's Rising," Buzzfeed News (November 5, 2020), https://www.buzzfeednews.com/article/ryanmac/facebook-internal-metric-violence-incitement-rising-vote.
[105] FB-CAP-00014390.

untrustworthy news publishers; as much as seventy percent of delegitimizing content from pages came from publishers with low NEQ scores.[106]



Violence and incitement on Facebook spiked immediately after Election Day *2020.*[107]

By the beginning of December, Facebook's metrics related to violent incitement had returned to pre-election levels, and the company felt comfortable rolling back the break glass measures. It set a target rollback date of December 8th.[108] After that date, at least thirty-four of the sixty-three break-glass measures were rolled back.

Not all of the break glass measured were rolled back at the same time, and some were not rolled back at all before January 6th. Several measures related to auto-deleting, demoting, and filtering content which might include incitement to violence, for example, were extended multiple times due to concerns about the prevalence of violence and incitement on the platform, especially in comments sections.[109]

The rollback process was touch-and-go as staff debated the merits of each measure.[110] There were mistakes: for example, on December 8th one staffer noted that three measures were deactivated "prematurely due to execution error"; the company declined to reactivate them because they were "not likely to be obvious" and it wasn't "worth the risk." These were a freeze on comments in groups with high rates of hateful and violent speech; a trigger to auto-disable

---

[106] FB-CAP-00011450; FB-CAP-00013022; FB-CAP-00013693.
[107] FB-CAP-00013552.
[108] FB-CAP-00015604.
[109] FB-CAP-00013552.
[110] A full list of the Break Glass measures deployed by Facebook, as well as the dates during which they were operative, has been made available to the Select Committee. This file is available as FB-CAP-00024827.

commenting in group threads with high rates of violent incitement; and a measure to prevent groups from changing their names to delegitimizing terms.[111]

Rollout and Deprecation of Select "Break Glass" Measures [112]

- Proportional demotion (5% confidence of violation -> 5% demotion), for violence and incitement and hate speech.
    - Launched for hate speech and graphic violence on 10/20 at 25%, boosted to 50% on 10/22, then to 100% on 10/28
    - Launched for violence and incitement at 50% on 10/26; increased to 100% on 10/29.
    - Reduced to 50% on 12/2, then deprecated on 12/3.
    - Relaunched in response to 1/6.
    - Deprecated on 1/25
- Demote videos designated "civic" from news pages with a low NEQ
    - Launched 10/12
    - Never deprecated
- Remove Feed boosts for non-recommendable Groups content
    - Launched 10/20 at 25%, boosted to 50% on 10/22, boosted to 100% on 10/28
    - Deprecated on 12/7
    - Relaunched in response to 1/6.
    - Deprecated 4/5
- Freeze commenting on posts in Groups that have a high rate of hate speech and violence and incitement comments
    - Launched 1/21 at 10%, then at 80% on 10/23
    - Deprecated on 12/1
    - Relaunched in response to 1/6
    - Deprecated 1/29.
- Remove all civic Grups from recommendations in "Groups you should join" to address low recall of groups associated with real-world harm
    - 10/20 launched
    - Never deprecated
- Filter low News Ecosystem Quality (NEQ) pages from Pages you May Like to prevent low quality and misinformation pages from becoming viral.
    - Launched 10/22
    - Reduced to 75% on 12/1, then 50% on 12/3, 25% on 12/8, and deprecated on 12/10.
    - Relaunched in response to 1/6

---

[111] FB-CAP-00013709.
[112] A full list of the Break Glass measures and their dates of activation/deactivation can be found in FB-CAP-00024827.

  - Deprecated 2/16
- Deploy the virality circuit breaker, which prevents the likelihood of URLs from new or unknown external domains that may contain misinformation from being boosted
  - 10/9 launched at 100x threshold
  - 10/23 launched at 25x threshold
  - 12/1 reduced to 75%, then 50% on 12/3, then 25% on 12/8
  - Deprecated on 12/10
- Remove all Groups created in the last 21 days from Recommendations in order to offset the low recall and detection of Groups potentially associated with violence and other harms.
  - 10/4 launched
  - Never deprecated
- Demote posts predicted to be hate speech at a 50% confidence level by 50%
  - 11/5 launched
  - 12/9 modified to keep on permanently with a confidence threshold of 70% (p70)
  - 1/6 relaunched at a confidence threshold of 50% (p50)
  - 1/25 modified to keep permanently
- Temporarily demote content that contains keyword matches for voter fraud or delegitimization claims. This measure was launched at 80% strength.
  - Launched 11/5
  - Deprecated 11/30
  - Relaunched on 1/6 for claim "Antifa was responsible for the violence at the Capitol"
  - Relaunched on 1/12 for claim "Donald Trump invoked or signed the Insurrection Act"
  - Deprecated 1/30
- Demote content from users who posted multiple pieces of third-party fact-checked misinformation in the past 30 days.
  - Launched 11/5
  - Reduced to 50% on 12/2, then deprecated 12/3
  - Relaunched 1/14
  - Deprecated 1/29.
- Demote low NEQ news and boost high NEQ news in order to increase the average quality of news in connected news feed
  - Launched 11/7
  - Reduced to 75% on 12/1, then 50% on 12/3, 25% on 12/8, and deprecated 12/10
  - Relaunched 1/13
  - Deprecated 2/16

Not everyone inside or outside the company agreed with the decision to roll these measures back. On December 3rd, for instance, Facebook leadership discussed a request from Senator Blumenthal that they maintain their protections for the general election through the Georgia

runoff on January 5th; their stance was that the break glass measures were for national level events which threatened to overwhelm their usual processes, and that a state election did not qualify.[113]

As the rollback process continued, Facebook moved to restructure the team behind the break glass measures. On December 2nd, Guy Rosen announced that the civic integrity team which led Facebook's election preparations would be reorganized into three pillars split across separate teams: one to deal with long-term responses to inauthentic behavior and other harms; one to support this work by creating tools and infrastructure; and one which develops active monitoring and mitigation strategies. [114] This announcement coincided with the departure of several team members from Facebook.

Facebook's stated purpose for the reorganization was to optimize its integrity efforts and integrate the lessons of civic integrity into the rest of the organization.[115] Others, such as Facebook whistleblower Frances Haugen, believe it inhibited integrity efforts at a critical juncture.[116]

Documents provided to the Select Committee suggest the reorganization was months in the planning, with Civic Integrity head Samidh Chakrabarti and other company leaders providing feedback on proposals in summer and fall of 2020. Internal company correspondence and interviews with other Facebook employees suggest that Civic Integrity was not beloved by other Facebook teams: it often clashed with the public policy team, for example, which had final approval over Civic Integrity's work, and it was not especially beloved by other integrity teams.[117] In an email exchange between Guy Rosen (VP for Integrity) and John Hegeman (the head of Newsfeed), Rosen noted that "we're exploring a few models in a very very tight group... rest assured that things like simplifying the relationship between Surface Teams and the central team (as well as the rocky relationship your team has with Civic Integrity specifically) are very top of mind."[118]

The tension between Civic Integrity and other teams was enough of a pain point that in Chakrabarti's biannual performance review, Rosen said that to "meet expectations" Charkabarti needed to have "zero drama in [cross-functional] collaboration." Chakrabarti felt this was a thinly veiled attempt to discourage him from "bringing dissenting viewpoints to critical topics," and noted it was a tactic Rosen used on others.[119]

---

[113] FB-CAP-00012137.
[114]FB-CAP-00005676; Memo on May 12, 2022, Select Committee Briefing with Meta.
[115] Memo on May 12, 2022, Select Committee Briefing with Meta.
[116] Transcribed interview with Frances Haugen.
[117] *See* memo on Select Committee interview with Sophie Zhang.
[118] FB-CAP-00007665.
[119] Facebook conducts performance reviews twice a year (see memo on Select Committee interview with Sophie Zhang); FB-CAP-00009657.

Public reporting claims that Chakrabarti went on leave around the time his team was restructured; he would later leave the company altogether in fall of 2021. On his return from leave in late January 2020, Samidh Chakrabarti forwarded feedback on Rosen's leadership of the integrity organization to Facebook's HR.[120]

In his feedback, Chakrabarti criticized Rosen for "prioritizing PR Risk over social harm"—a tendency which fatally inhibited the company's response to Stop the Steal. Chakrabarti believed that the "torrent of election delegitimizing content on Facebook" following the election could be as serious a threat to the company as the revelations about Russian election interference in 2016—but when Chakrabarti suggested the formation of a delegitimization working group to look into the issue, Rosen said no. Charkabarti quotes him as saying that even studying the problem would "just create momentum and expectation for action" that Rosen did not support.[121]

When Chakrabarti suggested the formation of a delegitimization working group to "study the torrent of election delegitimizing content on Facebook, a problem Chakrabarti believed "could be just as severe as the 2016 election challenges," Rosen told Chakrabarti that even studying the problem would "just create momentum and expectation for action" that he did not support.

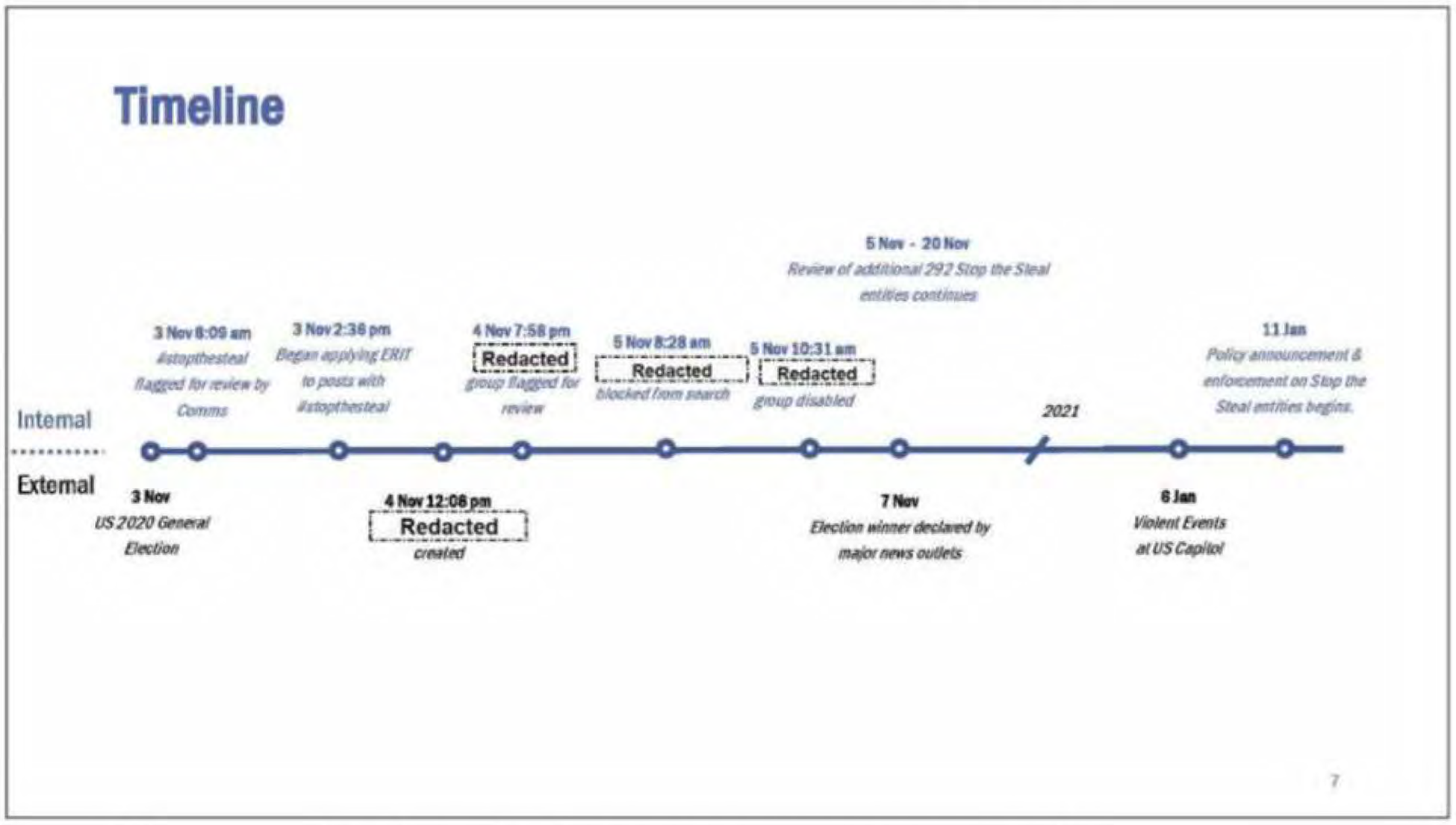
Meanwhile, Stop the Steal groups proliferated rapidly across Facebook and began to metastasize into a violent movement. During this period, Brian Fishman's Dangerous Organizations team closely following rallies in support of President Trump around the country, including the December 12th event in which Proud Boys participated. But it was not until "immediately after the first of January, when it became clear the rhetoric had changed across the web," that Fishman and his team began seeing signals he called "really concerning."[122] Fishman's fears became reality on January 6th, when a group of armed individuals breached the United States Capitol building. Some of them were involved with militia groups or other violent organizations, like the Proud Boys; others came motivated by their belief that the election had been stolen and determined to overthrow constitutional process to reverse it.

The individuals who participated in the insurrection that day had come to "Stop the Steal," a rallying cry which echoed across social media in the two months after Election Day. Facebook flagged the hashtag "Stop the Steal" for review the morning of Election Day—before voting had even concluded.[123] But the primary mechanism the movement grew through was Facebook groups.

---

[120] FB-CAP-00009657.
[121] FB-CAP-00009657.
[122] Brian Fishman Transcribed Interview 29:13.
[123] *Id.*

The first Stop the Steal group was created on November 4th by well-known figures on the right, including Ali Alexander and the Kremers (who organized the rally in Washington on January 6th that preceded the attack).[124] By the next day, November 5th, the initial Stop the Steal group had already accumulated more than 360,000 members and more than 7,000 posts, on which there more than 200,000 comments. Worried about calls for violence within the group, Facebook removed the group through what the company calls a "spirit of the policy" call—Facebook did not have a policy against election delegitimization, so instead, it made an ad hoc decision to remove the group due to the risk of violent incitement.[125]



Facebook-produced Timeline of Stop the Steal activity.[126]

As election delegitimating narratives spread throughout the conservative media ecosystem, the platform declined to act in a similarly decisive fashion for the copycat groups which quickly emerged.[127] It did, however, take the unusual step of restricting "Stop the Steal" from search results.[128] (Because searching for a term is an indication of user intent, Facebook is generally wary of restricting results from that feature.)[129]

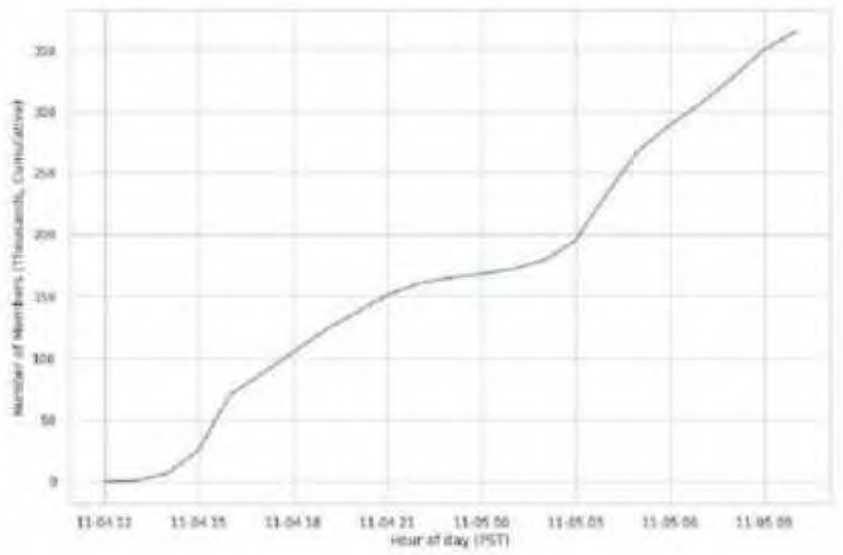
---

[124] FB-CAP-00001800, also available via Buzzfeed and whistleblower documents.
[125] May 18, 2022, Select Committee Briefing with Meta.
[126] FB-CAP-00013392.
[127] Facebook had election misinformation policies related to voter fraud but considered election delegitimization to be a matter of political opinion and was hesitant to act against it. See transcribed interview with Brian Fishman, p54 line 5. *See also* FB-CAP-00013392, May 18, 2022, Select Committee Briefing with Meta.
[128] FB-CAP-00013728.
[129] FB-CAP-00014770.

On November 7th, Facebook received an intelligence report monitoring activity on four other social media platforms: Parler, MeWe, Telegram, and MyMilitia. The report noted that the Proud Boys were initially angry at President Trump for losing the election, rather than at its alleged theft, and that Proud Boys leader Enrique Tarrio had been silent. The Three Percenters (or "Threepers), meanwhile, were mobilizing around nascent Stop the Steal protests.[130] Facebook continued to receive similar intelligence reports from a variety of sources. Most of these cover events during the month of November; by December 7th, as Trump's opportunities to overturn the election dwindled, Facebook's internal security team assessed that the protests were "losing momentum."[131]

Stop the Steal Group Information

- **Was live for 22 hours and 23 minutes.**
- Membership grew to 361,292
- 7,334 posts during the lifecycle of the group
- 206,196 comments
- **The majority of posts were made during the last few hours of the group's lifecycle**



24

Facebook-produced figure on the growth of the initial Stop the Steal group.

In the coming days and weeks, a bevy of groups arose to replace the original Stop the Steal group. They posed a quandary for Facebook: because it did not have a policy against election delegitimization with which to act swiftly and decisively, company staff were forced to wait for groups to accumulate sufficient strikes for content policy violations across a variety of other areas. Meanwhile, staff tracked the number of VPVs on the most far-reaching delegitimizing content as part of a regular rundown of problematic content on the platform; they did so for "situational awareness," noting that there was no policy against false claims of election fraud. The content flagged in this update received tens of millions of views on the platform.[132]

---

[130] FB-CAP-00004893.
[131] See FB-CAP-00004557, FB-CAP-00004567, FB-CAP-00004578, FB-CAP-00005489, and FB-CAP-00002563. For the claim that the protests were losing momentum, see FB-CAP-00010504.
[132] FB-CAP-00013420.

The growing size of the movement's offline adherents posed another challenge: as Brian Fishman told Select Committee staff, "what led to violence was misled people, informed by misinformation." Facebook was not the sole source of this misinformation: it came also from cable television personalities, other social media sites, and the President of the United States. Fishman argued that Stop the Steal represented a "genuine, misinformed view" and that banning election delegitimization would have been impossible without silencing a huge swath of the political right—including publishers, political figures, commentators, and other influential actors.[133] The company was in a political vice grip from which it could not escape without consequence.

But to Fishman's recollection, Facebook did not follow up on Chakrabarti's plea to study Stop the Steal or election delegitimization before January 6th.[134] Ultimately, 321 Stop the Steal groups proliferated across Facebook between election day and the end of November. Of these, Facebook took action against only forty-three.

Facebook took action against only 43 Stop the Steal Groups in November 2020.[135]

An internal Facebook report found that following the deletion of the first Stop the Steal group, the movement experienced "meteoric growth" as copycat groups sprung up to replace it. At one point, nearly all of the fastest growing civic groups on Facebook were related to Stop the

---

[133] Brian Fishman Transcribed Interview, 71:15, 92:14, 53:16.
[134] *Id.* at 77:15, 87-89.
[135] FB-CAP-00013392.

Steal.[136] These groups had more hate speech and violent incitement than other civic groups by a large margin.[137] Their membership also overlapped significantly with militia pages previously removed by the platform, which led Fishman to advocate that the company move more aggressively to take down Stop the Steal groups and content.[138] While the Stop the Steal movement had arguably not yet carried out acts of violence, it included calls for violence and unlike other protest movements, it was not explicitly *non-violent*; in Fishman's words, "they stood up next to folks that we knew had a track record of violence." Based on this, Facebook could have designated Stop the Steal a "violence-inducing conspiracy network" under a policy developed that fall to justify the removal of QAnon content. Facebook declined to take this step.[139]

As noted above, Facebook's automated systems were unable to detect harmful groups less than two weeks old—but as Facebook VP for Integrity Guy Rosen noted earlier that year, groups can grow exponentially in their first days after creation. This latency gave Stop the Steal groups time to balloon dramatically in membership before Facebook reacted. Stop the Steal was not unique in this aspect. Like antivaccine groups before it, Stop the Steal followed a well-established—and foreseeable—growth pattern.

A significant portion of this growth resulted from the work of "super-inviters": thirty percent of all invites to these groups came from just 0.3% of members.[140] These super-inviters were highly connected to one another—they frequently interacted through comments, messages, and tags. Initially, they may have been slowed somewhat by a break-the-glass measure limiting how many invitations to a group an individual user could send per day, but this measure was rolled back on December 16th.[141]

Super-inviters also circumvented these limits by using multiple accounts (a violation of Facebook's terms of service). When Facebook did catch and remove these groups, their organizers would often recreate them using backup groups and send invitations to the membership of the previous groups. As Fishman said to the Select Committee, this was typical of how dangerous groups grow, because "organizations always have organizers."[142]

---

[136] "Stop the Steal and Patriot Party: The Growth and Mitigation of an Adversarial Harmful Movement." A leaked version of this report is available via Buzzfeed here: https://www.buzzfeednews.com/article/ryanmac/full-facebook-stop-the-steal-internal-report. Another version, complete with graphics, has been provided to the Committee here: https://drive.google.com/file/d/1u10XeaKJ-XDe8P12n88LTwP0fA50mMF7/view?usp=sharing.

[137] "Stop the Steal and Patriot Party: The Growth and Mitigation of an Adversarial Harmful Movement."

[138] On this point, see also FB-CAP-00011546, which claims more than half of accounts involved in QAnon and Militiarized Social Movements belong to user segments related to conservative US politics and overlap with Stop the Steal.

[139] Brain Fishman Transcribed Interview, 54-60, 89.

[140] *Id.*

[141] Memo on May 12, 2022, Select Committee Briefing with Meta; FB-CAP-00024827.

[142] Brian Fishman Transcribed Interview, 82:9.

Fishman does not believe that the platform ever identified these organizers' offline identities—a task which might have fallen to his team. To his recollection he was not given a clear leadership directive to do so, and his priority after January 6th was to determine if the Oath Keepers and Proud Boys had organized violence directly on-platform. In the weeks after the election, Fishman's team tracked the Proud Boys and the Oath Keepers and found that both had difficulty coordinating on Facebook.[143]

What Fishman observed online during this period concerned him enough that he prepared for potential violence on the 6th. The day before, Fishman called for an IPOC, or a cross-team crisis coordination center, to be set up for the 6th.[144] If something happened, he knew employees would need to be on hand to take down livestreams of violent content, detect coordinated inauthentic behavior, and respond to any outbreak of violent incitement.[145]

On the morning of January 6th, Pacific Time, Fishman went for a walk as he took a regular standing phone call with Nathaniel Gleicher, Facebook's head of Security Policy. During the call, the two men received news that the United States Capitol building had been breached by an armed mob seeking to prevent the certification of the Presidential Election. Fishman ran back to his home, where he contacted the IPOC and began monitoring the situation.

As the as law enforcement moved to secure the Capitol building, Facebook staff began redeploying many of the election-related break-glass measures they rolled back in December.[146] Some of these—like the prohibition on recommendations for new and civic groups—would remain permanent.[147] Others were later rolled back again; these included the limit on bulk invitations like those that fueled Stop the Steal's growth.

Facebook's reengagement with the break glass measures continued for several days. On 1/6 or the day after, it relaunched an intervention which more powerfully demoted content evaluated as likely violative of the violence and incitement policy, keeping this policy in place until January 29th.[148]

On January 8th, Facebook delisted "Stop the Steal" from Groups search (In addition to the steps it took to remove the term from *main* search in November). This was the first time the company had ever intervened in the search function to return a "null result" for a term in Groups search. It did so because, in the words of one employee, the company recognized that Stop the Steal was "a

---

[143] *Id.* at 28:7.
[144] *Id.* at 15:9.
[145] *Id.* at 36:3.
[146] "Capitol Riots – BTG Response," provided to the Select Committee here: https://drive.google.com/file/d/1y5JbfZRolQsRWAPkTny6FJUO-LoyRFti/view?usp=sharing. See also FB-CAP-00024827 and FB-CAP-00016107.
[147] FB-CAP-0000282.
[148] FB-CAP-00034827.

potential gateway to violence and incitement," because "the only way to 'stop' would be extrajudicial, having exhausted all lawful remedies."[149]

On January 11th, Facebook began removing content containing the phrase "Stop the Steal" from the platform.[150]

The company also moved to put 8,500 users identified as a "bad risky actor cohort" (BRAC) in read-only mode, meaning they could not create groups, livestreams, or events, nor could they run pages. 700 of these were the admins and the 0.1% top commenters in pages removed as part of Facebook's recent policy against militarized social movements and violence-inducing conspiracy networks. The remaining 7,800 were followers of a specific Facebook page, the name of which remains unknown to Select Committee Staff.[151]

Because of the role Facebook Groups played in Stop the Steal's growth before January 6th, the fate of break glass measures controlling groups growth is especially significant. In the weeks after the attack, Facebook looked at the tradeoff between the speed of group growth and the risk of civic harm. Knowing what had already transpired, the company decided that the limit on group invitations should be deployed only in emergencies.

In February 2021, just weeks after the insurrection, Facebook's Growth team "urgently" requested the rollback of all groups-related break glass measures citing their negative impact on growth.[152] An assessment of these measures' impact noted that "lower invite limits had a consistently negative and [statistically significant] impact on confirmed invites (-16% to -29%), joins (-3% to -5%), and daily active groups (-1% to -2%) but not providing any [statistically significant] effect on any of the downstream metrics that we care most about."[153]

However, the authors *also* flagged Stop the Steal specifically as an example of the kind of risk mitigated by this measure, writing:

> "…we have anecdotal evidence that this measure does slow down fast-growing problematic groups: for example, we now understand that the infamous stop the steal group had 400 people rate limited on Nov. 4th and 250 people on Nov. 5th."[154]

In a briefing for Select Committee staff, Facebook indicated that this assessment was completed in February—nearly two months *after* the measure had already been rolled back. Facebook staff were dismissive of the claim that violent incitement, hate speech, and civic misinformation were

---

[149] FB-CAP-00016157.
[150] FB-CAP-00000352.
[151] FB-CAP-00006653; see also FB-CAP-00024827. The name of the relevant page has been redacted in documents provided to the Select Committee, likely for reasons related to the Stored Communications Act.
[152] FB-CAP-00020178.
[153] FB-CAP-00003700; Memo on May 12, 2022, Select Committee Briefing with Meta.
[154] FB-CAP-00003700.

widespread in quickly growing Stop the Steal groups which were largely unchecked by Facebook.[155] The company's focus on rapid growth and connectivity remains a liability.

Facebook staff reacted to the attack with anger and disappointment. Seeking to address staff morale, Facebook CTO Mike Schroepfer reshared a note to all employees from Mark Zuckerberg, adding his own cover note:

> "I'm saddened [sic] the attack on the most fundamental part of America: the peaceful transition of power. Hang in there everyone as we figure out the best ways to support our teams and manage discourse on our platform to allow for peaceful discussion and organizing but not calls for violence. I know I've had trouble focusing today as I'm watching events unfold. So if this is impacting you you [sic] are not alone. Hang in there."[156]

Many staffers reacted strongly and negatively. One wrote,

> "I'm struggling to match my values to my employment here. I came here hoping to affect change and improve society, but all I've seen is atrophy and abdication of responsibility. I'm tired of platitudes; I want action items. We're not a neutral entity."

Another pointedly noted that Facebook leadership failed to deploy all the possible measures to promote trust & safety on the platform:

> "You mention the list of things we've changed in the past few years but how are we expected to ignore when leadership overrides research-based policy decisions to better serve people like the groups inviting violence today. Rank and file workers have done their part to identify changes to improve our platform but have been actively held back."

To which another replied, "so many research-backed ideas get shut down. We need to do a better job making decisions from a research-first perspective."

Another drew a "straight line" from decisions made by Facebook years before to the events of January 6th:

.

> "Never forget the day Trump rode down the escalator in 2016, called for a ban on Muslims entering the US, we determined that it violated our policies, and yet we explicitly overrode the policy and didn't take the video down. There is a straight line that can be drawn from that day to today, one of the darkest days in the history of democracy

---

[155] Memo on May 12, 2022, Select Committee Briefing with Meta.
[156] "Comments on Zuck's Response to Capitol Riots," provided to the Select Committee here: https://drive.google.com/file/d/1rG2z3NTgcZ53kJJ0s5J4zmGEmWLDd8Px/view?usp=sharing.

> and self-governance. Would it have made a difference in the end? We can never know, but history will not judge us kindly."

Following the insurrection, Facebook filled gaps in policy and enforcement. The lack of a policy against election delegitimization meant that Facebook's integrity workers did not have a strong policy handle with which to address Stop the Steal's rapid growth.[157] In a policy proposal written after the January 6th attack, Facebook employees wrote that this policy gap meant "high-profile entities were able to serially spread such claims without crossing our falsifiable-misinformation threshold for enforcement."[158]

Facebook's lack of forethought on how to address coordinated harmful activity by groups of domestic users also allowed Stop the Steal to evade the company's moderation efforts. For years, the company prioritized investments targeting *coordinated inauthentic behavior* of the sort used in Russian efforts to interfere in the 2016 US Presidential election.[159]

Coordination itself is not a violation of Facebook policy (many social movements coordinate online), and not all harmful networks are inauthentic: recent movements like Stop the Steal, vaccine skepticism, and QAnon involve coordinated *authentic* activity, or purposeful collaboration among networks of users operating under their real identities.[160]

In 2021, Facebook working groups like the Disaggregating Harmful Networks Task Force addressed harmful coordinated authentic behavior with a new policy on "coordinated social harm." This policy allowed Facebook to act against networks of accounts which, while not inauthentic or otherwise violative of Facebook policy, engage in activity which heightens instances of hate speech, violent incitement, misinformation, and other harmful content.[161] However, the concept of "social harm" became difficult to objectively define or identify, so the policy was changed to "Coordinated Violating Networks"—group efforts to violate Facebook policies while evading enforcement. This shift made the company less reliant on a subjective, undefined notion of "harm" and allowed Facebook to draw on defined policies to identify adversarial networks.

Sometimes a central user (or group of users) sits at the core of these networks and may not engage in content policy violations themselves but coordinate or encourage others who do. These adversarial actors sometimes work to keep their content on just-the-right-side of a policy, or to

---

[157] Memo on May 18, 2022, Select Committee Briefing with Meta.
[158] FB-CAP-00001786.
[159] "Coordinated inauthentic behavior" is a term used by Facebook to describe accounts and pages which "work together to mislead others" about their identity and activities. See "Coordinated Inauthentic Behavior Explained," https://about.fb.com/news/2018/12/inside-feed-coordinated-inauthentic-behavior/. *See also* Memo on May 18, 2022, Select Committee Briefing with Meta.
[160] Memo on May 18, 2022, Select Committee Briefing with Meta; *see also* FB-CAP-00001786.
[161] *Id.*; In a document prepared by Facebook staff, authors recommended an "actor, network, and ecosystem" strategy for tackling the problem.

distribute strikes for violating policies across multiple accounts to avoid takedowns of content, groups, pages, etc. The new policy circumvented this problem by enforcing against a harmful *network*, rather than against individual accounts or pieces of content.[162]
The Coordinated Violating Networks policy dovetailed with the Dangerous Organizations team's policy against Violence Inducing Conspiracy Networks and Militarized Social Movements. Both teams struggled to deal with loosely organized movements which were mobilizing in ways that contribute to the risk of violence or other harms.

Stop the Steal took advantage of longtime struggles at Facebook to detect, moderate, and slow the growth of toxic groups. Groups are a "surface" of Facebook (like the newsfeed, events, or pages) which users can join to connect with other users interested in the group's theme. If a user joins a group, content from that group may appear in a user's newsfeed. Facebook may recommend groups to users through a feature called Groups You Should Join (GYSJ). Some groups are public—anyone can join—and others are private. If users are invited to a private group, they will have an opportunity to preview it. Groups also have administrators, who in some circumstances are responsible for approving members and posts in the group. Harmful content—from COVID-19 misinformation to hate speech—was rampant in "civic groups," or groups with a high degree of content commenting on political issues.[163] To identify civic groups, Facebook relied on AI systems to assess how much of a group's content was related to civic or political issues. At least according to one metric, a group was classified as civic if:

- The group's "subject elements" were civic;
- More than ten percent of the posts viewed were civic in the past seven days; or
- More than ten percent of the posts created were civic in the past seven days.

Once a group was marked as civic, it remained civic "for life."[164]

The labeling of civic groups was not foolproof: it relied on AI systems which struggled to properly label groups younger than 2-3 weeks old due to lack of available data. There was also a delay (or, "latency") when groups changed topics, and they could do so quickly. During the August 2020 protests in Kenosha, for instance, neighborhood discussion groups became, in the words of one Facebook employee, "a space for heated conversation."[165] Facebook's AI systems also suffered from "recall gaps," or shortcomings in their ability to identify problematic content including hate speech and violent incitement, especially in the comments under group posts.[166]

---

[162] Mem on May 18, 2022, Select Committee Briefing with Meta.
[163] FB-CAP-00010172.
[164] FB-CAP-00013164.
[165] FB-CAP-00010609; FB-CAP-00015560.
[166] "Recall" differs from "precision"; the former describes what fraction of a type of content an AI can identify from the total, while the latter describes how often an AI correctly discerns that a type of content matches what it is looking for. A low recall rate will produce false negatives; a low precision rate, false positives.

Notes from a September 29th document suggest that in at least some tests, civic classifiers had a false negative rate as high as thirty-five percent.[167]

Under some circumstances, Facebook will make groups non-recommendable to users so they will not show up in GYSJ if those groups have a risk of low-quality or toxic content but do not rise to the level of outright removal. At an August 2020 all hands meeting for Facebook's integrity organization (of which civic integrity was one component), presenters shared that most (around seventy percent) of the top 100 civic groups by viewport views (VPVs) were non-recommendable. Of these, the top three were dedicated to President Trump, right-wing commentator Candace Owens, and White House Press Secretary Kayleigh McEnany, and were flagged for high rates of hate speech and violence and incitement (V&I) in their comments (the group devoted to President Trump reportedly had five times more hate and V&I than any other group on Facebook).[168]

In August, Guy Rosen (Vice President for Integrity), Samidh Chakrabarti (head of the Civic Integrity team), and Tom Alison (who was then VP of Engineering) received an update stressing that "harmful civic groups grow faster than our integrity systems can handle… we are still very exposed to risk during US2020 as these groups proliferate and evade detection." In other correspondence, Facebook staff noted that groups can go from zero to millions of VPVs in a matter of weeks.[169] Facebook staff recommended setting up a team specifically to monitor fast-growing, potentially problematic groups in real time and creating protocols for dealing with harmful groups not detected by automated systems.[170]

Rosen replied to the update, acknowledging that the growth of problematic groups was a serious challenge. In response, Facebook's integrity organization created a groups task force to manage this problem and developed a tool to prioritize problematic groups for human review, called HERO-CO.[171] By December, the HERO-CO review process led to the removal of four hundred toxic and harmful civic groups with more than a billion VPVs, which Facebook called a "powderkeg risk."

The proposal Rosen received flagged that designating a point-of-contact on Facebook's policy team would be a necessary step toward implementing this project. Changes to Facebook's policies or product (as the platform is called internally) are reviewed not just by content policy or product teams for impact on user experience and engagement, but also by public policy staff for

---

[167] FB-CAP-00005842.
[168] FB-CAP-00009887. VPVs, or viewport views, are a metric by which Facebook usually measures how many users see a piece of content.
[169] FB-CAP-00008597.
[170] FB-CAP-00010702; see also FB-CAP-00011291.
[171] HERO-CO stands for "High Risk Early Review Operations for Complex Objects"; HERO is a more general process for reviewing viral content early in its spread. See briefing with Ryan Beiermeister. See also FB-CAP-00007267 for a general overview of HERO.

"defensibility" and "legitimacy," concepts related to PR and political risk. Rosen agreed that such a counterpart would be necessary next step.

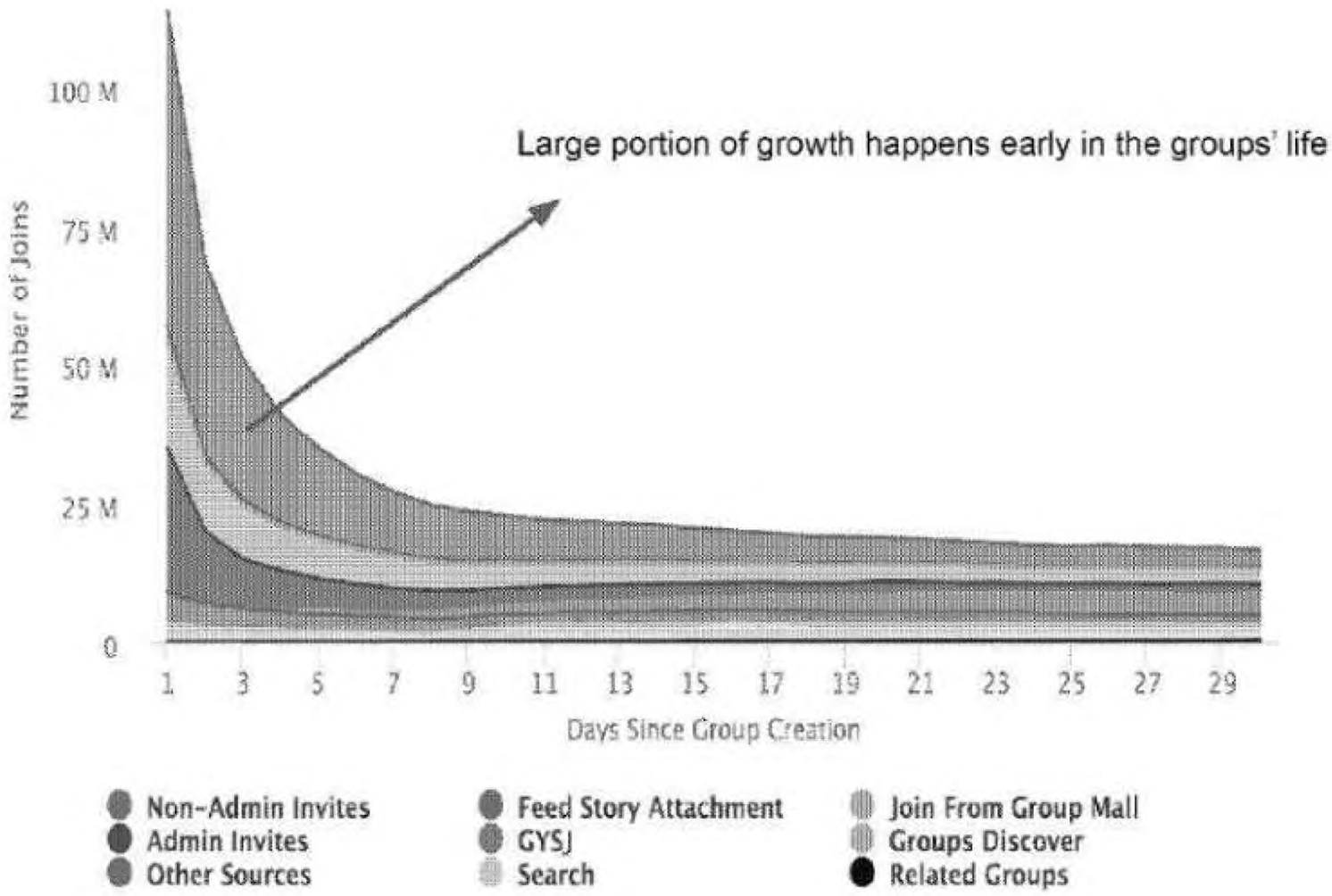
Rosen then set a separate email to a counterpart on the policy side of the company, writing that:

> "In parallel to reviews of groups which are being set up, the Groups Integrity team is figuring out how to put some brakes on the growth vectors these groups user [sic] to grow super fast (some of the largest civic groups are literally a couple of weeks old – so imagine these growing fast in October or November before we get our arms around them)… they seem mostly product levers (the usual growth vs. integrity tension – but Tom Alison is very supportive of slowing them down here), but I'm hearing that the people on your team working with Groups are nervous about a few aspects."

Andy O'Connell (then-head of Product Policy & Strategy) was skeptical. First, he noted that much of the content in question, while possibly concerning, did not necessarily violate a clear written policy. Indeed, the original proposal noted that the team would need to make "spirit of the policy" calls due to "known gaps in our protocols for enforcement," essentially foreseeing some of the challenges presented by Stop the Steal.

O'Connell also had reservations about slowing groups' growth by capping the number of daily invitations a user could send, recalling a prior case involving "groups with vaccine misinfo… 500 users accounted for 70,000 invites [to those groups]." In that case, Facebook COO Sheryl Sandberg was a "hard 'no'" on limiting the number of invitations those users could send (he did not recall why).[172]

[172] Rosen replied that the proposed limit rejected by Sandberg was on *users* in that instance, not on the groups themselves.



Facebook groups grow quickly in the days after creation, and invites are a major driver.[173]

Rosen received another email about groups less than a month later, on September 2nd, when a data scientist sent a note outlining what they believed were "urgent existential threats" faced by the company.[174] The sender was pessimistic about Facebook's prospects of successfully defending civic integrity before the election, and called for drastic action:

> "The US is in the midst of a violent crisis, less than 10 weeks form what will be the most divisive US election in modern history, and a great deal of the violence and division is playing out on Facebook. Our existing systems cannot catch even a small fraction of the hate, violence, or misinformation on Facebook. We have heavily overpromised regarding our ability to moderate content on the platform. We are breaking and will continue to break our recent promises about recommendations."

The sender suggested pausing all recommendations for groups through the 2020 election. They also asked Rosen to "drastically limit growth and connectivity on new entities leading into the election," and raised the above issue of "super-inviters" who send hundreds of group invites a day. Later in the exchange, another employee added that pausing groups recommendations would do nothing to address problematic groups that are already large. Given the violence of the summer, they wrote that:

---

[173] FB-CAP-00010441.
[174] FB-CAP-00012605.

> "It's totally reasonable to look around at the developments of the last two weeks and say, hey, facts on the ground actually did change. In particular, while we've long known about the horrible discussions happening in comments on our platforms, we've now clearly moved beyond the "it's just bad people talking" and into "and it's leading to offline violence…" we've made a number of commitments to the public that are trivially non-falsifiable, and that falsification exposes us to massive risk… we have no means to comply with these commitments in the next several months while they run the maximum risk of doing harm… finally, part of the reason for these proposals is *because* we don't seem to be able to find a sufficiently justifiable reason to take the groups down."

Another Facebook staffer described an "Armageddon scenario," in which a group reached tens of millions of VPVs and "dumped" misinformation on users before the election. Ironically, this threat played out *after* the election, imperiling the peaceful transfer of power, but did not meaningfully impact the voting process.

To Facebook's credit, the company took multiple steps to temporarily slow the growth of groups during the election period. Near the end of September, the company placed a pause on recommendations for groups less than 21 days old. Steps were also taken to require mandatory administrator approval for especially risky civic and health-related groups; a byproduct of this change was that Facebook could better assess the intent of group admins, a key factor in deciding whether or not a group should be taken down (groups with admins who intended to allow violating content can more justifiably be shut down). Additionally, Facebook limited "repeat offender" admins from creating new groups for thirty days and took steps to prevent the reemergence of previously removed "recidivist" groups.[175]

Ryan Beiermeister, who oversaw the workstream on abusive groups, announced these changes and laid out some of the biggest challenges in this area:

> "One of the biggest challenges in groups enforcement is the overall low recall of our classifiers for things like hate speech and V&I. That, and the fact that most of the concerning groups have very low rates of user reports – which we hypothesize is due to homogeneity in membership."[176]

Homogeneity in membership limited user reports because members of groups with frequent hate speech and violence may be less bothered by that content and therefore less likely to report it. This concentration of harm in particular communities was a vexing issue for Facebook, which was set up to respond to viral, widespread content but less so for harm concentrated in one community or user segment. Such "narrowcast" misinformation could be up to two or three

---

[175] FB-CAP-00004004.
[176] FB-CAP-00004004.

times as common as viral misinformation, but harder to detect.[177] This notion of harm which disproportionately occurred in certain communities was an ongoing research topic for Facebook, and later became an issue with Stop the Steal and election delegitimization.[178]

In another October update, Beiermeister also discussed a site event (Facebook parlance for a serious technical issue) during which groups did not receive strikes for violence and incitement for *months*. The update claims that hundreds of groups, profiles, pages, and accounts that should have been disabled were not, and once the issue was corrected more than 10,000 groups received strikes, with more than five hundred disabled as a result. The author of the update stressed that human review of the riskiest groups was an important process for holistically evaluating the integrity of groups as a surface.[179]

The company took further steps in mid-October: it would filter out recommendations not only for all groups less than 21 days old, but for all groups classified as civic in the United States (a position the Mozilla Foundation publicly called for around this time).[180] Both of these were considered temporary "break glass" measures around the election, though around March of 2021 both would be made permanent.[181]

The data behind this decision demonstrates the underlying problems with civic discourse in Facebook groups: 14.6% of impressions for civic groups in GYSJ were on groups later taken down for violating community standards, compared to 0.8% for non-civic groups. Similarly, users who encountered community standards violations were three times more likely to encounter hate speech, six times more likely to encounter violations of the "dangerous organizations and individuals" policy, and 1.6 times more likely to encounter misinformation in civic groups as compared with non-civic.

Facebook developed another Groups-related break glass measure in mid-October, limiting the number of invitations to a group a single user could send to 100, down from an initial limit which may have been as high as 2,250.[182] Unlike the measures related to recommendations, this one would not be made permanent; it was rolled back in December 2020 because of its harsh impact on groups' growth.[183]

---

[177] FB-CAP-00003914; FB-CAP-00003918.
[178] FB-CAP-00003925; FB-CAP-00003927; FB-CAP-00008803.
[179] FB-CAP-00010376.
[180] "Facebook Heeds Mozilla Call, Pauses Group Recommendations," Mozilla Foundation (October 30, 2020). Available at https://foundation.mozilla.org/en/blog/facebook-heeds-mozilla-call-pauses-group-recommendations. Facebook was already planning similar steps in advance of Mozilla's letter and discussed internally how to proceed without appearing to cave to outside pressure. See FB-CAP-00015337.
[181] FB-CAP-00003700; Memo on May 12, 2022, Select Committee Briefing with Meta; FB-CAP-00000282.
[182] The Select Committee has received inconsistent information on this limit. *See* FB-CAP-00003700; Briefing with Ryan Beiermeister.
[183] FB-CAP-00024827; Memo on May 12, 2022, Select Committee Briefing with Meta.

Analysis of Facebook activity by over three hundred January 6th defendants paints a picture of how individuals mobilized to violence engaged with these groups. The Select Committee's analysis found that January 6th defendants were largely *consumers*, rather than *creators*, of political disinformation on Facebook. However, many of the matched defendants consumed fringe Facebook content extensively, and some were members in hundreds of political Facebook groups.

In total, the Select Committee identified over 1,700 Facebook groups that contained at least one defendant. These groups were originally diverse in their causes but coalesced around keeping Trump in power and punishing Democrats and RINOs. Manual review of the posts that these groups contained revealed implicit and explicit calls for violence against political opponents in response to perceived grievances.

The Select Committee's investigation also demonstrated much higher levels of QAnon engagement among defendants than was established by previous studies that relied upon court documents as the source for ascertaining defendants' motivations. Only four percent of defendants were characterized as QAnon adherents based on those court documents. However, our analysis of defendants' Facebook activity, on the other hand, revealed that the real number of QAnon believers was at least three times as high. This is likely a conservative measure, as we exclusively identified defendant engagement with explicit QAnon rhetoric. While most of these QAnon groups were deplatformed by Facebook in October 2020, two groups containing defendants survived and posted increasingly violent material up to and through January 6th.

In the seven months leading up to the insurrection, QAnon-centric Facebook groups containing at least one defendant saw an average of 23 posts per day mentioning civil war, revolution, and/or 1776. Review by the Select Committee demonstrated that the vast majority of these were direct endorsements or promotions of civil conflict. Many posters asserted that political violence in any form was the last option available to Trump supporters in response to alleged Democrat crimes; many also suggested that Democrats were in league with Communists and foreign governments, further justifying civil war. All told, Facebook groups containing defendants clearly established environments where credulous and even hopeful discussions of civil war were tolerated, desensitized, and often supported.

The break glass measures also changed the type of content which was boosted or recommended to users on parts of the platform other than Groups. One important measure was "probable violating demotions," which one employee called "the most meaningful lever we've found to reduce hate speech prevalence."[184] This measure is simple: a computer judges the likelihood that a piece of content violates a policy, for example by containing hate speech, and then the content is demoted by a percentage directly corresponding to the computer's confidence in that

---

[184] FB-CAP-00012605.

assessment.[185] If a piece of content is judged ten percent likely to be hate speech, it is demoted in a user's newsfeed by ten percent.

The scale of activity on Facebook means that no amount of human review will ever successfully assess all the content on the platform in a timely manner, so the company uses machine learning systems to identify problematic content and groups. These systems are subject to a precision threshold, the height of which is very important: if it is too high, and the system will generate false negatives (harmful content or groups which remain online). However, if it is too low, there will be false positives (authentic political speech and communities which are wrongfully removed or demoted).

Because speech is contextual and political context is constantly changing, computers take time to learn and tend to be imprecise when making these judgements—in fact, they are much less precise than outside observers might expect. As one employee noted,

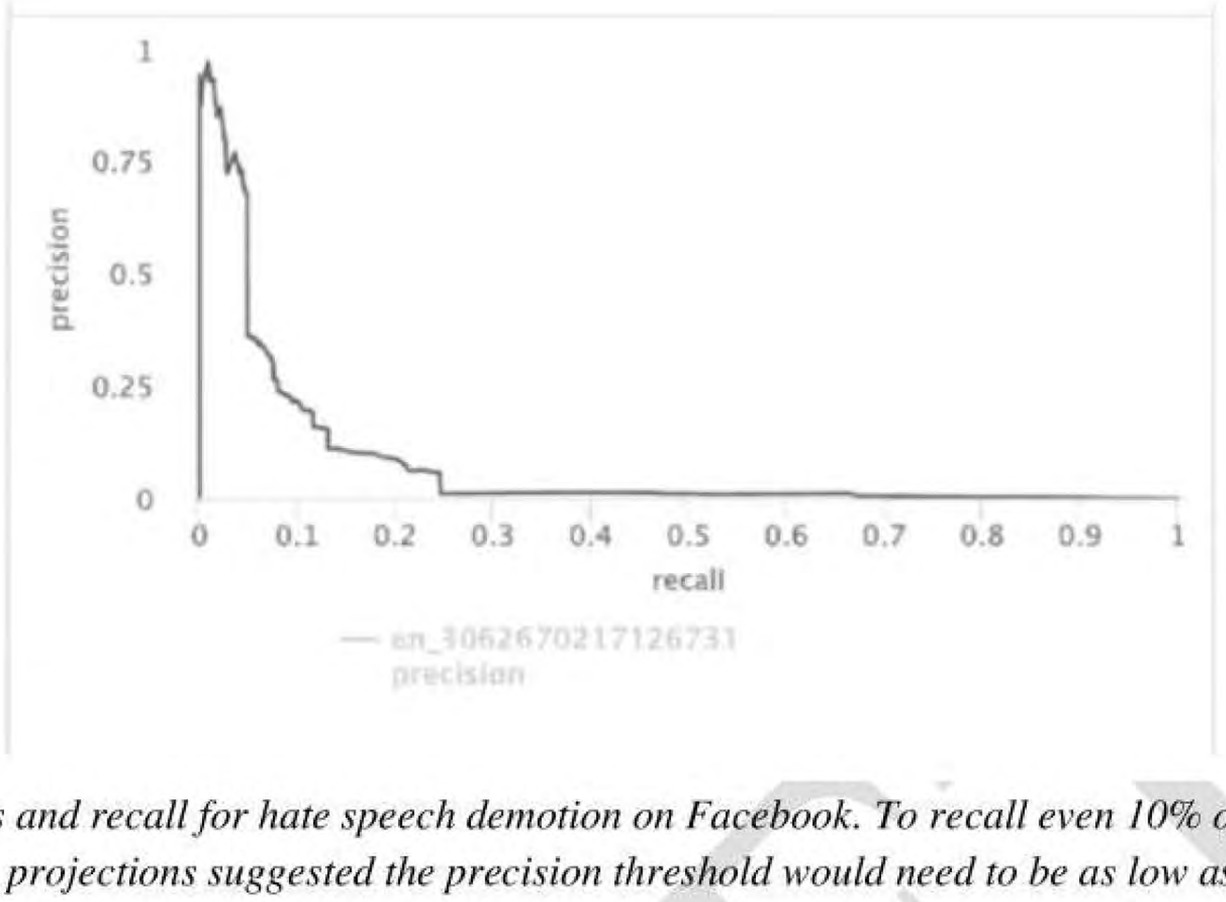
> "…at 95% precision, which is the threshold we use for auto-actioning, we get approximately 1% recall… to get even 10% recall we'd have to go below 25% precision. This is why demotion… worked so much better than anything else we tried—because by going down to low precision thresholds, we can catch up to 20x as much of the hate speech on Facebook as we currently catch with hard actions."

Further in the same email thread, this employee noted that the low precision rate has serious effects on how Facebook enforces policy against the violations it does discover. Before the election, Facebook lowered the number of strikes needed before a group could be removed, from five to three.[186] But if Facebook's automated detection catches only a small percent of hate speech or violent incitement—for example one percent—then three strikes might represent three hundred potential violations. This data scientist believed that this was the case and Facebook failed to catch the vast majority of this content.[187]

Others on the thread seemed to agree this approach was worth pursuing, or at least discussing. However, it had already been raised with Facebook leadership in other contexts. John Hegeman, Facebook's head of newsfeed, wrote that he would be "supportive" of this option, but that it would be "relitigating a prior decision" and that it "wasn't super likely" they would change it.

---

[185] Interview with Frances Haugen.
[186] Memo on May 12, 2022, Select Committee Briefing with Meta.
[187] FB-CAP-00012605.



Precision thresholds and recall for hate speech demotion on Facebook. To recall even 10% of hate speech on the platform, projections suggested the precision threshold would need to be as low as 25%.[188]

Hegeman was right that Mark Zuckerberg had already ruled on this topic. In April of 2020, Zuckerberg met with a task force looking at this and other "soft actions" Facebook could take prior to the election.[189] Zuckerberg initially the strength of this demotion out of concern for its impact on meaningful social interactions (MSI), a key metric for user engagement on Facebook. Eventually, he allowed demotions down to 100% as an emergency-only "break glass" measure.[190]

Many Facebook employees complained that leadership was often skeptical of potential interventions which might limit MSI, growth, or user engagement. Like probable violating demotions, many potent ideas were limited to emergency use only. For example, one break the glass measure penalized the reach of posts from users who frequently misinformation identified by third-party fact-checkers.[191] Others limited the spread of content based on its degree of separation from the original post.[192]

Many other interventions were also considered, and some were implemented as break-glass measure around the election.[193] These include:

---

[188] FB-CAP-00012605.
[189] "Mark Feedback on Soft Action Proposal," provided to the Select Committee here: https://drive.google.com/file/d/1USEfzcA6gKDVZpiuFQ9M6MEFtSNUwDNn/view?usp=sharing.
[190] Comparing "Mark Feedback on Soft Action Proposal" with FB-CAP-00014022 suggests that at least some of Zuckerberg's decisions were ultimately revisited with some success.
[191] Jackson, Conversation with Tom Cunningham; Select Committee interview with Facebook whistleblower. A version of Sparing Sharing may have been implemented following in the insurrection, in April 2021; see here: https://drive.google.com/file/d/15f6FeJvGRvs8CY-VMd-1IWCUYLbuw88V/view?usp=sharing.
[192] FB-CAP-00024827.
[193] For the complete list of Break Glass measures and their dates of activation and deprecation, *see* FB-CAP-00024827.

- Demoting posts from users who repeatedly shared misinformation; this lever was considered especially powerful and may have reduced misinformation VPVs by 42%. This measure was launched on November 5th, deprecated on December 3rd, and relaunched in response to January 6th before being shut of permanently at the end of January.

The Repeat Sharer demotion appears to have been especially effective.[194]

- Removing amplification levers for content from non-recommendable groups in users' newsfeeds; this resulted in a 4.4% reduction in VPVs of misinformation and other violations.[195] This lever was launched on October 20th and sequentially strengthened before being deprecated on December 7th, relaunched in response to January 6th, and finally deprecated at the beginning of April.
- Removing new and civic groups from recommendations;[196] some of these measures remained permanent.
- Freezing commenting in groups with high rates of hate speech and violence and incitement in the comments sections, as well as lowering the precision threshold for detecting hateful comments by twenty percent.[197] This measure was launched in late

---

[194] FB-CAP-00003994.
[195] FB-CAP-00003982; FB-CAP-00011450.
[196] FB-CAP-00014022.
[197] FB-CAP-00014022.

October, deprecated on December 1st, relaunched in response to January 6th, and finally deprecated at the end of January.

- A "Virality Circuit Breaker" which slowed the distribution of content with URLs linking to unknown domains—John Hegeman, the head of News Feed, wanted this to become permanent, but it did not.[198] It launched on October 9th and was deprecated on December 10th.
- Using Facebook's "News Ecosystem Quality" (NEQ) score to demote content from news publishers was delayed by Mark Zuckerberg in October due to concerns about altering what news sources people saw in the days before the election. This measure was launched after the election on November 7th, deprecated on December 10th, relaunched in response to January 6th, and then finally deprecated in mid-February.[199]
- Using the NEQ score to filter pages with no or low NEQ scores from recommendations surfaces.[200] This measure was activated earlier than the above, on October 22nd, and deprecated on the same date as the above.

Some of these measures were delayed, or diluted, but the civic integrity team and other Facebook staff worked diligently to implement what changes they could before election day.[201] In December, that team was restructured and many of these measures were rolled back only to be reimplemented after January 6th.

The world will never know what political conversations on Facebook would have looked like if these measures had remained in place through December and into January. Still, they do seem to have had a positive impact on election discourse on the service. Many if not most could feasibly have remained in place forever—but their function constrained the speed of growth, connectivity, and activity on the service. The legacy of the break glass measures is a testament to Facebook's ability to pump the brakes, and to the consequences of flooring the accelerator.

Many of the break glass measures were "soft interventions," which may be important for deterring the migration of users to more radical fringe platforms. A soft intervention is anything short of removal or a ban—demoting a post's algorithmic distribution, for example. Efforts to reduce exposure to harmful content have become more difficult with the emergence of fringe platforms which provide safe havens for activity that is not allowed elsewhere. After the January 6th attack, a third-party intelligence report drafted for Facebook describes how a combination of social media bans and law enforcement activity drove the migration of users from mainstream platforms to fringe, "alt" platforms with fewer content moderation standards.[202] Others moved to web forums which "specifically cater to extremist views." According to the

---

[198] FB-CAP-00011450; FB-CAP-00014717.
[199] FB-CAP-00011450; FB-CAP-00013022.
[200] FB-CAP-00014022.
[201] For an example of conversation around approving these measures, see FB-CAP-00013022. See also "Mark Feedback on Soft Action Proposal," provided to the Select Committee here: https://drive.google.com/file/d/1USEfzcA6gKDVZpiuFQ9M6MEFtSNUwDNn/view?usp=sharing.
[202] FB-CAP-00000855.

report, "researchers have already observed members of more violent neo-Nazi groups use these channels… to recruit disaffected QAnon conspiracy followers."[203] The migration of extremist users to these spaces inhibits both industry and law enforcement efforts to monitor and disrupt their activity.[204] Facebook received several similar reports from different sources on this ongoing migration.[205]

Nathaniel Gleicher, Facebook's head of security policy, told Select Committee staff in a briefing that when the company systematically removes content from people who believe what they are saying is true, the feeling that they have been censored reifies their beliefs.[206] A draft academic paper provided to the Select Committee corroborates these concerns: they found that the "great deplatforming" following January 6th led to "substantial intentional movement" to alternative platforms, especially Gab. As new users flocked to Gab, discourse there came to include more hate speech and more claims of voter fraud in the 2020 election.[207]

If widespread bans lead to platform migration and significant blowback, another possible course of action is greater reliance on "soft interventions"—anything short of a ban, like the forms of filtering and downranking employed as part of Facebook's "break glass" measures. These interventions allow platforms to limit the exposure of vulnerable individuals to misinformation and extremist content, giving them the opportunity to "move away" from extremism. Policy transparency and consistency also help ameliorate the perception of censorship—though these are undermined when platforms make exceptions for powerful political figures, personalities, and movements.[208]

Those kinds of exceptions typify the fears and concessions that limited Facebook's appetite for acting against Stop the Steal. The January 6th attack on the United States Capitol did not occur in a vacuum. As one Facebook employee told Select Committee staff, it was an anomalous event which nearly had a catastrophic outcome, made more likely by years of Facebook's failure to adopt stronger integrity measures. These failures had at least two key drivers. First, the company's organizational structure subordinates integrity teams to the policy team, which oversees both content policy and public policy—a clash of incentives that compromises decision-making on integrity issues in ways that may be unique to Facebook or are at least unusual among its peers. Second, the company feared allegations of bias from right-wing politicians, and for years the desire to avoid political reprisals has shaped Facebook policy choices in ways which reverberate across the political and media landscape.

---

[203] *Id.*
[204] *Id.*.; Memo on May 18, 2022, Select Committee Briefing with Meta.
[205] See FB-CAP-00004607, FB-CAP-00004641, FB-CAP-00004392, FB-CAP-00004081, and FB-CAP-00004631.
[206] May 18, 2022, Select Committee Briefing with Meta.
[207] "Effects of the Post-January 6th Deplatforming on Social Media Discourse," Buntain, Innes, Mitts, and Shapiro (February 2, 2022). Draft on file with Select Committee.
[208] May 18, 2022, Select Committee Briefing with Meta; Summary memo of conversation with J. Johnson

Perhaps no day better exemplifies this patter than May 28, 2020, when—during the protests following the murder of George Floyd—President Trump posted on Facebook that "I just spoke to Governor Tim Walz and told him the Military [sic] is with him all the way… any difficulty and we will assume control but, when the looting starts, the shooting starts."[209]

Trump's post placed Facebook in an uncomfortable position. Its staff, the public, and Democratic lawmakers were livid. On the other hand, the White House and its allies had already taken aim at the company for allegedly suppressing Republican viewpoints.

Public reporting has since revealed that after Trump's post, Facebook's head of global policy, Joel Kaplan, prepared three options for Mark Zuckerberg: first, the post could be removed as incitement to violence. Second, it could be interpreted as a discussion of the state's use of force, which was permissible under Facebook's community standards. Third, it could be read as a mere *prediction* of violence—also permissible.[210]

Later that day, Zuckerberg spoke to the President directly by phone.[211] They agreed that the post would remain online; a short while later, Trump posted a second time. In this post, he claimed his previous post was "spoken as a fact, not as a statement" and that he didn't "want this to happen"—an apparent overlap with Kaplan's second and third options for Zuckerberg.[212]

For his part, Zuckerberg addressed the post on Facebook, saying:

> "We looked very closely at the post that discussed the protests in Minnesota to evaluate whether it violated our policies. Although the post had a troubling historical reference, we decided to leave it up because the National Guard references meant we read it as a warning about state action, and we think people need to know if the government is planning to deploy force. Our policy around incitement of violence allows discussion around state use of force, although I think today's situation raises important questions about what potential limits of that discussion should be. The President later posted again, saying that the original post was warning about the possibility that looting could lead to violence. We decided that this post, which explicitly discouraged violence, also does not violate our policies and is important for people to see."[213]

According to Dmitry Borodaenko, a data scientist formerly employed by Facebook, head of content policy Monika Bickert defended Zuckerberg's decision to outraged staff by arguing the

---

[209] "Facebook Employee Leaks Show Betrayal By Company Leadership," Buzzfeed News, https://www.buzzfeednews.com/article/ryanmac/facebook-employee-leaks-show-they-feel-betrayed
[210] "The Infinite Reach of Joel Kaplan, Facebook's Man in Washington," Wired (March 2020), https://www.wired.com/story/facebook-joel-kaplan-washington-political-influence/.
[211] Trump and Zuckerberg share phone call amid social media furor, Axios (May 31, 2020), https://www.axios.com/2020/05/31/trump-facebook-zuckerberg-phone-call.
[212] *Id.*
[213] May 29, 2020, post by Mark Zuckerbeg, Facebook, https://www.facebook.com/zuck/posts/10111961824369871.

rationale for not removing the post was the same as Facebook's policy allowing videos of police violence: the content was newsworthy, and therefore exempt from removal. Staff pointed out the difference between a call for violence and documentation of state violence against civilians. In a later meeting, Zuckerberg echoed Bickert's argument despite staff objections.

Borodaenko believes this "newsworthiness exception" was Facebook's solution to a recurring problem: the need to find fresh policy justifications for avoiding action against Donald Trump's account. He dated this pattern all the way back to 2015, when Facebook declined to remove Trump's campaign announcement in which he proposed banning Muslims from entering the United States. Recent reporting claims that although many conservative staffers felt the video violated the company's hate speech policy, Kaplan advocated against removing it to avoid inviting "outrage from conservative America."[214]

Weeks before the January 6th attack on the United States Capitol, Kaplan again advocated for bending policy to avoid the ire of President Trump and his supporters. On November 17th, 2020, Zuckerberg testified before the Senate Judiciary Committee on social media and misinformation. During the hearing, he said that Facebook would not change its approach to President Trump's account after he left office; a company spokesperson said the company does not fact-check world leaders but would fact-check the President's account after his term ended.[215]

The following day, representatives of Facebook's policy team debated this ambiguity. Nick Clegg, Vice President for Global Affairs, said that he thought they had already confirmed to press that Facebook "deem[s] politicians who are no longer candidates or in office as eligible for fact-checking, ie [*sic*] there would be a significant change in our treatment of Trump." Kaplan responded:

> "This is likely true, but I think we should not rush to make this public commitment before we have fully thought through the consequences and options. Under our currently policies, Trump will likely be in almost immediate [repeat offender] status, and see his distribution massively reduced while he effectively and indefinitely remains the leader of the Republican Party. We may not want to contemplate it, but the reality is he will still be quite a unique user of our platform and applying our existing flawed [third-party fact-checking] program to him on day one will cause tremendous difficulties (as, of course, would not applying it)."[216]

---

[214] "The Infinite Reach of Joel Kaplan, Facebook's Man in Washington," Wired (March 2022), https://www.wired.com/story/facebook-joel-kaplan-washington-political-influence/.
[215] "How Twitter and Facebook plan to handle Trump's accounts when he leaves office," New York Times (November 17, 2020), https://www.nytimes.com/2020/11/17/technology/how-twitter-and-facebook-plan-to-handle-trumps-accounts-when-he-leaves-office.html.
[216] FB-CAP-00012275. It is likely that "RO status" means "repeat offender status," and 3 PFC refers to "third party fact-checking."

These incidents are part of a troubling pattern. During the summer of 2020, Borodaenko discovered a whitelist of political accounts insulated from fact-checking measures.[217] The whitelist had malfunctioned or been accidentally deactivated; as a result, Facebook was issuing strikes to about seventeen politicians each day for spreading false information on Instagram, when ordinarily they would receive none. Further investigation unearthed a document showing more than thirty instances where an account was due to receive a strike for spreading false information but was flagged for review by Facebook's public policy team. Public policy removed the strikes, protecting those individuals and pages from being placed under repeat-offender status and having their content demoted. In one instance, Joel Kaplan intervened directly to reverse a strike against right-wing personality Charlie Kirk.[218]

Almost all the accounts on this list were right-of-center. They included:

- Candyce Owens
- Glenn Beck
- Donald Trump Jr.
- Eric Trump
- PragerU
- Dennis Prager
- Turning Point USA
- Charlie Kirk
- The Epoch Times
- The Gateway Pundit
- FreedomWorks
- Alliance Defending Freedom

While Joel Kaplan's role at Facebook is frequently criticized, multiple former staffers told the Select Committee that Mark Zuckerberg is happy with Kaplan's role at Facebook and that attention paid to Kaplan ultimately shifts responsibility away from Zuckerberg.[219]

Another incident in this pattern is from 2019, when Facebook included Breitbart as a trusted partner in its "News Tab, "which features content from included outlets for which the company compensates them. Angry staffers challenged company leadership to justify Breitbart's inclusion and pointed to internal metrics on the trustworthiness of news publishers, which showed that Breitbart was among the least trusted by Facebook users. Campbell Brown, head of global news partnerships at Facebook and a member of Kaplan's policy team, claimed that Breitbart had

---

[217] Summary memo of interview with Dmitry Borodaenko.
[218] This story has also been reported on publicly. See "Facebook Fired An Employee Who Collected Evidence Of Right-Wing Pages Getting Preferential Treatment," Buzzfeed News (August 6, 2020), https://www.buzzfeednews.com/article/craigsilverman/facebook-zuckerberg-what-if-trump-disputes-election-results.
[219] See summary memo of interview with Tom Cunningham.

changed and deserved a second chance.[220] Breitbart's inclusion as a News Tab partner became an issue of controversy again in 2020, when staff called the company to remove it due to its incendiary coverage of the protests following George Floyd's murder.[221]

As part of its investigation, the Select Committee performed an analysis of Facebook data to test the claims made by Republican critics that Facebook actively suppresses conservative voices. The analysis yielded results at-odds with those claims. The Committee found that between October 1, 2020, and January 6th, 2021, 47% of top-ten Facebook posts were those posted by Donald Trump alone (based on impressions). Among the posts that appeared in Facebook's daily top-ten posts between October 1st and January 6th, right-leaning posts appear in Facebook's daily top **ten three times as often** as left-leaning and nonaligned posts combined. Moreover, right-leaning posts reached seven times as many Facebook users than left-leaning posts.

This is an even more concerning trend when considering the nature of the content shared by right-wing influencers and in right-wing groups. Specifically, right-wing Facebook communities have been proven to share significant volumes of content coming from sites which repeatedly post incendiary, divisive, unreliable content (sometimes called "repeat offenders").[222] For example, the Carter Center conducted analysis of 871 Facebook groups between August 17, 2020, and January 20, 2021, inclusive of both left- and right-wing groups. Approximately 75% of those groups contained links from established "repeat offenders"; repeat offender content was shared more often within right-leaning Facebook groups.[223] Moreover, 98.7% of members of the right-leaning Facebook groups in the sample were exposed to links to repeat offender content.[224]

Tom Cunningham, another data scientist formerly employed at Facebook, described Mark Zuckerberg's 2018 speech at Georgetown University as a philosophical turning point for the company.[225] While Facebook made significant strides in integrity work after the 2016 US

---

[220] Summary memo of interview with Dmitry Borodaenko; see also "Facebook will begin paying some outlets for their content as it introduces its News tab," Washington Post (October 25, 2019), https://www.washingtonpost.com/technology/2019/10/25/facebook-will-begin-paying-some-outlets-their-content-it-introduces-its-news-tab/.

[221] Summary memo of interview with Dmitry Borodaenko; see also "Facebook chose to keep Breitbart on News Tab and gave it special treatment — even after employees warned of its embellished and hyper-partisan coverage of events like the George Floyd protests," Business Insider (October 24, 2021), https://www.businessinsider.com/facebook-files-breitbart-news-tab-employee-objections-2021-10.

[222] "Repeat Offender," in this context, is distinct from, but related to, the designation given by Facebook to outlets which independent fact-checkers find to push false content multiple times, leading to reduced distribution for their content. This is similar to the terminology used by the Election Integrity Partnership (EIP), convened by Stanford Internet Observatory and the University of Washington's Center for an Informed Public and joined by the National Conference on Citizenship, Graphika, and the Digital Forensic Research Lab. The EIP uses the term "repeat spreaders," which refers to accounts or individuals that, in relation to the 2020 election, "regularly shared false claims or misleading information about voting procedures."

[223] Michael Baldassaro, Katie Harbath, and Michael Scholtens." The Big Lie and Big Tech Misinformation Repeat Offenders and Social Media in the 2020 U.S. Election." The Carter Center. August 2021.

[224] *Id.*

[225] "Zuckerberg: Standing for Voice and Free Expression," October 17, 2019. https://www.washingtonpost.com/technology/2019/10/17/zuckerberg-standing-voice-free-expression/.

Presidential election, Cunningham characterized this progress as lacking a coherent strategy or vision. After Zuckerberg's Georgetown speech, however, Cunningham said that Facebook employees were told not to use terms like "trust" or "quality," which reflected objective assessments of content, for fear of blowback from publishers like Breitbart and aligned political leaders. Instead, Zuckerberg wanted the company to ground itself in subjective measures of user value, a stance framed publicly as libertarian and pro-free expression.

Cunningham was frustrated by the company's insistence on using subjective indicators of quality determined by signals from users. The company had already developed objective indicators of the quality of news sources and content but using them carried political risk: interventions based on these signals disproportionately affected the political right. Sensitive to the political ramifications, Facebook's policy team favored "fairness in outcome" over "fairness in decision-making" and avoided viewpoint-neutral interventions which disproportionately penalized right-leaning content.[226] Borodaenko corroborated this view, saying that Facebook's approach "replaces objectivity with balance."[227]

Unable to use objective signals of content quality, Facebook's integrity professionals began to justify their work through counterintuitive arguments about what users *truly* value, despite the signals they send.[228] Promising integrity interventions such as demotion of "deep reshares" (content that is shared multiple degrees of separation away from its original source) and network centrality (a way of ranking news sites by those frequently linked to by others) were rejected or reversed.\ A similar fate awaited "sparing sharing," an intervention which penalized the distribution of posts from users who frequently shared misinformation. Cunningham called this a "reasonably good signal" and a form of "duct tape" used to clean up Facebook newsfeeds following the 2016 election. These measures appear to have been converted into emergency only "break-glass" measures—but Cunningham's account suggests they were at one time live, or at least proposed as permanent fixes rather than temporary solutions.[229]

In September 2020, a departing employee wrote that they had "seen promising interventions from integrity product teams with strong research and data support be prematurely stifled or severely constrained by key decision makers—often based on fears of public and policy

---

[226] Summary memo of interview with Tom Cunningham; "Last Day at Facebook," provided to Select Committee here: https://drive.google.com/file/d/1qxAc-ML_nJFiPaYm2tBWqNhoq_hGTlzZ/view?usp=sharing.
[227] Summary memo of interview with Dmitry Borodaenko.
[228] Facebook's head of News Feed, John Hegeman (someone Cunningham considered generally friendly toward integrity efforts), hoped to replace engagement-driven metrics like meaningful social interactions (MSI) with new metrics like the "feed satisfaction survey" after the election. He hoped that this would ensure News Feed teams had "a strong incentive not to over-optimize for the [engagement-based] proxy metric even when it isn't creating a great experience." VP for Integrity Guy Rosen was also excited about these changes—they agreed that MSI might be "cool for cat videos" but was bad for civic content. This would appear to be in line with the trends Cunningham identified, although it may have also helped the company avoid "prioritizing engagement over integrity," which appeared to be a priority for Hegeman. See FB-CAP-00007665 and FB-CAP-00011549.
[229] Summary memo of interview with Tom Cunningham.

stakeholder responses." They cited sparing sharing as an example of an already functioning safeguard rolled back and alleged that because of "fears over potential public and policy stakeholder responses," Facebook was "knowingly exposing users to risk of integrity harms."[230] In December 2020, Cunningham announced to his colleagues that he was leaving the company despite challenging work and great pay. He made this choice because Facebook was having a "net negative influence on politics in Western countries," and company leadership was not "involved in a good faith effort to fix this" despite "pockets of excellent work inside the company." He cited research showing that Facebook was a net driver of political polarization in the United States, and wrote that he had,

> "…seen a dozen proposals to measure the objective quality of content on News Feed diluted or killed because either (1) they have a disproportionate impact across the US political spectrum, typically harming conservative content more; or (2) they cannot be framed in terms of subjective quality ("what the users want"). There has been a big push in integrity towards basing demotions on subjective quality, and this has led to (in my opinion) a tangled and inefficient approach to integrity, where demotions must nominally target subjective quality, but actually target objective quality. I think that Facebook already does, in practice, take a position on objective quality, and always will, but it's very averse to admitting that publicly, and that causes things to be far more complex and inefficient than they need to be."

Finally, Cunningham wrote that he felt Facebook's content policy decisions were "routinely influenced by political considerations" to "avoid antagonizing powerful political players." He complained that while the company's content-policy processes could "easily be made independent," they instead went through the Public Policy team. [231]

This arrangement—which Cunningham felt was peculiar to Facebook—was a common point of contention among integrity professionals. In June 2020, after President Trump posted that "when the looting starts, the shooting starts," Kaushik Iyer—an Engineering Director who worked on civic integrity at Facebook—wrote a workplace note to his colleagues recommending that the policy organization, which was headed by Joel Kaplan and included both content and public policy, be broken up to separate out the incentives for decision-makers.[232]

Samidh Chakrabarti, head of the Civic Integrity team, similarly suggested on workplace that public and content policy should "live in separate orgs." Vice President for Integrity Guy Rosen later told him that "such talk did not befit a leader at the company."[233] In July 2020, Chakrabarti described the friction between his team and the policy organization as a "lowlight" for his

---

[230] FB-CAP-00005348.
[231] "Last Day at Facebook," provided to Select Committee here: https://drive.google.com/file/d/1qxAc-ML_nJFiPaYm2tBWqNhoq_hGTlzZ/view?usp=sharing.
[232] FB-CAP-00010299.
[233] FB-CAP-00009657.

colleagues. Chakrabarti also listed "key decisions," such as the newsworthniess exception and the decision not to fact check politicians, as lowlights alongside the blocking of new protections for at-risk countries due to new policy requirements. In January 2021, he would reflect that Rosen rarely stood up for them in cross-team meetings: Rosen had instructed the team to focus on implementing rather than influencing policies, which "marginalized the role of Product Management within the integrity org" and "led to worse product outcomes."[234]

Facebook's senior leadership was aware of these concerns. In September, talking points for Chief Operating Officer Sheryl Sandberg at the company's integrity summit included a list of "watchout" topics. One of these was "continued requests to 'break up' the policy org"; another warned that Facebook employees viewed the company's "hate speech challenges as policy problems, not detection" or "enforcement gaps"; employees felt that the company's publicly provided statistics on hate speech removal were misleading and that policies were written so narrowly that a great deal of harmful content remained on the platform. Others were concerned about the platform's role in increasing political polarization.[235] (Notably, years before Joel Kaplan killed a project called "common ground" that hoped to address polarization on the platform.)[236]

Facebook's tolerance of increasingly radical speech and hyper-partisan media may have accelerated polarization and extremism in the United States. In his transcribed interview with Select Committee staff, Brian Fishman said politicians have normalized and instrumentalized political violence. In his words, "the more our politics use violent, militarized rhetoric, the harder it is to distinguish a real threat from an exhortative political claim," and "we're in a place where this kind of thing is going to be more common."[237] Fishman fears there will be violence in the 2024 election, if not before.[238]

He also worries that irresponsible rhetoric from political leaders has placed the burden of arbitrating political discourse on unelected social media companies which are ill-positioned for this task with any sort of legitimacy, saying that "the unwillingness of political leadership to set reasonable boundaries on what is acceptable speech and what is not defers responsibility to social media companies in ways we just should never want to give it to them."[239]

In assessing social media's contribution to the January 6th attack and the rise of right-wing extremism, it is important not to lose sight of the offline drivers of the erosion of democratic norms in the United States. The mob which attacked the Capitol received hyper-partisan, misleading, and incendiary messages on social media—but they also received them from

---

[234] FB-CAP-00007284; FB-CAP-00009657.
[235] FB-CAP-00010567.
[236] Brian Wofford. *The Infinite Reach of Joel Kaplan*, Wired (Mar. 10, 2022), available at https://www.wired.com/story/facebook-joel-kaplan-washington-political-influence/.
[237] Brian Fishman Transcribed Interview 118:7, 115:9.
[238] *Id.* at 123:5.
[239] *Id.* at 118:7.

traditional media (especially cable news television) and from the President of the United States and his political allies.

Social media is just one part of this larger media ecosystem, but it is an increasingly significant one.[240] It matters when large platforms reward certain kinds of discourse. In 2018, Facebook staff published an internal report claiming that representatives from political parties across Europe came to the platform with concerns about changes to Facebook's ranking algorithms, which determine what types of content users are likely to engage with and boosts that in their feeds. The parties found that after those changes, they were rewarded more highly for angry, negative posts about their opponents—and so an increasing proportion of their messaging struck that tone. But privately, they asked the companies if their pages could be governed by a previous version of the algorithm which less handsomely rewarded vitriol and outrage.

Later, Facebook would consider decreasing the weight it gave to "angry" reacts in its ranking algorithm—however, cognizant of the fact that US political campaigns factored those reactions into their communications strategies, the platform discussed holding off until after the election.[241]

The experience of those European party officials has an eerie parallel in American politics. Eric Barber, a state lawmaker from West Virginia who participated in the attack on the US Capitol, later told Select Committee staff that the beginning of his radicalization was when he noticed his campaign received more engagement on social media if they posted angry, negative messages. Stepping back from 2020 and the events of January 6th, it is possible to imagine that the cumulative effect of charged political discourse online contributed to the mainstreaming of radical extremist ideas that led individuals to attack the US Capitol on that day. Social media reflects the state of our politics, but it is also a partial driver of them.

In the absence of democratic legitimacy, Facebook struggles with the tradeoffs of content policy. Integrity work can involve complex tradeoffs. Many of Facebook's break glass measures used imprecise computer algorithms to guess whether or not content is violative and either take action against that content or, in some cases, escalate it for human review. Because of the sheer scale of content, humans could not manually review more than a fraction of the content every day, and review capacity was a concern for Facebook during the election.[242]

During a briefing with the Select Committee, Facebook stressed that this approach carries a high risk of false positives which it accepted temporarily during the election period. The company

---

[240] *Id.* at 119:7, 120:2.
[241] "Political Party Response to '18 Algorithmic Change," a document provided to the Select Committee by an independent whistleblower. See also FB-CAP-00001781 and FB-CAP-00013172.
[242] *See* FB-CAP-00010105, which notes that "Global Operations capacity is nearly maxed out"; the COVID-19 pandemic strained Facebook's content moderation efforts and combined with the election created a bottleneck where the supply of staff struggled to meet demand.

said that these false positives (content, accounts, pages, or groups which are wrongly removed or demoted) have real consequences in the form of communities and conversations disrupted.[243]

As other Facebook employees and documents mention, not just protecting but *maximizing* user "voice" is a key priority for the company.[244] A June 2020 planning document for integrity efforts called "increasing focus on protecting voice an "intrinsic part of integrity work" and highlighted the creation of a team to study overenforcement of Facebook policies.[245]

That same month, CEO Mark Zuckerberg himself was so concerned about overenforcement that he requested briefings on the subject from senior Facebook employees after the company acted against a group Zuckerberg joined.[246] Zuckerberg wanted the company to move toward a posture where the values and intentions of groups held more sway over how Facebook enforced policy within that group. In discussing this, one employee suggested that the company might not enforce certain "low-severity violations," such as ""low-tier hate speech," in private groups at all.

Despite its concerns with user voice and avoiding the perception of censorship, Facebook, as a corporation, is not a public forum.[247] It is not an impassive carrier of information— Facebook actively shapes and influences what is at the top and bottom of users' newsfeeds. In some cases, it does so for integrity and safety reasons, and often it does so to boost engagement with content on the platform (which, in turn, allows it to serve more advertisements). In internal correspondence, Facebook employees recognize the company's elevated responsibility to moderate content when Facebook is actively boosting it.[248]

While hard interventions—the removal of accounts, groups, or content from the platform—undeniably limit "voice," soft interventions such as downranking arguably do not because users have no constitutional right to boosted distribution on Facebook's service. Facebook's core value of "maximizing voice" rarely seems to grapple with this distinction or with whether the risk of false negatives is in fact higher than the risk of false positives.[249] When asked to comment on the

---

[243] Memo from May 12, 2022, Select Committee Briefing with Meta.
[244] Brian Fishman Transcribed Interview, 70:11.
[245] FB-CAP-00008372.
[246] FB-CAP-00008372; FB-CAP-00008650; FB-CAP-00008907.
[247] Ironically, surveys of Facebook users show that while perception of censorship is increasing, it is a "low-intensity" harm which users ranked as less bothersome than bullying or other negative online experiences. See FB-CAP-00011151.
[248] FB-CAP-00000282.
[249] As stated above, Beiermeister did not dwell on this argument. It is important to spell out, however, because Mark Zuckerberg's 2018 "pivot" to free speech coincides with a series of policy decisions that reflect fears that accusations of anti-right-wing censorship would lead to regulatory consequences for the company. See, for example, memos summarizing Select Committee interviews with former Facebook data scientists Tom Cunningham and Dmitry Borodaenko.

distinction between hard and soft interventions and their respective impacts on speech, one Facebook representative merely stated to Select Committee staff that both seem "pretty bad."[250]

It is true that Facebook lacks the political legitimacy of a truly public civic space—even though if often feels like one to its users. Brian Fishman and other platform employees interviewed by the Committee believe that companies can help make up for this legitimacy gap by becoming more transparent about the steps they take to remove or demote potentially violative content on the platform.[251] Fishman also said that he believes navigating the tradeoffs between false positives and false negatives should involve broader conversations between government, industry, and civil society.[252] He warned that collective reflection on the use of AI to make these kinds of decisions around online speech are an imperative, because an event like January 6th "will happen again."[253]

[250] Memo from May 12, 2022, Select Committee Briefing with Meta.
[251] Brian Fishman Transcribed Interview, 76:16.
[252] *Id.* at 75:6.
[253] *Id.* at 62:5.

V. Alphabet (YouTube)

On November 17, 2020, Facebook CEO Mark Zuckerberg and Twitter CEO Jack Dorsey testified before the Senate Judiciary Committee on their handling of misinformation during the 2020 election, leading some observers to ask: "Where is Susan Wojcicki, YouTube's CEO?"[254] Despite widespread election misinformation on YouTube, the company receives relatively little public scrutiny compared to Facebook and Twitter. Some speculate this is simply because video is more difficult to analyze than text.[255] YouTube is, however, worthy of further investigation.

The video sharing platform nets over 14 billion visits each month, from 1.7 billion unique users.[256] It is second only to Google among the most-visited websites in the world.[257] A Pew Research Center study revealed that 73% of U.S. adults say they use YouTube, situating the video platform atop the user rates of all other social media, including Facebook, Instagram, Snapchat, Twitter, and WhatsApp.[258] And while platforms like Twitter, Facebook, and Instagram qualify as social networking sites because of their emphasis on relationships and connections between users, YouTube is in a league of its own as a vessel for influence. Videos generated by content creators tend to be hyper-focused, resonating with viewers' interests, needs, and struggles, and the dissemination of new content on a weekly or even daily basis enables YouTube creators to have a frequent and sustained presence in a viewer's life.

Unlike Twitter and Facebook, the Select Committee found YouTube's preparations for the election to differ little from its general practices. While the company did apply a unique label to videos which made claims regarding election fraud, though this label was more generic and less prominent than other platforms. YouTube published a new policy against election delegitimization after the December 9th deadline for states to certify their electoral college votes. The policy did not apply retroactively, allowing videos published in the month between the election and December 9 to continue attracting viewers.

Since 2019, the company has downranked content assessed to be "borderline" in its recommendation and search features. However, Alphabet representatives told Select Committee staff that it was not possible to assess how many views are received by borderline videos in specific policy areas like election delegitimization, because the company does not retain that data.[259] Similarly, counsel for Alphabet confirmed that the company does not track data or

---

[254] Evelyn Douek, *Why Isn't Susan Wojcicki Getting Grilled By Congress?* Wired (Nov. 17, 2020), available at https://www.wired.com/story/why-isnt-susan-wojcicki-getting-grilled-by-congress/.
[255] "Election misinformation continues staying up on YouTube," *The New York Times* (November 10, 2020). https://www.nytimes.com/2020/11/10/technology/election-misinformation-continues-staying-up-on-youtube.html.
[256] "Digital 2022 Global Overview Report." Hootsuite. https://hootsuite.widen.net/s/gqprmtzq6g/digital-2022-global-overview-report
[257] id.
[258] Perrin & Anderson, 2019
[259] Notes from July 13, 2022, Select Committee Call with counsel for Alphabet.

produce trend assessments related to specific topics, such as violence and incitement that occurred at specific junctures between Election Day and January 6th.[260]

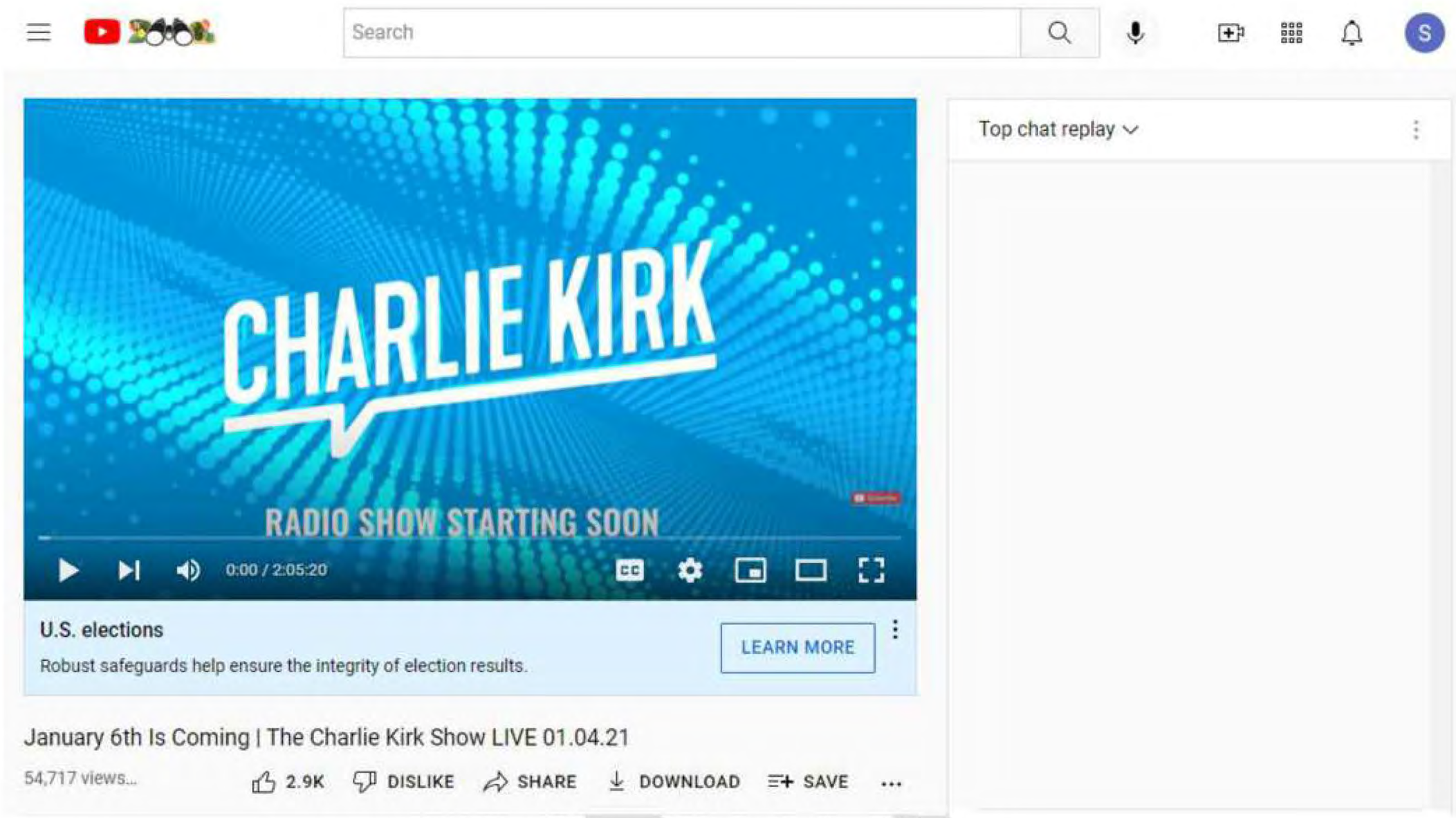
YouTube's content moderators remove violative content and demote "borderline" content in recommendations. The company relies on a mix of automation and human review; before it releases a new policy, it trains algorithms on similar content so it can proactively recognize it. This process is iterative, and the algorithm improves with time. When the algorithm is not able to make an assessment with a high degree of confidence or precision, the decision is reviewed by humans; Alphabet now employs more than 22,000 content moderation staff for these purposes. When content is assessed by humans, YouTube uses nine-member panels of independent reviewers to make decisions. When these panels find that content does not violate YouTube policies but comes close, they designate this content "borderline." Borderline content is demoted in search and recommendations.[261]

YouTube's election strategy included efforts to boost authoritative content. During the election, the company boosted authoritative news to the top of its search results; seventy percent of search results for election-related content were from authoritative sources. There is no static list of authoritative news producers—instead there is a dynamic list of authoritative content. The company uses a range of signals to designate content "authoritative," including inputs from Google News; independent review panels on misinformation; the reputation of the outlet or creator; and whether or not it is satire.

YouTube also labeled election fraud claims—but did so anemically. Like Twitter and Facebook, YouTube labeled election fraud content in lieu of removing it. Unlike Twitter and Facebook, it is not clear from available evidence if YouTube experimented with the size, placement, and language of these labels. Facebook, for instance, found that the exact wording of a label and its perceived neutrality—or lack thereof—can greatly affect how users perceive it. Twitter's labels were located prominently above the relevant content, and some were interstitials requiring users to click through them to view violative tweets. YouTube's labels are relatively small, located below content, and carry a relatively neutral message.

---

[260] *Id.*
[261] Memo on May 16, 2022, Select Committee Briefing with Alphabet.



YouTube video page with election fraud label.

When asked, briefers from Google were not able to provide information about whether these labels were experimented with or evaluated for the effectiveness of their text, positioning, color, or other factors related to their content or appearance.

YouTube's election fraud policy did not apply retroactively and did not result in account suspensions until after January 6th. On December 9, 2020, YouTube put into place a policy against videos claiming election fraud. This date was chosen because it was the day after the "safe harbor" deadline for states to certify their federal election results. From 12/9 – 1/6, the company removed more than 2000 videos for violating this policy.[262]

Notably, the company did not apply this policy retroactively. The decision to retroactively apply a policy and remove violative videos *already on the platform* is made on a case-by-case basis. In this case, YouTube chose the safe harbor deadline as the final date to debate the integrity of the voting process—but it considered claims of election fraud before December 9th to be permissible political discourse and left them on the platform after the policy was instituted.[263] This provided YouTube the appearance of neutrality but allowed continuing damage to faith in the election process.

---

[262] Memo on May 16, 2022, Select Committee Briefing with Alphabet.
[263] *Id.*

Ordinarily, videos in clear violation of YouTube policy receive a warning, followed by "strikes" to the uploader's account. Strikes last for ninety days.[264] Each strike is followed by a suspension: one week for the first strike, two weeks for the second, and then a permanent suspension for the third. For new policies like the election fraud policy, however, YouTube maintains a thirty-day grace period during which it will remove videos in violation of its policy but issue no strikes. YouTube followed this practice for election fraud content, meaning the grace period ended on January 7th—the day after the attack on the US Capitol. Further, the policy took a narrow view of what constituted "election fraud." It primarily covered claims of irregularities in the voting process; other claims that the election was administered "illegally" or was otherwise illegitimate were not considered violative.[265]

Additionally, YouTube did not take steps to ban the term "Stop the Steal" as an ipso facto violation of its policies until January 11, 2021.[266] Just as with its broader election fraud policy, this reflected an approach that only belated considered election denialism to be an urgent threat, even after it was willing to adopt a more muscular approach following the safe harbor deadline.

President Trump's account suspension was an exception to YouTube Policy made in exigent circumstances. On January 6th, President Donald Trump's account received a warning for violating YouTube's election fraud policy after posting a video of his Rose Garden speech. YouTube removed the video but did not issue a strike because of the grace period (which later ended one day early on January 7th). On January 12th, his account was suspended for one week in accordance with standard YouTube policy for violating YouTube's policy against incitement to violence.[267] On January 26, however, YouTube took the unusual step of extending that suspension indefinitely.[268] On March 4th, YouTube CEO Susan Wojcicki announced that YouTube would maintain this suspension until it judged the danger of violence had passed.

Google officials told Select Committee staff that they will make that assessment using signals from a variety of sources, including government statements, law enforcement activity, violent rhetoric on YouTube, and intelligence reports from Google's own teams.[269] Select Committee staff confirmed with the briefers that there is no benchmark or set of metrics that would inform YouTube's assessment that the risk of violence has receded; it is purely a judgement call based on available signals.[270]

No discussion of YouTube could be complete without addressing its highly controversial recommendation algorithm. A key debate for those who study the use of YouTube by internet

---

[264] GOOG-HSCI-00000001.
[265] Memo on May 16, 2022, Select Committee Briefing with Alphabet.
[266] Notes from July 13, 2022, Select Committee Call with Alphabet Counsel.
[267] GOOG-HSCI-00001076.
[268] "YouTube extends ban on Trump amid concerns about further violence," *The Guardian* (January 27, 2021), https://www.theguardian.com/us-news/2021/jan/26/youtube-trump-ban-suspension.
[269] "A conversation with YouTube CEO Susan Wojcicki," Atlantic Council (March 4 2021), https://www.atlanticcouncil.org/event/youtubes-wojcicki/.
[270] Memo on May 16, 2022, Select Committee Briefing with Alphabet.

subcultures and reactionaries more specifically is the power of algorithms versus the power of the more structural, systemic issues with what is allowed on the platform. The power of the algorithm represents a key debate across the various disciplines with which social media, extremism, and tech intersect, wherein some argue that the platform's curation mechanisms enabled extremists to thrive on YouTube, while others believe that YouTube would be a hotbed for radical ideas even without the algorithm.

Aware of the above concerns, YouTube made over thirty revisions to its recommendation algorithm in 2019.[271] One change diversified the topics of recommended videos. Others altered how the algorithm's uses the hundreds of signals that are used in ranking videos, including likes, dislikes, and survey results. [272] These changes also demoted "borderline" content in recommendations.[273]

External academic analysis of these changes and their impact is not conclusive.[274] There does seem to be emerging consensus that YouTube's recommendation systems do not drive ordinary, non-radicalized users down deep rabbit holes toward overtly extremist content. However, hardline, right-of-center conflict does seem to be "stickier" than content on the political center or left; viewers who consume it spend more time watching political content than corresponding users of different perspectives, and the recommendations of users who consume radical content still deliver content from far-right and channels and influencers. Meanwhile, YouTube serves as an important content repository for far-right users across the internet: many individuals arrive at borderline content on the platform not through recommendations, but through links to YouTube content posted on Facebook, Twitter, and Reddit, as well as across alt-tech and fringe platforms..

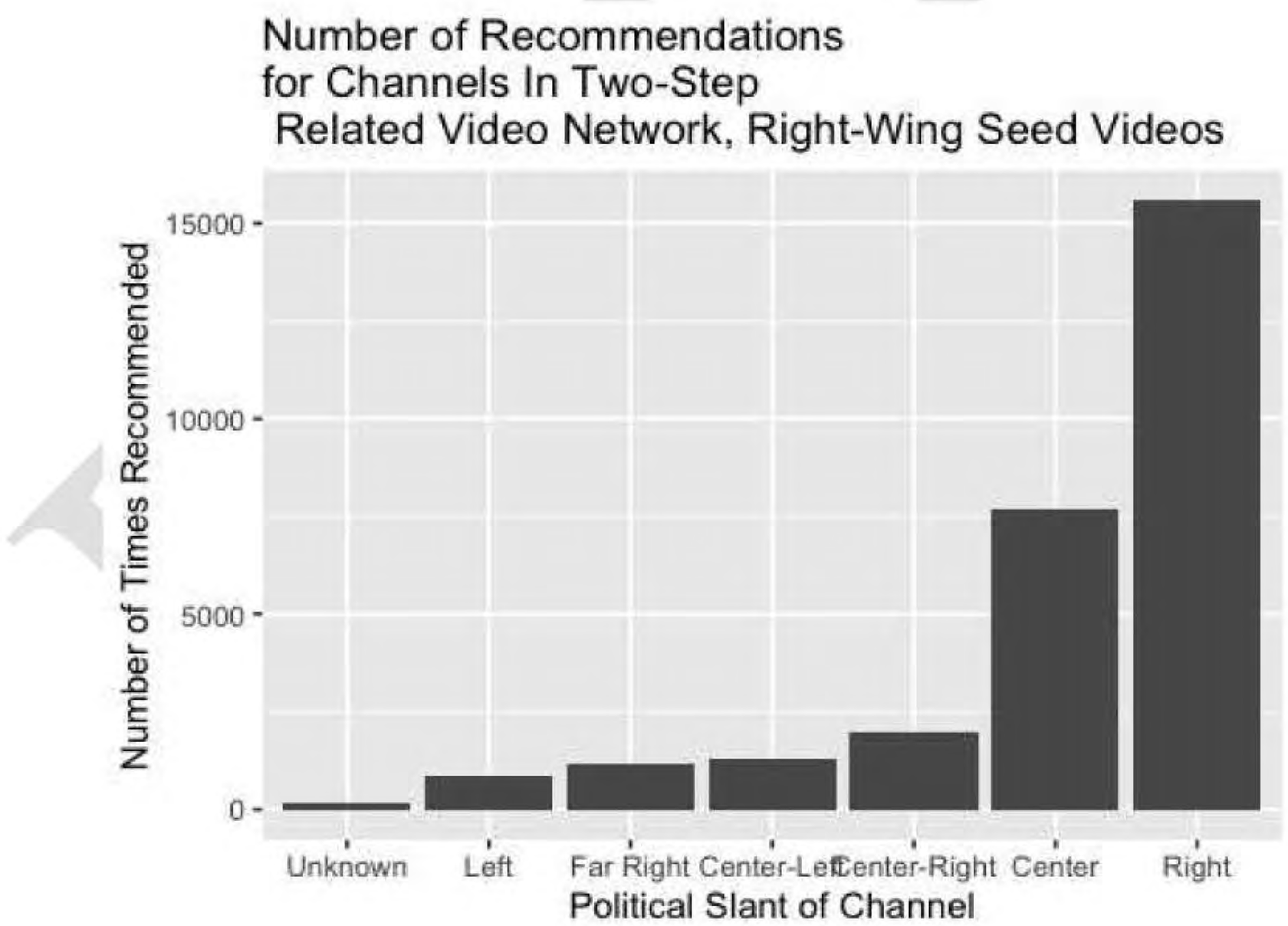
In an October 9, 2020, letter to Representative Lauren Underwood, YouTube claimed that the 2019 changes to its algorithm led to "a 70% average drop in watch time of borderline content coming from non-subscribed recommendations in the U.S."[275] In 2021, the company claimed publicly that experiments showed demoting "salacious or tabloid-type content" actually increased watch time by 0.5% because some users find it "off-putting."[276] YouTube has not

[271] GOOG-HSCI-00000386; GOOG-HSCI-00001378; GOOG-HSCI-00006804.
[272] GOOG-HSCI-00000741.
[273] YouTube trains its content moderation and recommendation algorithms with the help of panels of external, human evaluators; these evaluators may label content as "borderline" if it is not clearly violative but still potentially harmful or problematic. *See* GOOG-HSCI-00001370.
[274] *Consider* "Examining the consumption of radical content on YouTube," Homa Hosseinmardi, Amir Ghasemian, Aaron Clauset, Markus Mobius, David M. Rothschild, and Duncan J. Watts (August 2, 2021), available here: https://www.pnas.org/doi/10.1073/pnas.2101967118; "Exploring Bias and Information Bubbles in YouTube's Video Recommendation Networks," Baris Kirdemir & Nitin Agarwal (January 1, 2022), available here: https://link.springer.com/chapter/10.1007/978-3-030-93413-2_15; "Auditing radicalization pathways on YouTube," Manoel Horta Ribeiro, Raphael Ottoni, Robert West, Virgílio A. F. Almeida, and Wagner Meira (January 27, 2020), available here: https://dl.acm.org/doi/abs/10.1145/3351095.3372879; "Election Fraud, YouTube, and Public Perception of the Legitimacy of President Biden," James Bisbee, Megan A. Brown, Angela Lai, Richard Bonneau, Joshua A. Tucker, and Jonathan Nagler, forthcoming, draft provided to Select Committee.
[275] GOOG-HSCI-00001378.
[276] GOOG-HSCI-00006804.

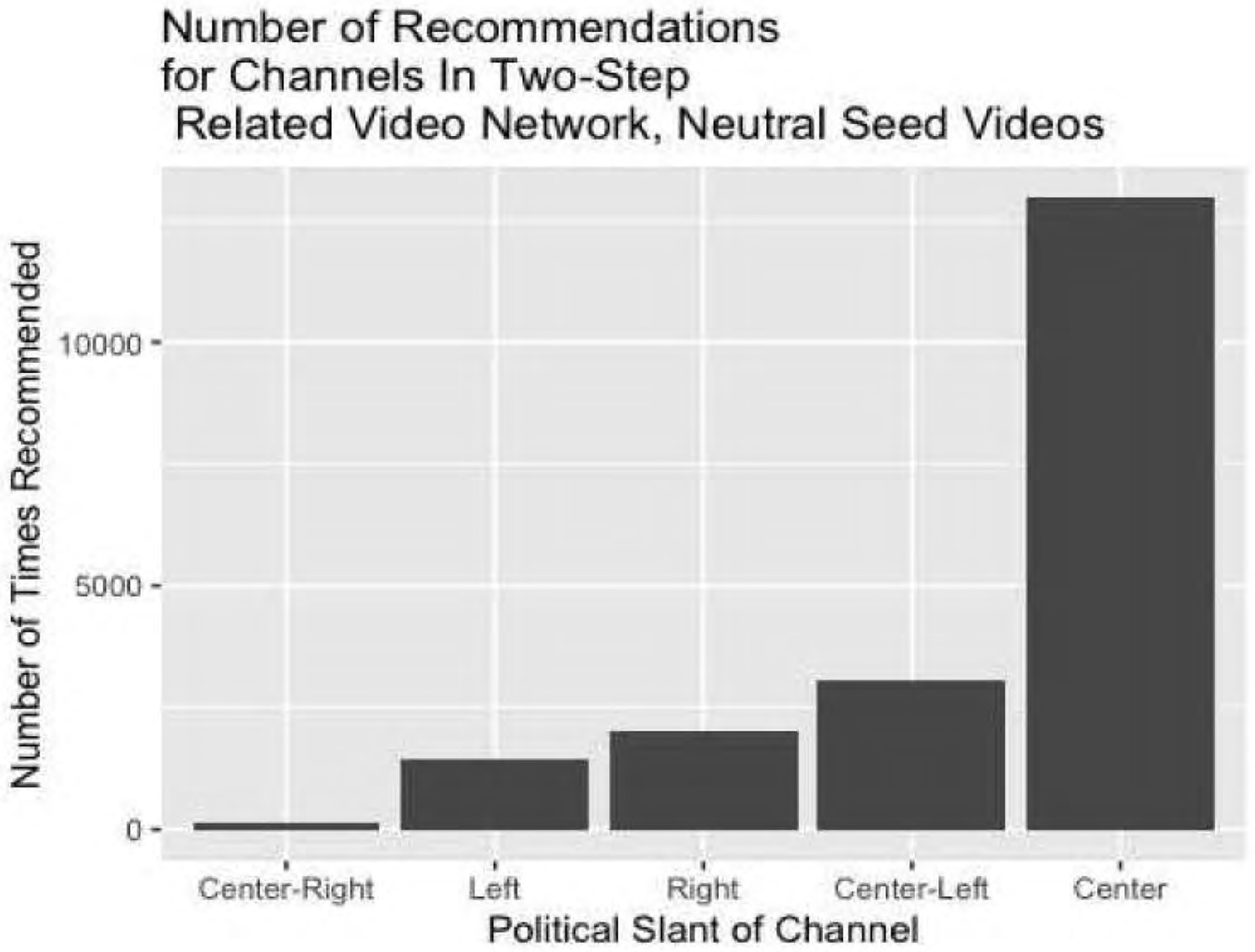
provided the Select Committee any information on how many users bypass the recommendation by subscribing to such content on purpose; the above academic studies, however, suggest that users who consume this borderline content often do so on purpose, not because they are taking algorithmic direction.

In a May 2022 briefing, Select Committee staff asked YouTube if it saw a decrease in *overall* views on borderline content, or on borderline content related to election fraud. The briefers said that YouTube does not collect this data, making it hard to assess how popular this content is overall.[277]

The Select Committee ran experiments on YouTube to test the effects of the recommendation algorithm. Select Committee staff gathered a set of videos as "seeds" to explore related videos. These seed videos were composed of five "neutral" videos and five "right-wing" videos, all of which cover allegations of election fraud within the context of the 2020 presidential election. Select Committee staff categorized them qualitatively based on assessment of the video content published by each channel. Then, they collected all related videos, and then all of the related videos for *those* videos, creating a two-step process for discovering YouTube recommendations. The results are summarized in two charts:



[277] Memo on May 16, 2022, Select Committee Briefing with Alphabet.



The clear takeaway from these charts is that right-wing videos on election fraud link out to other right-wing and far-right content on a **much more frequent basis** than neutral videos on election fraud. It's not close, either: right-wing to far-right content made up nearly 60% of the top 100 channels recommended for the right-wing seed videos and only 10% for the neutral seed videos. Perhaps even more striking, there were **no far-right channels** in the top 100 channels for the neutral seed videos, while far-right videos – i.e., more extreme than the seed videos – made up a substantial portion of recommendations for the right-wing seeds.

Regardless of these changes to the algorithm, YouTube's policy problems persist. Leading up to the 2020 U.S. presidential election, campaign ads for both President Donald Trump and Vice President Joe Biden were aired on YouTube channels for white supremacists like the Identitarian movement.[278] This enables the actors responsible for the channel to collect money from both campaigns.[279] The key takeaway here is not that extremists were able to make a profit off mainstream American presidential campaigns, but rather that international white supremacist groups were still operating openly on the platform as of July 2020. This bears a stark contrast to the banning of individuals partaking in far-right hate speech on YouTube including David Duke, Richard Spencer, and Stefan Molyneaux, which happened in the same month.[280] As long as such content is readily available, anyone can see it.

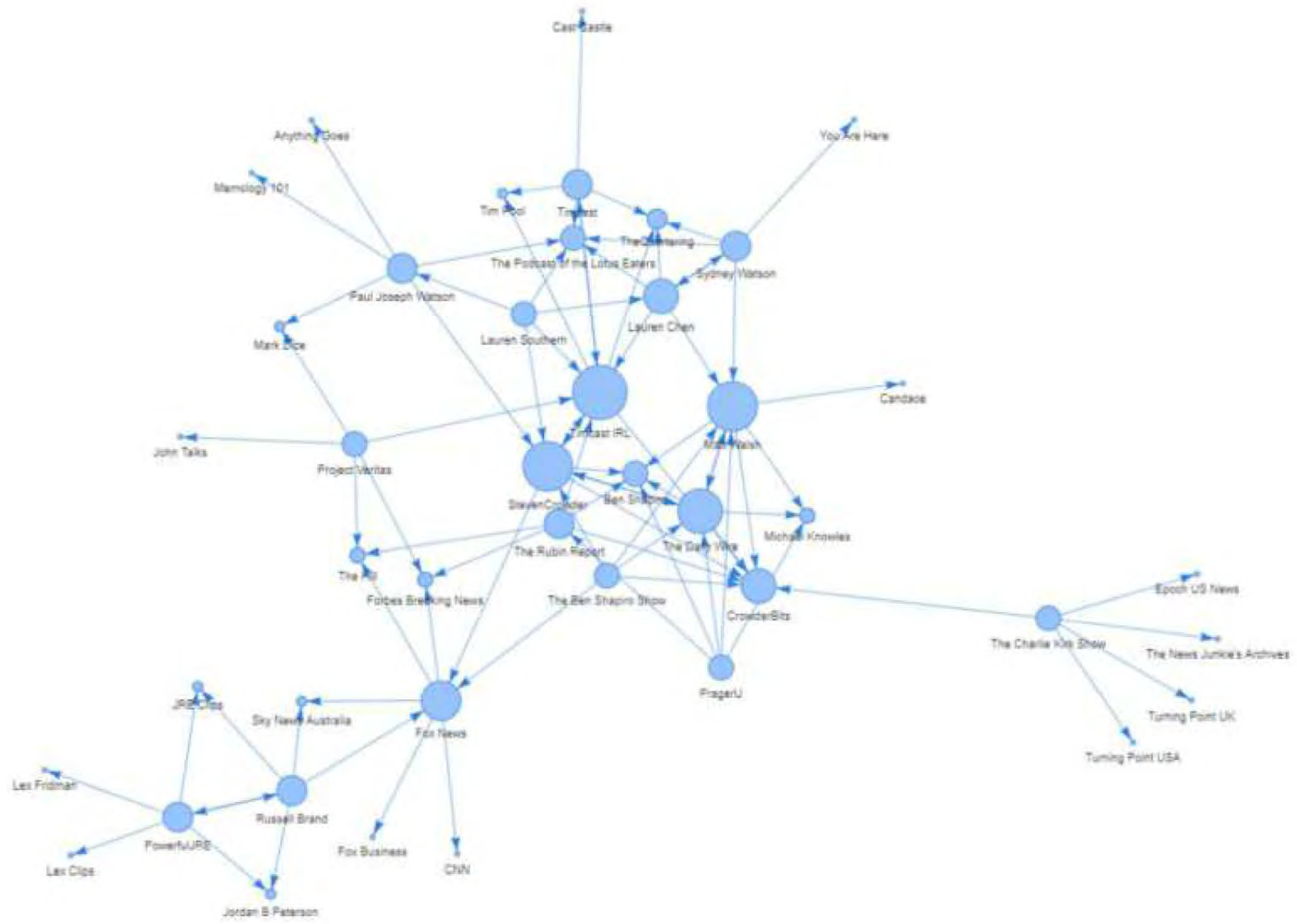
---

[278] Mark Scott, *YouTube Runs Trump, Biden Campaign Ads Alongside Videos from White Supremacists and Russian Media*, Politico (July 7, 2020), available at https://www.politico.com/news/2020/07/07/youtube-trump-biden-campaign-ads-russia-white-supremacist-350650.
[279] *Id.*
[280] Alex Hern, *YouTube Bans David Duke and Other US Far-Right Users*, The Guardian (June 30, 2020), available at https://www.theguardian.com/technology/2020/jun/30/youtube-bans-david-duke-and-other-us-far-right-users.

The Committee conducted a second analysis of YouTube users who frequently used Stop the Steal rhetoric in YouTube comments between Election Day and January 6th. These users were statistically more likely to comment on multiple right-wing channels and can spread conspiratorial narratives at disproportionate levels by serving as bridges between extreme and mainstream influencers. The network topology of YouTube commenters on 18 right-wing influencer channels shows that, between Election Day and January 6th, Stop the Steal commenters created extensive connections between channels and communities.

In the network map below, blue nodes are channels and other nodes are commenters. The blue channel nodes are sized based upon the number of Stop the Steal comments they hosted between November 3, 2020, and January 6, 2021. Analysis based on audience overlap data provided by Tubular Labs (specifically its "Audience also watched" function) demonstrates that right-wing influencers share significant portions of their audiences with each other. This overlap data also shows that seemingly benign but conspiratorial influencers like Russell Brand can act as bridges themselves for audiences to cross over from mainstream to right-wing communities.

The audience overlap data for these influencers further confirms that they have built a strong, shared community



The audience overlap data for these influencers further confirms that they have built a strong, shared community

VI. Reddit

Reddit differs from Facebook or Twitter in that, rather than rely on individual users and their connections, it cultivates communities around shared interests and pseudonymity.[281] This functionality is similar to that seen in Facebook groups, Telegram channels, and subgroups within other forums (e.g., TheDonald.win as a community within the broader ".win" sites).

Reddit did not play a noteworthy role in enabling insurrectionists to plan, collaborate, or engage in outwardly violent discussion in the lead up to the attack on the Capitol. But like other sites that allow for people with shared interests to reverberate their beliefs within the confines of a bounded community, Reddit enables the creation and cultivation of echo chambers. The most notorious of those echo chambers was found in r/The_Donald, a long-running subreddit that was used as a forum for supporters of President Trump and had been the focus of repeated reports of violent content and hate speech.[282] The subreddit had almost 800,000 subscribers at its height of popularity and, according to experts, was impactful because of its users "funneling content shared or created by its users to an audience of hundreds of thousands of Trump supporters and potentially millions of general Reddit users."

Brad Parscale tweeting about his use of r/The_Donald

On June 29, 2020, Reddit finally shut down r/The_Donald The decision to shut down r/The_Donald followed months of back-and-forth between Reddit administrators and the subreddit's community moderators, who had failed to flag content that included calls for armed activity by militias in response to COVID-19 lockdowns and potential violence against government officials.[283] Documents provided to the Select Committee demonstrate that the quarantine process dragged on for months, with subreddit moderators providing several detailed appeal months after the initial action.[284] Reddit confirmed that users on the subreddit

---

[281] Dr. Sam Bernard, "Statement for the Record: Reddit and r/The_Donald," submitted to the Select Committee on March 31, 2022.
[282] Craig Timberg and Elizabeth Dwoskin, *Reddit Closes Long-Running Forum Supporting President Trump After Years of Policy Violations*. Washington Post (Jun. 29, 2020), https://www.washingtonpost.com/technology/2020/06/29/reddit-closes-long-running-forum-supporting-president-trump-after-years-policy-violations/; *see also*.
[283] JAN6_0630; JAN6_0615 (the latter document is Reddit's initial notification to subreddit users that the community was quarantined, and the reason cited was "content that encourages or incites violence," most recently "towards police officers and public officials in Oregon.")
[284] JAN6_0617.

continued to upvote and engage with violative content even after the offending posts were removed, suggesting that the problems with the community may not be easily fixed.[285]

When a subreddit is quarantined, its functionality and ability to come across Reddit users' feeds is limited; for example, it from search results, prevent advertisements from appearing, and access to general user feeds. However, a quarantine does not prevent members from viewing or posting in the subreddit, even if users are prompted with a warning that the subreddit is quarantined.[286]

By the time that Reddit shut down the notorious subreddit—a year after its initial quarantine—most of the communities' users had already migrated to TheDonald.win, a new platform that had been created by the moderators to serve as a content moderation-free zone for the same audience. Moderators were directing users to the new site by the time Reddit finally banned the subreddit.[287] In fact, Jody Williams, a moderator of the subreddit who would go on to buy TheDonald.win's domain, confirmed that he and other moderators made posts "encouraging people to move over" to the new site.[288]

Williams said that the moderators of the subreddit actually believed that they would be banned "right away," and were surprised when they were given a protracted quarantine period in which they were able to freely promote a new, more openly extremist website.[289] However, Jessica Ashooh, Reddit's Senior Director of Policy, said that she could not recall any posts about TheDonald.win during the period in which r/The_Donald was under quarantine.[290] In fact, Ashooh emphasized the drop-off in activity on the subreddit after the quarantine was imposed, and characterized most of the posts as "inward-looking"; that is, about the subreddit itself.[291] This characterization of the final months of r/The_Donald neglects to mention the way that moderators such as Williams deliberately used their continued ability to post on the platform as a means to promote their new website, which was, according to Williams, "almost a like-for-like copy of Reddit, at least functionally speaking."[292]

Williams explained that the moderators on r/The_Donald were for the most part "working very hard" to remain on Reddit, which he said was important given that the website would have a "better reach than any alternatives."[293] He stated explicitly that they were worried a lot of users

---

[285] Memo on May 19, 2022, Select Committee Briefing with Reddit.
[286] Memo on Feb. 8, 2022, Select Committee Briefing with Reddit.
[287] Robert Peck, *The Hate-Fueled Rise of r/The_Donald—And Its Epic Takedown*. Wired (Aug. 3, 2020), https://www.wired.com/story/the-hate-fueled-rise-of-rthe-donald-and-its-epic-takedown/. In a briefing with the Select Committee, Reddit explained that although quarantining a subreddit froze much of the activity of the subreddit, moderators were still able to make and approve posts that were visible to subscribers. *See* Memo on Feb. 8, 2022, Select Committee Briefing with Reddit.
[288] Deposition of Jody Williams, 32:8.
[289] *Id.* at 32:16-25.
[290] Memo on May 19, 2022, Select Committee Briefing with Reddit.
[291] *Id.*
[292] Jody Williams Deposition, 13:23-24.
[293] *Id.* at 22:11-12, 23:13-15.

"wouldn't convert" to another site if the subreddit would shut down, which would diminish the community's capacity to influence the 2020 election.[294] **This demonstrates the power of mainstream major social media sites and the potential influence they can have by deplatforming violent, extremist communities. However, it also demonstrates the cost of a delayed reaction to those communities.**

Williams testified that the length of the quarantine "helped us a lot, of course" and said that the userbase of TheDonald.win was "a whole lot more than we expected" because of the opportunity to advertise the new site on Reddit for so long.[295] Williams explained that the moderators of r/The_Donald did not expect to have this opportunity, and agreed that the "belief" was that the subreddit would be shut down as soon as they started promoting TheDonald.win.[296] The transition between r/The_Donald and TheDonald.win a preplanned strategy, to the extent that the subreddit was frozen "at the same time" that TheDonald.win went live, in order to "encourage" users to use the new site instead.[297] **Despite this, the Committee was unable to establish that Reddit viewed the emergence of TheDonald.win as an influencing factor in deciding to shut down r/The_Donald, which may have facilitated to the growth of the community.**

The banning of the r/The_Donald coincided with the implementation of a new policy against hate speech that allowed Reddit to look more broadly at community dynamics when determining whether to take action against a subreddit.[298] Reddit indicated that this new policy was targeted at r/The_Donald, but that it ended up being the final straw for the subreddit.[299]

Reddit's actions against the subreddit, however protracted, were apparently of interest to the Trump Administration. In October 2019, Ory Rinat, former White House chief digital strategist, emailed Reddit to ask for more information about an article that alleged the "suffocation" of r/The_Donald.[300] This outreach is especially interesting given public reports that Dan Scavino repeatedly engaged with the subreddit and mined it for content to use on Trump's Twitter feed.[301] Reddit informed the Committee that this type of outreach was atypical, but that it was the only kind of contact it received from the Trump Administration on this issue and had no impact on the ultimate disposition of r/The_Donald.[302]

Reddit's delayed reaction to r/The_Donald shows the problem with its user-mediated content moderation strategy, which continued throughout 2020. Although Reddit has maintained that it

---

[294] *Id.* at 23:13-15.
[295] Jody Williams Deposition, 32:22-33:25.
[296] *Id.* at 36:14-17.
[297] *Id.* at 31:20-32:6.
[298] JAN6_0666; *see also* Memo on May 19, 2022, Select Committee Briefing with Reddit.
[299] Memo on May 19, 2022, Select Committee Briefing with Reddit.
[300] JAN6_0756.
[301] Justin Hendrix, *TheDonald.win and President Trump's Foreknowledge of the Attack on the Capitol*, Just Security (Jan. 12, 2022), available at https://www.justsecurity.org/79813/thedonald-win-and-president-trumps-foreknowledge-of-the-attack-on-the-capitol/.
[302] Memo on May 19, 2022, Select Committee Briefing with Reddit.

was not a major source of planning or calls for coordinated violence on January 6th, its content moderation efforts had several critical failings that were essentially tied to the decentralized way that the platform processed concerning content. Indeed, the platform enabled the spread of election-related disinformation; the Election Integrity Partnership (EIP) found that r/The_Donald was one of the top two misinformation-spreading sites in their entire dataset spanning 15 platforms.[303] As of August 2020, Reddit did not have a single content moderation policy related to the election.[304]

The EIP highlighted several incidents wherein Reddit allowed for the amplification and proliferation of falsehoods relating to the 2020 election. For example, conservative influencer and verified Twitter user, Elijah Schaffer, posted photos of a mail-in ballot disposal incident that allegedly took place in Sonoma, CA, on September 25, 2020. He posted the photos to Twitter, which were subsequently spread across Gab, Parler, and Reddit. It is worth noting that the photos were from 2018 and showed empty envelopes that had been lawfully thrown away, rather than ballots that were being illegally discarded.[305]

A right-wing activist posted a video to Facebook that engaged with the conspiracy theory surrounding the use of Sharpie permanent markers at voting booths, and the consequent inability for those ballots to be counted. According to the investigation by the Election Integrity Partnership, the video was later shared across YouTube, Twitter, Rumble, TikTok, Parler, and Reddit.

These incidents speak to Reddit's role in the social media ecosystem: it is a conduit of information, enabling the mass spread of content including links, images, and discussion within echo chambers of like-minded users. Indeed, though the aforementioned Sharpiegate conspiracy theory was swirling in early- to mid-November of 2020, it is still being circulated on social media – including Reddit.[306]

Following the 2020 election, Reddit produced an internal after-action report that identified key items of "constructive criticism" that encourages more centralization of reporting processes and earlier planning for known risky events. The memo further noted that there was "no central tracking of all election-related escalations" and that the community moderators were overwhelmed by the influx of content.[307]

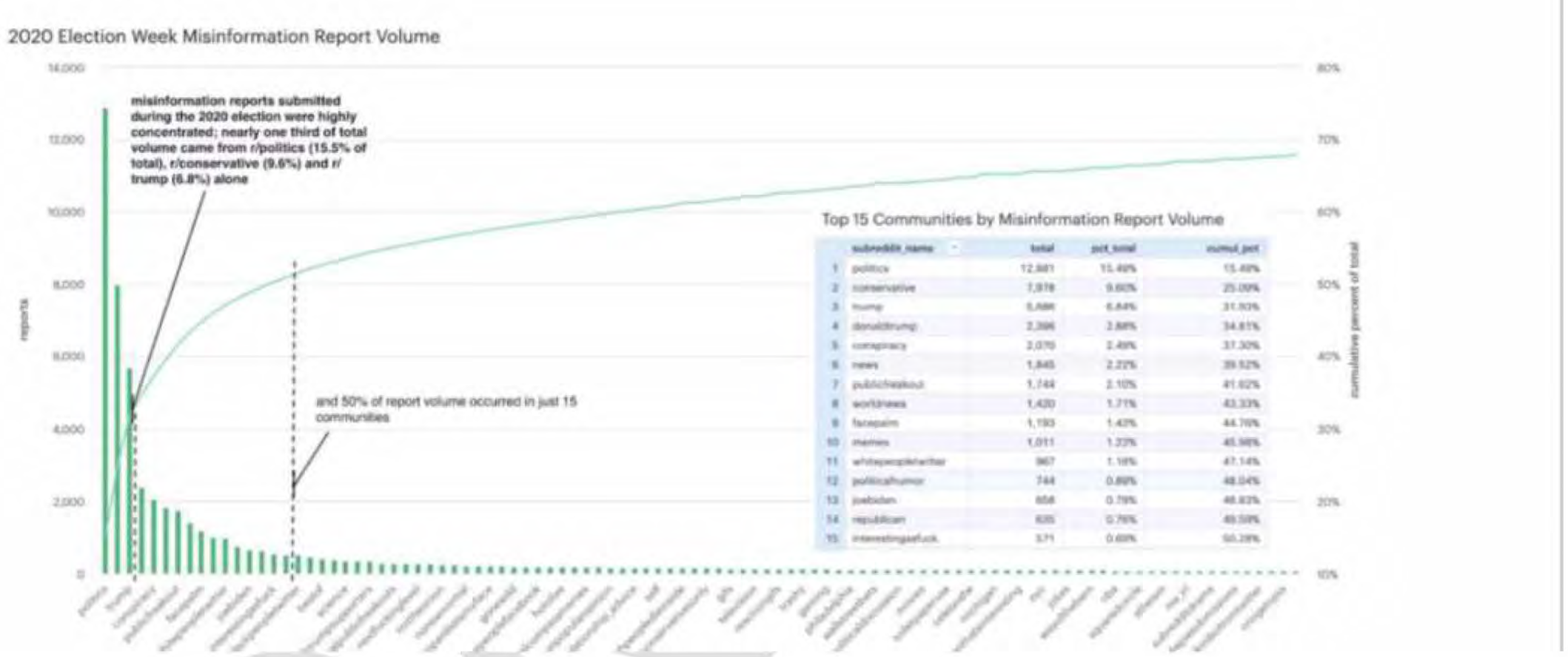
Reddit's content moderation operation was not fully prepared for violence on January 6th. Although contemporaneous documents provided by Reddit indicate that there was not a dramatic

---

[303] The platforms analyzed were Facebook, Instagram, Twitter, YouTube, Pinterest, Nextdoor, TikTok, Snapchat, Parler, Gab, Discord, WhatsApp, Telegram, Reddit, and Twitch.
[304] "The Long Fuse: Misinformation and the 2020 Election." The Election Integrity Partnership. 2021. Pg. 214.
[305] "The Long Fuse: Misinformation and the 2020 Election." The Election Integrity Partnership. 2021. Pg. 58.
[306] As of August 2022, per a review by Select Committee social media analysts.
[307] JAN6_0285.

increase in violative content on January 6th, the platform did not appear to remedy the problems it raised in its post-election after-action report. According to Reddit, employees had a constantly updated open-source dashboard throughout the post-election period to monitor related threats, which did not yield significant indications of coordinated calls to violence.[308]

Contemporaneous internal documents from Reddit support these findings,[309] but it is also clear that Reddit was not prepared for massive spikes in violent content. Internal meeting notes from the post-election period show that there were concerns that there is no ability to monitor the Reddit chat function, and that there may need to be "multi-reddit monitoring" if marches on capitols occurred. However, the agenda states they are "unlikely to need all-hands on deck."[310]



Internal Reddit analysis of election-related misinformation

The fact that Reddit was not the site of massive organizing before and during January 6th should not diminish the problematic content that *was* on the site related to the attack. Email traffic indicates that there was a need to censor personal information related to Speaker Pelosi and operations that were mobilized after the fact to control problematic content.

However, a Reddit employee wrote that these automatic operations were "the best real-time solution but impossible to stay on top of it b/c of the amount of inbound content." The same email chain shows that users were banned on January 6th because of glorification of violence,

---

[308] Memo on Feb. 8, 2022, Select Committee Briefing with Reddit.
[309] *See, e.g.*, JAN6_0279, JAN6_0324.
[310] JAN6_0294.

although this included both anti-protester and pro-protester sentiment.[311] On January 8, 2021, Reddit permanently banned r/donaldtrump for repeated violations.[312]

Reddit's content moderation practices came under strain in the face of organized far-right extremist movements. Just as Reddit experienced protracted debates with moderators of r/The_Donald, its user-mediated content moderation efforts also faced challenges later in 2020. While Reddit was flagging prominent pieces of election disinformation, it faced pushback from moderators of pro-Trump forums about whether and how fast content should be removed.[313] The overwhelming majority of subreddits that received prompts to better monitor election-related content were pro-Trump forums.[314]

Given that 97% of content is moderator-removed on Reddit, this presents a major problem with how the platform removes disinformation amidst coordinated campaigns surrounding specific events, such as the election.[315] Reddit indicated that lightening the load of user-moderators is a "constant theme" of Reddit's strategy and that, unlike other platforms under investigation by the Committee, it does not believe it is the job of user-moderators to do base-level content moderation.[316] However, the experience of r/The_Donald, where Reddit told the Committee that it was unable to find better moderators to control the policy-violating content, exposes dangers that are inherent to user-mediated content moderation.[317]

Internal analysis by Reddit showed that the vast majority of election-related mis- and disinformation occurred in just 15 subreddits, include r/conservative, r/trump, r/donaldtrump, and r/conspiracy. In addition, over 100K users were banned for behavior during the election season, belying Reddit's assertions that it was not a major vector of dangerous narratives in 2020.[318]

Overall, Reddit appears to have had a better handle on potentially violent content and disinformation on its platform than some of its peers, but the clumsiness of its content moderation operation left room for such narratives to take hold on far-right corners of the website. This reflects a broader pattern of delayed responses, that allowed r/The_Donald to stay active for long enough for its users to migrate to its more extreme successor, TheDonald.win. **According to Jody Williams, theDonald.win had "immediate" access to a userbase of "hundreds of thousands of people" who were associated with r/The_Donald, which proved to be an enormously important launching pad for the website that would become a staging ground for the attack on January 6th.**[319]

---

[311] JAN6_0279.
[312] JAN6_0324.
[313] JAN6_0285.
[314] JAN6_0568.
[315] JAN6_0351.
[316] Memo from May 19, 2022, Select Committee Briefing with Reddit.
[317] *Id.*
[318] JAN6_0285.
[319] Jody Williams Deposition, 33:1-14.

VII. TikTok

TikTok is a social media application for sharing short-form videos. The platform is owned by ByteDance, a technology company based in China. Started in 2018 as the international equivalent of a Chinese app called Douyin, TikTok grew quickly into one of the world's largest social media platforms.[320] Because of its rapid growth and the nature of the content shared on the platform, it is a space ripe for the quick, unmitigated spread of information – for better or worse. Online harms experts have described TikTok's content moderation guidelines as thorough when compared with those of other mainstream social media platforms; the issue is, according to those experts, shoddy enforcement of the existing policies.[321] Indeed, experts have criticized the platform for overly censoring people of color, members of the LGBTQ+ community, and – according to TikTok's internal documents – the "ugly" and "poor," while disinformation and extremist content are allowed to remain.[322]

Internal documents provided to the Select Committee demonstrated that TikTok was working quickly to establish content moderation policies in advance of the US 2020 election.[323]As it was standing up these policies, TikTok wrestled with unique political scrutiny; its origins in China had attracted suspicions in national security policy circles and outright hostility from President Trump, who tried to ban the platform from operating in the United States in the fall of 2020.[324]

Overall, TikTok does not appear to have been a major source of news and information for the perpetrators of the January 6th attack. However, the platform's growing popularity has increasingly made it a place for political content in the months since January 6th. TikTok continues to attract the mix of hyper-partisan commentators, conspiracy theorists, and extremists active on other platforms. For these reasons, its approach to civic integrity is relevant to understanding and preventing the risk of political violence in future elections.[325]

TikTok's approach to civic integrity was very much under construction in mid-2020 as it rushed to prepare for the election. TikTok policy documents from June 2020 reveal that the platform did not "have policy coverage to sufficiently address the broad scope" of mis- and disinformation. Over the next month, TikTok worked to establish three new policies on synthetic media (sometimes called "deepfakes"), manipulated media (deceptively edited media sometimes

---

[320] "TikTok: The story of a social media giant," *BBC* (August 5th, 2020); "TikTok to rank as the third largest social network, 2022 forecast notes," *TechCrunch* (December 20, 2021).
[321] Dr. Gabriel Weimann and Natalie Masri, "Hate on TikTok." GNET. July 7, 2020, https://gnet-research.org/2020/07/07/hate-on-tiktok/; Abbie Richards, Interview with the Select Committee, February 1, 2022.
[322] Abbie Richards, Interview with the Select Committee, February 1, 2022; Alex Hern, "TikTok 'tried to filter out videos from ugly, poor or disabled users'," *The Guardian*, March 17, 2020, https://www.theguardian.com/technology/2020/mar/17/tiktok-tried-to-filter-out-videos-from-ugly-poor-or-disabled-users.
[323] TT16SC_0001398.
[324] "Trump issues orders banning TikTok and WeChat from operating in 45 days if they are not sold by Chinese parent companies," CNN (April 28, 2022).
[325] "On TikTok, Election Misinformation Thrives Ahead of Midterms," *New York Times* (August 14, 2022).

referred to as "cheapfakes"), and coordinated inauthentic behavior (a term used by several other platforms as well as experts to describe groups of actors who use fake assets—such as accounts and pages, among others—to deceive users). TikTok also passed newly approved policies on medical and dangerous misinformation as well as false news and election misinformation to its, at the time, recently created Integrity Operations Team.[326]

The number of staff dedicated to this issue was small but growing. The size, responsibilities, and names of relevant teams were still being defined just months before the 2020 election.[327] An Anti-Abuse team responsible for implementing these policies had only one member when the policies were written.[328] As for the Integrity Operations Team, which was also responsible in part for these policies, one staffer was unaware TikTok had such a team. (A colleague had to explain that it referred to a group of content moderators whose function was being expanded.)[329]

Like other large platforms, TikTok also wrestled with how to moderate misleading content without attracting accusations of censorship, especially when the mis- and disinformation benefitted the political right. On July 28, 2020, one TikTok staffer modified the descriptions of a policy proposal on mis- and disinformation because "otherwise it may pick up much of Fox News."[330] (In a briefing with TikTok's Head of Safety, they told Select Committee staff that it is normal for Trust & Safety teams to discuss edge cases.)[331]

TikTok created a 24/7 "War Room" to prepare for the election. This center included fifty people from across different policy and functional teams, including those dedicated violent extremism, hate speech, investigations, and fact-checking, among others. The War Room was meant to handle escalated policy questions and identify threats and trends. On January 6th, the War Room monitored for extremist organization activity.[332]

In a briefing with Select Committee staff, TikTok representatives did not recall receiving warnings from law enforcement about violence before January 6th. Stronger signals came from third party threat detection teams that monitored the rest of social media for trends that might migrate to TikTok, such as Stop the Steal and SharpieGate.[333]

TikTok was slow to designate violent extremist organizations on the platform despite receiving intelligence briefings on the subject. TikTok's userbase skews younger than those of Facebook or Twitter. Its primary demographic is also younger than most of those arrested for attacking the

---

[326] TT16SC_0001408.
[327] TT16SC_0001501 – 1503.
[328] TT16_SC0001420.
[329] TT16SC_001474.
[330] TT16SC_0001408.
[331] Memo on May 24, 2022, Select Committee Briefing with TikTok.
[332] For more on TikTok's January 6th operations, *see* TT16SC_00000717 and TT16SC_00000609.
[333] Memo on May 24, 2022, Select Committee Briefing with TikTok.

US Capitol on January 6th, almost two thirds of whom were older than 35.[334] This does not mean that extremism had no home on TikTok in the lead up to the attack on the Capitol; on the contrary, militias like the Three Percenters maintained a presence on the platform.[335] TikTok was slow to recognize some of these groups: documents provided to the Select Committee suggest that the "Boogaloo" movement, for example, was not designated a violent extremist organization until April 2021, much later than when Facebook made the same decision.[336]

However, with regard to other groups and movements, TikTok's approach to trust & safety overlapped significantly with peer platforms. TikTok uses proactive keyword sweeps and machine learning tools for identifying iconography to detect violative content and accounts associated with banned organizations and movements. Proactively detected videos and accounts can be sent to human moderators for review; "hunt teams" can also search for networks of violating accounts or other threats that the Safety Team should be aware of.

Like its peers, TikTok also uses soft actions to demote content. In September of 2020, the DNC sent a message to TikTok raising concerns about QAnon content on the platform. TikTok's response emphasized that this content was in violation of its community guidelines and that company policy was to "remove content and ban accounts."[337] But like Twitter, Facebook, and YouTube, TikTok sometimes relies on soft actions to "grayzone" borderline content; if a video includes a certain number of flagged terms, TikTok automatically reduces its distribution while the Safety Team assesses whether it is violative content or critical commentary. TikTok will also prevent such videos from appearing on the "for you" page, where they are algorithmically recommended to users. They can also be banned from search.[338]

TikTok estimates this strategy was effective. Regarding QAnon content, the company believes that 90% of videos in the "gray zone" received fewer than a thousand views, while only 1% received over ten thousand.[339] That said, as with YouTube, even if content is not recommended and/or barred from search, some users inevitably find it by following specific individual accounts or through links distributed on other platforms. Therefore, the borderline content strategy is flawed; some of the videos which broke through the "gray zone" received millions of views.

---

[334] "AMERICAN FACE OF INSURRECTION: Analysis of Individuals Charged for Storming the US Capitol on January 6th, 2021," Chicago Project on Security & Threats (January 5, 2022). https://d3qi0qp55mx5f5.cloudfront.net/cpost/i/docs/Pape_-_American_Face_of_Insurrection_(2022-01-05)_1.pdf?mtime=1641481428.

[335] "Far-right militias are using TikTok to organize and recruit new followers," Media Matters for America (January 12, 20221). Available at: https://www.mediamatters.org/january-6-insurrection/far-right-militias-are-using-tiktok-organize-and-recruit-new-followers.

[336] TT16SC_0001615.

[337] TT16SC_0000277.

[338] *Id.*

[339] TT16SC_0001749.

Broad demotions like this can be effective, but they also cause collateral damage and are not foolproof.[340] Like Facebook, the company worried about "overkill" in this strategy, noting that some of the top videos from gray-zoned accounts were not violative.[341]

TikTok also has a strike system for account violations: strikes are weighted based on the policy violation. The company also looks closely at "sole purpose" accounts— for example, if an account mostly dedicated to sports posts a few QAnon video, that video will be removed but the account might not be demoted, watch-listed, or banned. If the level of violative content increases, the account may eventually be removed.

On other key issues, TikTok's policies were *ahead* of its peers. In contrast with Facebook, TikTok was able to respond to Stop the Steal quickly because its election misinformation policy already prohibited delegitimization.[342] TikTok staff told the Select Committee that this policy applies at all times globally and is one of the most severely prohibited forms of harmful misinformation. Recent reporting provides an example of how determined users can circumvent this policy: while #StoptheSteal is banned, users who search for alternatives like #stopthestealll can still find relevant content.[343]

According to an audit of 29 tech platforms conducted by Tech Against Terrorism for the Select Committee, "TikTok clarified on January 7, 2021, that Trump's speeches, where he reiterated claims of a fraudulent election, were being removed on the grounds that they violate the company's misinformation policy."[344] This action was not taken widely across other platforms.

Also unlike Facebook, TikTok's Safety team sits under different leadership than its public policy, communications, and legal team. This structure may help separate out conflicting incentives between content moderation and government relations.[345]

TikTok's policy on implicit violence seems more assertive than Twitter's: it removes implicit threats of violence from the platform, such as videos in which users pantomime cocking a gun.[346]

Like other platforms, TikTok waited until it was too late to act decisively and enforce policies against videos featuring President Trump. According to an audit of 29 tech platforms conducted by Tech Against Terrorism for the Select Committee, "TikTok clarified on January 7, 2021, that

---

[340] *Id.*
[341] TT16SC_0001749.
[342] For more on the development of the misinformation policy, see TT16SC_0001398.
[343] "On TikTok, Election Misinformation Thrives Ahead of Midterms," *New York Times* (August 14, 2022).
[344] Tech Against Terrorism, "Examining Tech Platforms' Moderation Actions and Policies Related to the Attack on January 6 and Their Impact on Online Extremist Behaviour," submitted to the Select Committee on March 30, 2022, in response to a request for an expert witness statement for the record.
[345] Memo on May 24, 2022, Select Committee Briefing with TikTok.
[346] *Id.*

Trump's speeches, where he reiterated claims of a fraudulent election, were being removed on the grounds that they violate the company's misinformation policy."[347]

TikTok acknowledges that its recommendation system can create harmful rabbit holes of radical content.[348] It is worth noting that TikTok's recommendation system is different than that of YouTube's or Facebook's, which are largely based on who a user is already following and whose content a user is actively engaging with; YouTube, in particular, emphasizes connections over context. While external studies examining TikTok's recommendation algorithm have already been undertaken and provide useful insights, the Select Committee designed its own experiment. The Committee's TikTok study revealed that even with minimal engagement, the platform's recommendation algorithm can still steer users down "rabbit holes."

For example, one of the profiles we created – a fictional 41-year-old female from Acton, Massachusetts – quickly accelerated from random content with no apparent theme to overtly right-wing content. After 15 minutes of scrolling – with no engagement beyond simply watching videos, meaning no liking, downloading, or sharing of the videos – the algorithm began showing our user TikToks centered around the Johnny Depp v. Amber Heard court case. 100% of the TikToks shown to "Alice" about this court case were pro-Depp and/or anti-Heard. It took 55 minutes of scrolling with no engagement to reach pro-Elon Musk TikTok content.

From there, the content veered away from Depp v. Heard and instead featured a mix of the following: Joe Rogan podcast clips; pro-US military content; conspiracy-centric content, featuring conspiracies about Jeffrey Epstein, President Vladimir Putin being ill, humans collectively living in a simulation, Freddie Mercury's death, the Titanic's sinking, 9/11, and the Freemasons, among others; religious content, specifically that which aligned with Christianity; and additional pro-Elon Musk content.

It took 1 hour and 8 minutes for "Alice" to be served pro-Trump content after absolutely no engagement with *any* specific TikTok videos. Within seven minutes of reaching pro-Trump TikToks, "Alice" reached esoteric (i.e., "third eye") content. The next video was one featuring Joseph Goebbels, the chief propagandist for the Nazi Party and Third Reich.

Immediately thereafter, our fictional user was shown a TikTok from @thedailywire, featuring right-wing influencer, Ben Shapiro. While casually scrolling for another 1 hour and 42 minutes, our fictional user was increasingly fed content that fell clearly into the categories of anti-feminism, anti-LGBTQ+, spiritual awakening and the metaphysical, various conspiracy theories, Christianity, and Nazism. Additional content featured Joe Rogan, Elon Musk, Donald Trump, Ron DeSantis, all demonstrated in a neutral or positive way. This is just one of the Committee's

---

[347] Tech Against Terrorism, "Examining Tech Platforms' Moderation Actions and Policies Related to the Attack on January 6 and Their Impact on Online Extremist Behaviour," submitted to the Select Committee on March 30, 2022, in response to a request for an expert witness statement for the record.
[348] TT16SC_00000791; TT16SC_00001749.

experiments that further evidenced the power of TikTok's recommendation algorithm in creating rabbit holes toward potentially harmful content.

In a briefing with Select Committee staff, company representatives said that their Safety Team works with the Product Team on projects related to filter bubbles. One technique for addressing this is "dispersion," or efforts to reduce density in the number of users who are referred primarily one type of content. This can be done for conspiracy theories and political extremism but also for non-political harms like promotion of eating disorders.[349]

TikTok's process for review of high-profile accounts applies to a small number of users, mostly in the realm of culture rather than politics. Like Twitter and Facebook, TikTok uses a second layer of review for high-profile individuals. According to company representatives, only four or five people can add to this list and the company is working to "unify" its approach globally over time. As with Facebook's "Crosscheck" program, TikTok says the purpose of this list is to avoid high-profile false positives. In the United States, TikTok representatives told Select Committee staff that about two hundred people are on this list and many of them are celebrities, rather than political figures. Neither presidential candidate in 2020 maintained a presence on TikTok.[350]

[349] Memo on May 24, 2022, Select Committee Briefing with TikTok.
[350] *Id.*

VIII. Key Alt-Tech and Fringe Platforms

Beyond the larger, more established tech companies, smaller social media platforms also played a key role in enabling the organization of far-right groups in advance of January 6th. In some instances, these platforms played a key role in allowing actors to orchestrate the attack or send out coordinated calls for violence. According to the expert witness statement provided by Tech Against Terrorism to the Select Committee[351]:

> "During the two months between the November 3, 2020 election and January 6, 2021, the 'Stop the Steal' movement gathered momentum and support from a range of extremist groups online including the Proud Boys, the Three Percenters, the Boogaloo Bois, as well as other radical pro-Trump activists and adherents to the QAnon conspiracy theory.[352] During these months, these groups and individuals were active on a range of online platforms to discuss their views on the elections, plan offline events, and incite each other. Based on our monitoring and third-party press reporting, this took place on a wide range of platforms including Facebook and Twitter, as well as more fringe 'alt-tech' spaces like Telegram, Parler, Gab, and MeWe.[353]"

Several of these sites are relatively recently founded "alt-tech" platforms designed to mirror the affordances of existing mainstream platforms. Parler, for instance, mirrors Twitter; Gab mirrors Facebook; and the "dot win" communities mirror Reddit (where TheDonald.win originated as the "r/the_donald" subreddit). "Alt-tech" sites such as Parler and Gab expanded in the run-up to the 2020 election as more mainstream platforms began to take some actions against violent content and users who were spreading disinformation.[354] The lack of formal, proactive moderation policies makes it difficult for these platforms to detect and deter violent content, making them staging grounds for the attack.[355]

These sites in particular were borne from users' desires to spread disinformation, encourage violence, and livestream and/or offer words of support for terrorist attacks without content moderation, or, in their words, "censorship." One expert testified to the Select Committee that these platforms were beset by "problems that had been present on the mainstream platforms" but

---

[351] Tech Against Terrorism, "Examining Tech Platforms' Moderation Actions and Policies Related to the Attack on January 6 and Their Impact on Online Extremist Behaviour," submitted to the Select Committee on March 30, 2022, in response to a request for an expert witness statement for the record.

[352] Holt, Jared, *After the Insurrection: How Domestic Extremists Adapted and Evolved After the January 6 US Capitol attack,* Atlantic Council (Jan. 4, 2022), available at https://www.atlanticcouncil.org/in-depth-research-reports/report/after-the-insurrection-how-domestic-extremists-adapted-and-evolved-after-the-january-6-us-capitol-attack/.

[353] Criezis Meili, Galloway Brad (2021), From MAGA to the Fringe: What was Happening Online Before the 6 January Insurrection and What Can We Do Now? Global Network on Extremism & Technology.

[354] Expert Statement of Candace Rondeaux.

[355] As described in the paragraphs below, Parler had *some* formal content moderation policies in place, though they pale in comparison to mainstream platforms. Other alt- or fringe platforms were founded in direct hostility to the concept of content moderation and had few to no policies in this area.

which were "exacerbated by an even greater lack of control and responsibility… and the coalescing of separate anti-government groups into single online communities, creating intellectually coherent and self-supporting narratives completely detached from reality."[356]

Fringe platforms are those that are relatively unique in their affordances. In particular, these platforms have been used to post terrorist manifestos, livestreams of terrorist attacks, and terrorist propaganda, including manuals that provide guidance on how to carry out attacks. Some experts refer to these platforms as *niche*, wherein "users readily encounter explicit extreme content."[357] This category includes forums like 8kun and 4chan, and messaging platforms like Telegram. Though the aforementioned forums were founded for fairly benign reasons, they have metastasized into hotbeds for antisemitism, racism, misogyny, violent rhetoric, hate speech, and even child sexual abuse material. The QAnon conspiracy movement originated on 4chan and then spread to 8kun before making its way to mainstream platforms like Facebook, YouTube, and Twitter, among others. It is no surprise, then, that each of these platforms played a role in creating the conditions – and even enabling the planning and coordination – for the attack on the US Capitol on January 6th.

For example, Parler's lack of content moderation resulted in a flood of violent content that caused employees to reach out to federal law enforcement. Parler's promise as a moderation-free alternative platform made it a breeding ground for extremism and violent content in the run up to January 6th. Although Parler was founded as a "free speech alternative" platform in 2018, it saw a massive increase in its userbase during 2020 and spikes in activity that coincided with Black Lives Matter protests and Stop the Steal events.[358] By the time of the election, prominent conservatives and far-right extremist groups, such as the Oath Keepers, had both migrated to Parler, making the platform a centerpiece of the campaign to undermine the legitimacy of President Trump's loss.[359]

Emails obtained by the Committee demonstrate the depth of extremism on Parler leading up to the attack, to the extent that Parler employees were concerned about the possibility for violence on the day of January 6th. The Select Committee's investigation found that while Parler was more tolerant of extremist content than mainstream platforms, the shocking increase in violent rhetoric and explicit preparations between users to commit acts of violence concerned the platform enough that it reached out to law enforcement with concerns.[360]

---

[356] Kingdon, A. & Fuller, C. "The Rise of Alt-Tech and the Role of Gab in the January 6th Insurrection." Submitted to the Select Committee on April 7, 2022.
[357] Williams, H.A. & Evans, A.T. "Extremist Use of Online Spaces." RAND. Submitted to the Select Committee on April 25, 2022.
[358] *Id.*
[359] *Parler: Where the Mainstream Mingles with the Extreme* (Nov. 12, 2020), available at https://www.adl.org/blog/parler-where-the-mainstream-mingles-with-the-extreme; Mike Isaac and Kellen Browning, *Fact-Checked on Facebook and Twitter, Conservatives Switch Their Apps*, New York Times (Nov. 11, 2020), available at https://www.nytimes.com/2020/11/11/technology/parler-rumble-newsmax.html.
[360] *See,* e.g., CTRL0000007435.

The implications of Parler's minimalist content moderation policies are clear from the nature of the threats that it communicated to federal law enforcement prior to the attack. While communications between law enforcement and Reddit—another platform that utilizes a user-mediated content moderation system—focused on discrete instances of voter intimidation or election-related misinformation, the threats that Parler elevated to the FBI were alarmingly violent and specific, in some cases advocating for civil war.[361] At one point, a Parler employee told the FBI in a January 2, 2021, email, that they were "concerned about Wednesday."[362]


From: [redacted]
Sent: 12/24/2020 12:03:28 PM
To: [redacted]@fbi.gov]

https://parler.com/profile/[redacted]/posts

[redacted]

We need to mass an armed force of American Patriots 150,000 on the Virginia side of the Potomac river prepared to react to the congressional events of January 6th

And follow what the Declaration of Independence has dictated to do in situations of absolute despotism..

That whenever any Form of Government becomes destructive of these ends, it is the Right of the People to alter or to abolish it,

Prudence, indeed, will dictate that Governments long established should not be changed for light and transient causes; and accordingly all experience hath shewn, that mankind are more disposed to suffer, while evils are sufferable, than to right themselves by abolishing the forms to which they are accustomed.

But when a long train of abuses and usurpations, pursuing invariably the same Object evinces a design to reduce them under absolute Despotism, it is their RIGHT, it is their DUTY, to throw off such Government, and to provide new Guards for their future security.

read less

0 0 0

One of the posts that Parler reported to law enforcement prior to January 6th.

These calls for violence did not abate after the attack itself. Parler also saw calls for an armed invasion of D.C. *following* January 6th, with calls to "shoot your way" into the city if it was blockaded on Inauguration Day.[363] It also included calls for mass targeting of civilians in the weeks prior to January 6th.[364]

Nevertheless, throughout this time, Parler appeared to rely largely on a content moderation structure that was reliant on internal reports from users and a "Community Jury" to make enforce decisions on content.[365] In general, it sought to emulate a content moderation approach that was "viewpoint neutral" and was allegedly modeled on the First Amendment.[366]

To that end, it is not clear what kind of proactive monitoring of content was present prior to January 6th or put in place following the attack. Former CEO John Matze claims that he was

---

[361] *See, e.g.*, CTRL0000007432, CTRL0000007439, CTRL0000007439.
[362] CTRL0000007435.
[363] CTRL0000007444.
[364] CTRL0000007811.
[365] CTRL0000007811; CTRL0000007907 (Terms of Service mentioning the Community Jury).
[366] CTRL0000007907.

pushed out of his position by the Board of Directors because he was advocating for stronger content moderation measures to remove content from groups such as QAnon, which he felt was necessary to prevent a recurrence of the attack on the Capitol.[367] In his deposition with the Committee, Matze refused to answer any substantive questions, including about whether Parler saw an increase in coordinated calls to violence by extremists in the weeks leading up to the Capitol attack.[368]

The Committee was unable to procure a voluntary interview with Amy Peikoff, Parler's Chief Policy Officer, who said that in the weeks after the election that Parler "realized we need[ed] to do more" to prevent violent content on the platform.[369] However, Peikoff was also in charge of Parler's policy operation in November 2020 when an update clarified that "fighting words"—not protected under the First Amendment—would *not* be considered a violation of the guidelines *ipso facto*.[370] This undercuts Parler's claim that it was simply modeling its moderation strategies off First Amendment guarantees; just as extreme content was reaching a zenith with the 2020 election, Peikoff and her team dialed back Parler's ability to remove potentially inflammatory posts on the platform, demonstrating a commitment to the far-right spaces described in the following section.

Though the original version of Parler that was attributed with hosting massive volumes of incendiary rhetoric and election disinformation before the January 6th attack – referred to as Parler 1.0 – was removed only four days after the fact, much of the data has been captured and archived. It is all publicly available. A team of experts at the New America Foundation leveraged that dataset to produce the most fulsome analysis of Parler's role in the Capitol attack to date.

New America narrowed the dataset to 102 high-profile Parler influencers, comprised of individuals like X, Y, and Z. The team found that these influencers often shared links to content hosted by "repeat offenders," or those who frequently spread mis- and disinformation on other platforms, such as Facebook. These "repeat offender" sites included:

- The Gateway Pundit, which comprised 12.5% of all links shared by Parler influencers
- Fox News, which comprised 3.5% of all links shared by Parler influencers
- Breitbart, which comprised 2.9% of all links shared by Parler influencers

Other noteworthy sites that were shared:

- YouTube – 8.4% of links shared by Parler influencers

---

[367] Bobby Allyn, *Parler CEO is Fired After "Constant Resistance" Inside the Conversative-Friendly Site*, National Public Radio (Feb. 3, 2021), available at https://www.npr.org/2021/02/03/963832594/parler-ceo-is-fired-after-constant-resistance-inside-the-conservative-friendly-s.
[368] John Matze Deposition, 18:7-9.
[369] National Public Radio, *Parler Insists It Would Not Knowingly Tolerate Criminal Activity on Its Site* (Jan. 15, 2021), available at https://www.npr.org/2021/01/15/957141101/parler-insists-it-would-not-knowingly-tolerate-criminal-activity-on-its-site.
[370] Simon Weisenthal Center, *Parler: An Unbiased Social Platform?* (Nov. 2020), available at https://www.wiesenthal.com/assets/pdf/parler_report_final-2020.pdf.

- Twitter – 7.5% of links shared by Parler influencers
- Rumble – 3.3% of links shared by Parler influencers

These findings demonstrate the significance of the broader media ecosystem in spreading narratives in a sustained manner; users are incredibly like to be barraged with mis- and disinformation regardless of what platforms they use. They also revealed that **Parler was a prominent locus of influence for allies of President Trump.** Several Republican elected officials and media figures served as influencers on Parler. All 13 objectors to the certification of the election had Parler accounts:

- Devin Nunes (CA-22) – 5,100,000 followers, joined February 4, 2020
- Ted Cruz (TX) – 4,900,000 followers, joined June 3, 2020
- Matt Gaetz (FL-01) – 914,000 followers, joined June 24, 2020
- Jim Jordan (OH-04) – 816,000 followers, joined June 25, 2020
- Andy Bigs (AZ-05) – 342,000 followers, joined May 28, 2020
- Paul Gosar (AZ-04) – 264,000 followers, joined May 7, 2020
- Marjorie Taylor Greene (GA-14) – 228,000 followers, joined June 28, 2020
- Louie Gohmert (TX-01) – 110,000 followers, joined May 27, 2020
- Debbie Lesko (AZ-08) – 47,000 followers, joined June 25, 2020
- Jodey Arrington (TX-19) – 33,000 followers, joined May 28, 2020
- Jim Banks (IN-03) – 12,000 followers, joined June 24, 2020
- Rick Crawford (AR-01) – 8,800 followers, joined June 10, 2019
- Ron Estes (KS-04) – 160 followers, joined November 18, 2020

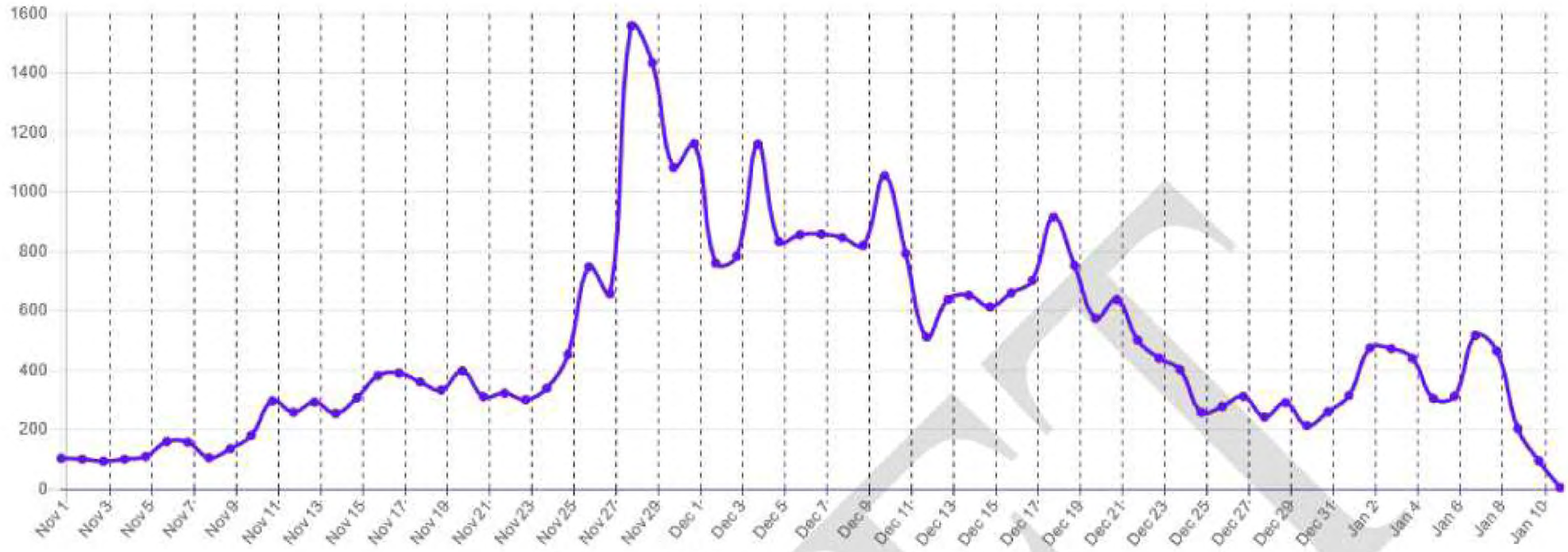
Some of the content posted by the above objectors reached millions of impressions per day. According to an analysis of the content posted by the objectors, they chose to use Twitter as a means to thank groups and media and engage with policy matters such as taxes and immigration. Parler, however, was used by the objectors to post about election integrity, COVID-19 restrictions, and their fixation on Joe and Hunter Biden.[371] This demonstrates the perceived differences between Twitter audiences and Parler audiences, and how influencers felt empowered to operate on each respective platform.

These Members of Congress maintained a presence on a platform where insurrectionary and conspiratorial language was much more common than on mainstream platforms. Select Committee Analysis found calls for the use of violence against politicians, references to

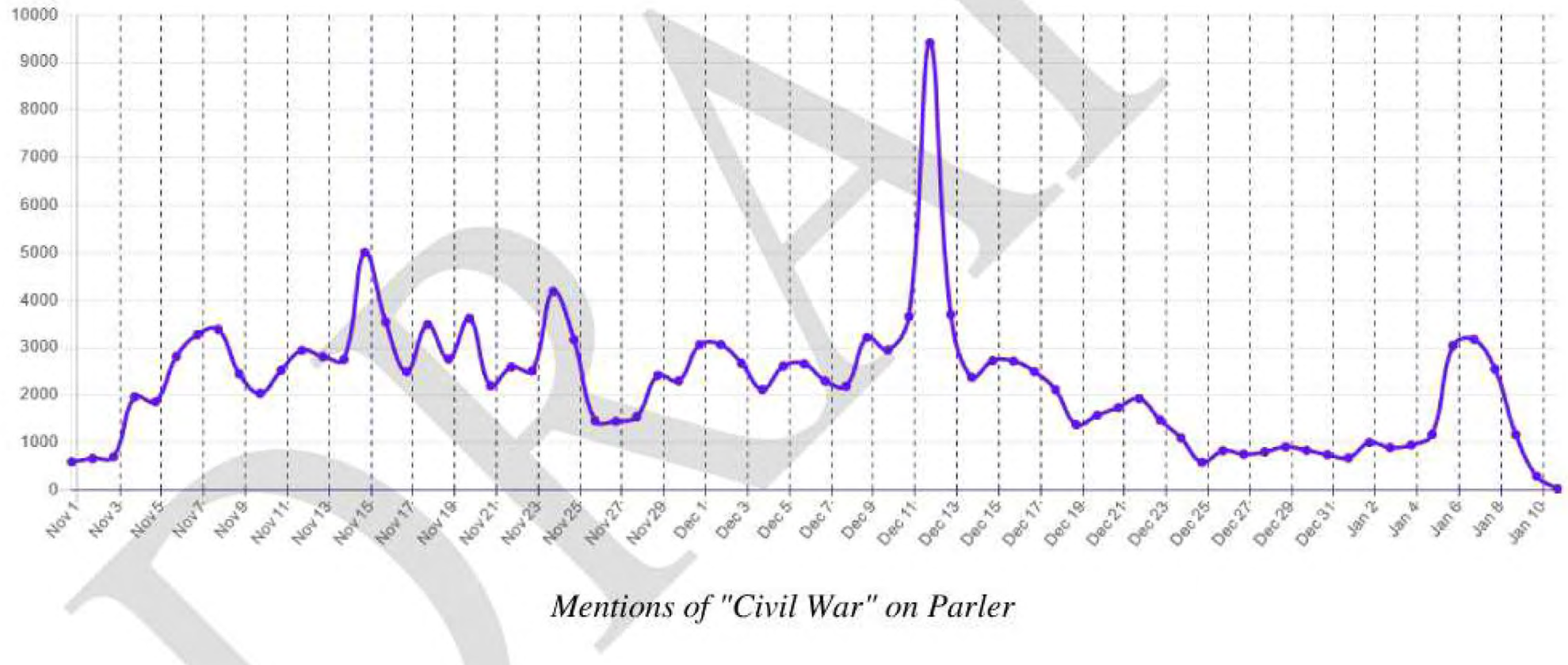
---

[371] Candace Rondeaux, Michael Simeone, Tom Taylor, Dave Troy, Shawn Walker, Ben Dalton, and Cuong Nguyen. "Parler and the Road to the Capitol Attack: Investigating Alt-Tech Ties to January 6." New America Foundation. Last updated January 5, 2022.

conspiracy theories like the Insurrection Act, and even the use of neo-Nazi terms like "Day of the Rope" all increased significantly on Parler after Election Day.[372]

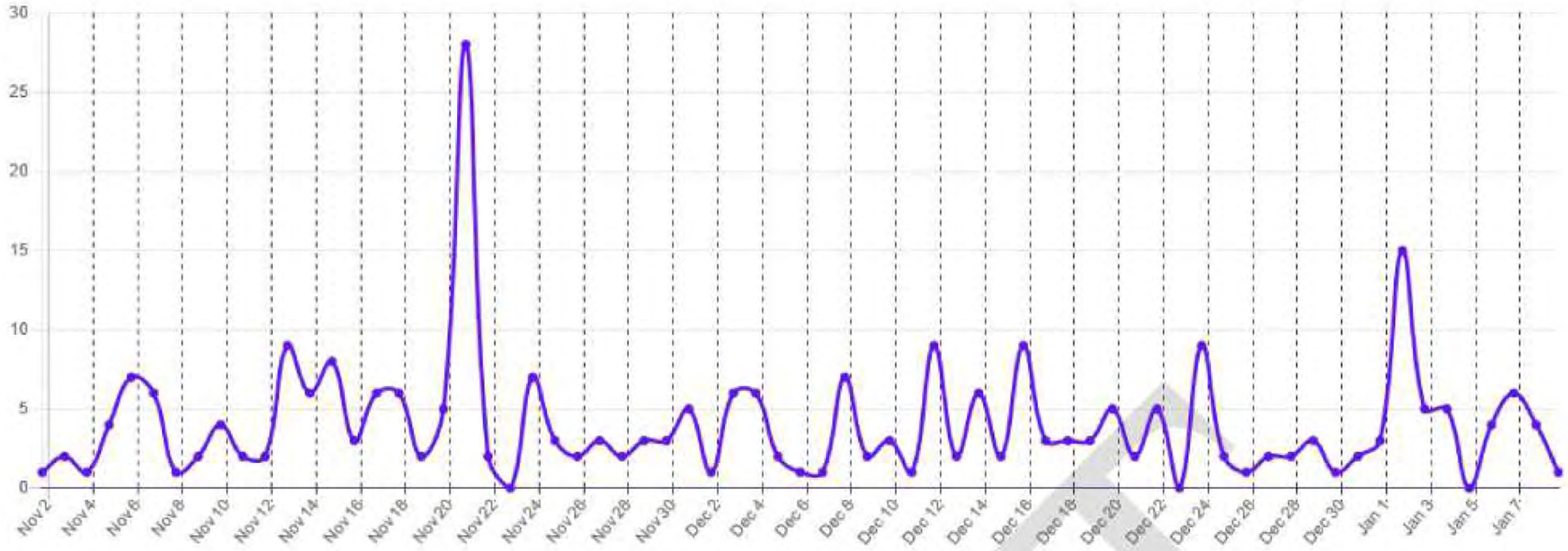


Mentions of "Firing Squad" on Parler



Mentions of "Civil War" on Parler

[372] For more analysis of violent trends on Parler, see memo by Select Committee staff, "Platform-specific findings," Meghan Conroy & Alex Newhouse, August 31, 2022.



Mentions of "Day of the Rope" on Parler

The Trump Organization was negotiating for the purchase of Parler until late 2020, showcasing the importance of the platform to President Trump's base. President Trump's apparent interest in purchasing Parler after the election demonstrates the platform's success in becoming a key part of the far-right media ecosystem. Public accounts have shown that President Trump and Parler were in negotiations over the platform during 2020, especially as conservative dissatisfaction with Twitter and Facebook grew.[373] Internal Parler documents show that Trump Organization officials—including Eric Trump—were involved in these negotiations well into December 2020.[374]

The fact that these negotiations were ongoing while users on Parler were actively plotting the January 6th attack is a striking example of how closely the Trump orbit coalesced with far-right extremists on social media. Parler CEO John Matze invoked the Fifth Amendment when asked about his communications with the Trump family and President Trump himself, including whether conversations about the potential acquisition of Parler continued after January 6th.[375]

Other alternative platforms and fringe sites had even less capacity—and desire—to detect violent moderation than Parler and contributed to spread of violent content prior to January 6th. Like Parler, these platforms catered to the extreme far-right played an important role in coordinated calls for violence on January 6th. Unlike Parler, however, the Select Committee has far less visibility with the extent to which employees communicated these concerns to federal law enforcement. These platforms were also defined by the lack of formal content moderation and a heavy reliance on its users to report illegal content.

Telegram is one such platform. While Telegram is similar to WhatsApp or Signal in terms of user experience, it has long been the home of radicalization, as well as extremist coordination and

---

[373]Ryan Mac and Rosie Gray, *Parler Wanted Donald Trump On Its Site. Trump's Company Wanted a Stake*, BuzzFeed News (Feb. 5, 2021), available at https://www.buzzfeednews.com/article/ryanmac/trump-parler-ownership.
[374] CTRL0000007468, CTRL0000007469.
[375] John Matze Deposition, 39:10-41:7.

collaboration. Previously deemed the "epicenter of terrorist propaganda," especially within the context of ISIS,[376] it has continued to be a haven for white supremacist and militant accelerationists, among others. It was used to promote state capitol protests in the lead up to January 6th, and played an ample role in spreading dangerous disinformation and violent rhetoric about the election.[377]

Moreover, various white supremacist, neo-fascist, and pro-Trump groups experienced growth throughout election season – especially after Trump told the Proud Boys to "stand back and stand by" during the first presidential debate in September 2020.[378] Telegram, like many of the other platforms detailed in this report, is especially useful in sharing content from other platforms; for example, Vincent James, a right-wing YouTuber and influencer, used his Telegram channel to share Trump's December 19th tweet, adding that the tweet was Trump "calling all patriots to the White House on January 6th." He also posted a link to a petition for President Trump to "Invoke [the] Insurrection Act to Take Back Our Republic With Military."[379] Much like other social media and communications platforms, Telegram was used to coordinate and advertise protest events as well as violent memes and rhetoric. For example, the Philadelphia Proud Boys channel shared a photo in mid-December, ahead of the Million MAGA March on December 12th, that read "Shatter Their Teeth."[380] This was all possible, in part, due to Telegram's lack of election-related policies.[381]

Gab became one such platform in the run-up to the attack, with content including discussion of overwhelming police with large crowds and hashtags such as "Storm the Capitol," "civil war," and "Fight for Trump,"[382] though these trends were seen on mainstream platforms, as well. Gab has refused to share the content of its communications with law enforcement and further states that it only received a single notification from law enforcement about interference in U.S. electoral processes.[383]

More than any other platforms under investigation by the Committee, these extreme platforms relied almost entirely on reports from users to surface violent content. In a briefing provided to the Committee, Gab CEO Andrew Torba explained that there was only *one* employee

---

[376] Amarasingam, A. "Telegram deplatforming ISIS has given them something to fight for." VOX-Pol. January 1, 2020. https://www.voxpol.eu/telegram-deplatforming-isis-has-given-them-something-to-fight-for/.
[377] "From Shitposting to Sedition." Moonshot. January 2021. https://moonshotteam.com/resource/from-shitposting-to-sedition-2020-us-elections-report/.
[378] "From Shitposting to Sedition." Moonshot. January 2021. https://moonshotteam.com/resource/from-shitposting-to-sedition-2020-us-elections-report/.
[379] "Evidence of Planning and Violent Discussion in Run-Up to Capitol Riot," Institute for Strategic Dialogue, submitted to the Select Committee.
[380] DFRLab's Domestic Extremism Daily Monitoring Notes, provided to the Select Committee.
[381] "The Long Fuse: Misinformation and the 2020 Election." The Election Integrity Partnership. 2021. Pg. 214.
[386] Craig Timberg, *Gallows or Guillotines? The Chilling Debate on TheDonald.win Before the Capitol Siege*, Washington Post (Apr. 15, 2021), available at https://www.washingtonpost.com/technology/2021/04/15/thedonald-capitol-attack-advance-democracy/
[386] Craig Timberg, *Gallows or Guillotines? The Chilling Debate on TheDonald.win Before the Capitol Siege*, Washington Post (Apr. 15, 2021), available at https://www.washingtonpost.com/technology/2021/04/15/thedonald-capitol-attack-advance-democracy/

responsible for reviewing user reports and otherwise monitoring for illegal content. Around January 6th, Torba estimated that Gab had approximately one million users.[384] In advance of the attack, Gab made no effort to streamline reporting processes for potential threats; Torba told Committee staff that he could not recall any conversations about whether allusions to violence against elected officials or calls to storm the Capitol should be referred to law enforcement.[385]

Fringe platforms gained traction as users migrated from mainstream platforms. The rise of these moderation-free platforms was driven in part by the deplatforming of far-right and conservative users on mainstream platforms throughout the course of 2020. The case in point for this phenomenon is TheDonald.win, which was the successor website to the banned subreddit r/The_Donald (discussed in further detail earlier in this report). In late 2020, users on TheDonald.win shared advice about bringing firearms into Washington, D.C., described the proper kind of zip ties for detaining members of Congress, and shared diagrams of tunnels beneath the Capitol complex.[386] Posts on the platform were extremely clear about their users' intentions. One such post read: "If we occupy the Capitol building, there will be no vote."[387]

The owner of the website, Jody Williams, was a former user moderator on r/The_Donald. He ultimately shut down the website after January 6th and described his unsuccessful efforts to remove the most extreme and violent content in the lead-up to January 6th, including posts from QAnon conspiracy theorists, white supremacists, and Holocaust deniers.[388] Like other alt-tech platforms, it does not appear that TheDonald.win or its successor website had a sophisticated content moderation operation, relying on individual moderators to report content.[389]

Williams testified that TheDonald.win had some automated measures to look for content that was in violation of their rules proactively, largely focused on efforts to dox individuals.[390] While Williams asserted that user-moderators of the website did try to take down specific threats of violence, he also said that some of the moderators were so upset by President Trump's election loss that "they were approving some things that we never would have approved just two months

---

[386] Craig Timberg, *Gallows or Guillotines? The Chilling Debate on TheDonald.win Before the Capitol Siege*, Washington Post (Apr. 15, 2021), available at https://www.washingtonpost.com/technology/2021/04/15/thedonald-capitol-attack-advance-democracy/

[386] Craig Timberg, *Gallows or Guillotines? The Chilling Debate on TheDonald.win Before the Capitol Siege*, Washington Post (Apr. 15, 2021), available at https://www.washingtonpost.com/technology/2021/04/15/thedonald-capitol-attack-advance-democracy/
[386] Craig Timberg, *Gallows or Guillotines? The Chilling Debate on TheDonald.win Before the Capitol Siege*, Washington Post (Apr. 15, 2021), available at https://www.washingtonpost.com/technology/2021/04/15/thedonald-capitol-attack-advance-democracy/

[387] Ken Dilanian & Ben Collins, *There Are Hundreds of Posts About Plans to Attack the Capitol. Why Hasn't This Evidence Been Used in Court?* NBC News (Apr. 20, 2021), available at https://www.nbcnews.com/politics/justice-department/we-found-hundreds-posts-about-plans-attack-capitol-why-aren-n1264291.

[388] Craig Timberg and Drew Harwell, *TheDonald's Owner Speaks Out on Why He Finally Pulled Plug on Hate-Filled Site*, Washington Post (Feb. 5, 2021), available at https://www.washingtonpost.com/technology/2021/02/05/why-thedonald-moderator-left/.

[389] Notes from Select Committee Call with Sean Duggan, Patriots.win.

[390] Jody Williams Deposition, 42:7-15.

before that," particularly posts related to January 6th.[391] Williams told the Committee that there were posts, including directed threats and doxing, that some moderators were no longer trying to shut down following President Trump's election loss, and he worried that the posts could damage President Trump and his movement.[392]

Williams indicated that the top moderators in charge of TheDonald.win were among those who were ignoring the site's internal rules to allow violations to stay online, including one post with instructions about how to tie a noose that was "pinned" by a moderator shortly before January 6th.[393] This sort of behavior eventually convinced Williams to leave the site, because he believed that senior moderators were willing to allow "inciteful" content to remain online because they themselves were upset about the election.[394]

Similarly, Gab CEO Andrew Torba, who posted on the day of January 6 itself that "in a system with rigged elections there are no longer any viable political solutions," refused to answer the Committee's questions about how his personal views may have affected the treatment of content on Gab pertaining to the attack.[395] Torba also indicated that no changes have been made to make Gab more responsive to threat of violent extremism after January 6; changes to the content moderation system were driven instead by a concern about other types of content and the explosion in users that occurred in January 2021, which Torba attributed to an exodus of users from mainstream platforms after President Trump was banned.[396]

These largely anonymous message boards made it easier to make explicit plans for violence ahead of January 6th. 4chan and 8kun, two popular message boards that allow individuals to post anonymously, also allowed for the incubation of violent content in a way that evaded both effective moderation and law enforcement disruption. Communications between 4chan and federal law enforcement in the wake of January 6th suggest that the company is unable to retain most posts and does not cooperate with law enforcement requests absent "emergency involving death or serious physical injury."[397] 4chan has repeatedly denied to law enforcement that is able to identify whether individuals have ever posted on their site, creating a powerful shield against government investigations and potential prosecution of violence.[398]

---

[391] Jody Williams Deposition, 65:1-11.
[392] *Id.* at 69:7-14.
[393] *Id.* at 87:2-88:8.
[394] *Id.* at 69:10-24.
[397] *See, e.g.*, CTRL0000007965.
[397] *See, e.g.*, CTRL0000007965.
[397] *See, e.g.*, CTRL0000007965.
[398] CTRL0000008019.

Example of content 4chan said it could not trace despite law enforcement request.

In terms of planning of January 6th itself, 8kun played a more important role than 4chan. Founded in 2013 as 8chan—an even less moderated cousin of 4chan—the platform has been linked to child pornography, anti-Semitism, Nazism, hate crimes, and mass shootings.[399] More recently, it was the birthplace of the QAnon conspiracy theory.[400]

Prior to January 6th, anonymous 8kun users were openly discussing a violent attack on the Capitol. On January 5, one user wrote: "You can go to Washington on Jan 6 and help storm the Capitol. As many Patriots as can be. We will storm the government buildings, kill cops, kill security guards, kill federal employees and agents, and demand a recount."[401] Posts on 8kun went far enough to debate the politicians that users should target once they got.[402] On TheDonald.win, anonymous users discussed plans to bring weapons, posted schematics of the Capitol, and brainstormed battle tactics, including using a flagpole as a weapon.[403] However, Jody Williams testified that he did not remember FBI ever proactively reaching out to TheDonald.win to follow up on any of these posts prior to January 6th.[404]

As for 8kun, the site's owner, James Watkins, told the Committee that it was his commitment to only take down posts that were violations of the U.S. Code, a task that is ostensibly carried out by a small team of volunteers.[405] This effectively meant that 8kun willingly ceded its platform to

---

[399] Julie Carrie Wong, *8chan: The Far-Right Website Linked to the Rise in Hate Crimes*, The Guardian (Aug. 4, 2019), available at https://www.theguardian.com/technology/2019/aug/04/mass-shootings-el-paso-texas-dayton-ohio-8chan-far-right-website

[400] Kelly Weill, *QAnon's Home 8kun is Imploding—and Q Has Gone Silent*, The Daily Beast (Nov. 13, 2020), available at https://www.thedailybeast.com/qanons-home-8kun-is-implodingand-q-has-gone-silent?ref=scroll.

[401] Ben Collins and Brandy Zadrozny, *Extremists Made Little Secret of Ambitions to "Occupy" Capitol in Weeks Before Attack*, NBC News (Jan. 8, 2021), available at https://www.nbcnews.com/tech/internet/extremists-made-little-secret-ambitions-occupy-capital-weeks-attack-n1253499.

[402] Kari Paul et al., *Far-Right Website 8kun Again Loses Internet Service Protection Following Capitol Attack*, The Guardian (Jan. 15, 2021), available at https://www.theguardian.com/technology/2021/jan/15/8kun-8chan-capitol-breach-violence-isp.

[403] Ryan Goodman and Justin Hendrix, *The Absence of "The Donald,"* Just Security (Dec. 6, 2021), available at https://www.justsecurity.org/79446/the-absence-of-the-donald/.

[404] Jody Williams Deposition, 89:19-23, 90:8-12.

[405] James Watkins Deposition, 24:15-25:8.

its most extreme users, with a promise of an unmoderated space. Watkins testified that this minimalist policy would allow violent threats related to January 6th to remain online, including a call to action after January 6th to "keep killing Jews and leftards all over America." Watkins said that under 8kun's policy, such a post would not have to be removed.[406] There were some posts that were so specifically violent—including one from January 9, 2021, that said users should not come back to D.C. on January 17 if they "are not ready to die" because the day "will end in bloodshed"—that Watkins admitted that they should have been removed from the platform.[407] However, he said that such posts should remain online because "they're history now," and did not say that 8kun had considered any expansion of its content moderation practices given the examples of prohibited content that were not removed in a timely fashion.[408]

Another platform that catered to militia groups during the COVID-19 lockdown and beyond, MyMilitia, also had no ability to take proactive measures; instead, their content moderation consisted of six or seven moderators reviewing posts that are reported by users.[409] In a deposition with the Select Committee, MyMilitia owner Josh Ellis explained that these moderators—who were just private volunteers—were only able to review content from public forums, and not private chats that were utilized by militia groups throughout the country, meaning that the forum relied on the members of the militias themselves to report violent content.[410]

Just as TheDonald.win and Parler experienced growth throughout 2020, MyMilitia is a prime example of the migration of extremist audiences to platforms where they were freer to post violent content in the run-up to the attack on the Capitol. Ellis told the Select Committee that about 10,000 more users signed up on the militia networking site from the onset of the COVID-19 pandemic, increasing its user base by a third, which Ellis attributed to news coverage of MyMilitia as a "good place to meet – other patriots who, you know, want to make sure that we remain a constitutional republic."[411] MyMilitia was, indeed, largely centered around coordinating offline meetings. For example, in mid-October 2020, MyMilitia users began advertising an event to be held in Florida. The "American Patriot Rally" was to be held on October 24, 2020, and featured Josh Ellis as a speaker.[412] Around the same time, Ellis used his podcast to recruit attendees for a "Free Kyle" rally for Kyle Rittenhouse to be held in Waukegan, IL, on October 30th.[413] The podcast is still available to watch on YouTube.[414]

---

[406] James Watkins Deposition, 106:25-107:17.
[407] *Id*. at 111:10-16.
[408] *Id*. at 111:17-112:9.
[409] Josh Ellis Deposition, 14:17-15:23.
[410] *Id*. at 17:4-22.
[411] *Id*. at 19:22-20:8.
[412] DFRLab's Domestic Extremism Daily Monitoring Notes, provided to the Select Committee.
[413] DFRLab's Domestic Extremism Daily Monitoring Notes, provided to the Select Committee.
[414] As of September 5, 2022.

Moreover, throughout October, MyMilitia users were anticipating – excitedly – that "SHTF" (shit hits the fan) in the wake of Election Day. Much of the discussion centered around being prepared and/or ready for what was to come.[415] Some posts included[416]:

- "I've been wondering about possible scenarios that could happen if indeed fighting begins after the election. And I've had some concerns about what armed combat in a potential conflict could mean for all of us that choose to stand against the tyranny of the leftists."
- "Past Nov 3 the gloves come off. If you defy curfew, if you're a part of a riot, then your part of a coup and the police will have orders to open-fire shoot-to-kill, so be prepared to pay with your life if you think you're going to bully this nation into submission. We are NOT going to let this nation succumb to blackmail or an Obama 'color revolution'. If the police have to kill Americans trying to stage a coup d e'tat, well, freedom isn't free. This is as real as t gets folks. Be prepared."

The Facebook knockoff MeWe was similar to MyMilitia in its hosting of anti-government actors and their violent discussions. In the MeWe group "Three Percenters – 1776 Patriots United," one user claimed that an Indiana militia was going to deploy to the state's Capitol building on Election Day because "antifa has promised to assault every red state, kidnap GOP representatives, and murder them."[417] This particular MeWe group contained discussions of "clearing the streets" of "commies," wherein members asserted, "after a thousand die the rest will grow up. They need to be reminded this is the home of the brave."[418] The group featured memes such as BLM and Marxist hunting permits, as well as targeted violent rhetoric like, "When the shooting starts, I will absolutely go out, find evil, and kill it."[419]

As Election Day loomed, the intensity of the threats continued to escalate, reaching a point where users were debating whether to "start with the locals with Biden signs" and "then work your way out," or to "start at the TOP, and work your way down. HVT's [high value targets] first!"[420]

Another MeWe group titled "AMERICAN CIVIL WAR 2.0" contained similarly violent discourse, with one user urging his fellow MeWe users to "grab your armor, rifle and combat" and to "load & deploy with your local militia." Other users asked about deploying III% militia members at polling stations.[421] This particular chat featured posts like the following:

---

[415] DFRLab's Domestic Extremism Daily Monitoring Notes, provided to the Select Committee.
[416] DFRLab's Domestic Extremism Daily Monitoring Notes, provided to the Select Committee.
[417] DFRLab's Domestic Extremism Daily Monitoring Notes, provided to the Select Committee.
[418] *Id.*
[419] *Id.*
[420] *Id.*
[421] *Id.*

> "Civil war is inevitable. And mark my words, they're gonna start shit before the sun sets on Election Day. It will go wide on the 4th, if they stick to their plans. My intel is pretty accurate and that's what it's telling me. […] Leftists have already fired the first shots, and have murdered MANY conservatives. Believe me, when the Right eventually hits back, the bloodshed will be Biblical."[422]

Boogaloo groups on MeWe also emphasized the need to "get with your local goons and start training," and complained that the online movement is "dying" due to "all the [platforms'] censorship."[423] To MeWe's credit, the platform began terminating Boogaloo groups, forcing them to other messaging platforms like Keybase.[424]

Ellis of MyMilitia denied any awareness that there was activity on his site about protests at state capitols or vote centers in the weeks following the election.[425] This exact type of activity unfolded on various platforms – including MyMilitia – after November 3rd. For example, Bucks County Women for Trump and PAPN (Proud American Patriots Network) used MyMilitia to market an event on November 21, 2020, at the Doylestown, PA courthouse, encouraging possible attendees to, "show the deep state that we still support our president."[426] In that same vein, the One Republic Society planned a December 12th Patriot Rally in Columbia, SC, posting flyers for the event on MyMilitia.[427]

This lack of awareness is yet another example of how these platforms were unable to take down content that violated their own, minimal terms of service, as the atmosphere in closed chats and essentially unmoderated boards turned uglier after Election Day. Indeed, as President Trump's legal efforts to overturn the election results failed time and time again, MyMilitia became a hotbed for discussion regarding whether it was time to "suit up" or hold off, acknowledging that the time may soon arrive.[428] Protests continued to be coordinated and advertised on the forum throughout December 2020.

After the election, MeWe users similarly began to discuss more kinetic responses to perceived election fraud. In particular, one user wrote that "The Kentucky militia needs to mobilize… arrest the Governor if need be. Same with Ohio. I'm working on Indiana."[429]

These platforms' users very closely followed the cues of President Trump in the weeks prior to January 6th. Jody Williams told us that the audience on TheDonald.win was extremely reactive to President Trump's narrative of a stolen election. Users were still generally

---

[422] *Id.*
[423] *Id.*
[424] *Id.*
[425] Josh Ellis Deposition, 32:14-19.
[426] DFRLab's Domestic Extremism Daily Monitoring Notes, provided to the Select Committee.
[427] *Id.*
[428] *Id.*
[429] *Id.*

"pretty upbeat" the day after the election because they did not believe that the results were finalized, but as President Biden's victory became clearer, according to Williams, content grew angrier and clearly believed there was "misconduct" in the election and began to use the language of "Stop the Steal."[430] And while Williams testified that users on TheDonald.win had been talking about traveling to D.C. since the election, but that after the President's December 19 tweet, "anything else was kind of shut out, and it just was going to be the 6."[431] James Watkins, owner of 4chan, also said that the President "did call" his supporters to D.C. when he tweeted out his message on December 19 and said that he made his own decision to attend the rally on January 6 because the President asked him to do so.[432]

Internal content moderation logs obtained by the Select Committee show the barrage of racist, violent content that flooded TheDonald.win in the days and weeks following Election Day, such as one post that said, "I'd buy season tickets to watch public executions of traitorous cucks."[433] While the moderation logs obtained by the Select Committee show that many of the violent posts were removed, the steady stream of posts—coupled with the testimony that many moderators were *re*-approving violent content after it had been removed—shows the depths of the problem on TheDonald.win. Internal emails show that the website's domain host was aware of moderators pinning "what looks like a call to shoot people" in mid-December 2020.[434]

The posts also show a reliance on the claims of election fraud that had been disseminated by President Trump and his allies, with one user posting, "RUDY, LIN WOOD, BYRNE, SYNDEY. NOW IS THE TIME TO PUT UP or STFU ABOUT YOUR EARTH SHAKING REVEALS, AND RECIEPT'S!" (*sic*).[435] One post, made by a user who Williams identified as a top moderator that was willing to keep up problematic content prior to January 6th, declared after the attack that "The Donald will continue, as it always has, to follow President Trump's lead."[436]

Even these darkest corners of the internet had a nexus to Trump's inner circle. Ron Watkins, a former 8kun administrator, the son of 8kun owner Jim Watkins and currently a House candidate in Arizona, cultivated a mainstream social media following spreading election disinformation. Internal Trump campaign documents show that Rudy Giuliani's team believed Watkins to be a relevant influencer who should be leveraged in the campaign's efforts to spread election fraud narratives.[437] Leading up to January 6th, Watkins was retweeted by the former

---

[437] CTRL0000030069.
[437] CTRL0000030069.
[437] CTRL0000030069.
[437] CTRL0000030069.
[437] CTRL0000030069.
[437] CTRL0000030069.
[437] CTRL0000030069.
[437] CTRL0000030069.

President several times before January 6th and supplied an affidavit in a lawsuit filed by Sidney Powell focusing on claims of fraud by Dominion Voting Systems.[438]

Separately, public reporting suggests that Dan Scavino monitored these far-right platforms, including TheDonald.win, where users reacted to a post by Scavino by interpreting it as "literal war drums" and suggesting an armed march on the Capitol.[439] While the Committee was unable to confirm that Scavino or any other Trump allies maintained accounts on TheDonald.win, Jason Miller did send Mark Meadows a text on December 30, 2020, declaring "I got the base FIRED UP" and linking to a post on TheDonald.win, indicating some level of familiarity with the website that was simultaneously seeing a flood of violent content leading up to January 6th.[440]

Donald.win users react to a video posted on Twitter by Dan Scavino on Dec. 30, 2020.

These were longstanding tactics used by Trump campaign staff. According to analysis by Justin Hendrix of *Tech Policy Press* and *Just Security*, during the 2016 election "a team in the war room at Trump Tower was monitoring social media trends, including the r/The_Donald subreddit…**and privately communicating with the most active users to seed new trends**'

---

[438] Rachel Greenspan, *Trump Shared a Video That Featured a Former Administrator of QAnon's Fringe Message Board Spreading Election Misinformation*, Business Insider (Nov. 19, 2020), available at https://www.insider.com/trump-tweets-video-voter-fraud-dominion-former-qanon-message-board-2020-11; Drew Harwell, *To Boost Voter-Fraud Claims, Trump Advocate Sidney Powell Turns to Unusual Source: The Longtime Operator of QAnon's Internet Home*, Washington Post (Dec. 1, 2020), available at https://www.washingtonpost.com/technology/2020/12/01/powell-cites-qanon-watkins/.

[439] TheDonald.win, *Dan Scavino—THIS IS NOT OVER!* https://patriots.win/p/11RO7PPPMH/dan-scavino--this-is-not-over-/c/ (accessed Apr. 27, 2022); Justin Hendrix, *TheDonald.win and President Trump's Foreknowledge of the Attack on the Capitol*, Just Security (Jan. 12, 2022), available at https://www.justsecurity.org/79813/thedonald-win-and-president-trumps-foreknowledge-of-the-attack-on-the-capitol/.

[440] *See* MM014441.

(emphasis added). In May 2017, *The Atlantic* pondered, "Is Donald Trump a Secret Redditor?" 'Probably not, but his staffers might be,' the article concluded, **identifying examples when content the President tweeted had appeared on The Donald just *before* he tweeted it."**[441] Many of those who viewed Trump's December 19th "be there, will be wild" tweet as a call to arms on January 6th embraced the notion that violence was not only inevitable, but *preferable*.

Jody Williams told Select Committee staff that the audience on TheDonald.win was extremely reactive to President Trump's narrative of a stolen election. Users were still generally "pretty upbeat" the day after the election because they did not believe that the results were finalized, but as President Biden's victory became clearer, according to Williams, content grew angrier and clearly believed there was "misconduct" in the election and began to use the language of "Stop the Steal."[442] And while Williams testified that users on TheDonald.win had been talking about traveling to D.C. since the election, he said that after the President's December 19th tweet "anything else was kind of shut out, and it just was going to be the 6."[443] James Watkins, owner of 8kun, also said that the President "did call" his supporters to D.C. when he tweeted out his message on December 19th and said that he made his own decision to attend the rally on January 6th because the President asked him to do so.[444]

One post, made by a user who Williams identified as a top moderator that who willing to keep up problematic content prior to January 6th, declared after the attack that "The Donald will continue, as it always has, to follow President Trump's lead."[445]

In these relatively unmoderated spaces, users frequently used extreme language, most notably anti-Semitic, homophobic, and racist terms and slang that are outright banned on mainstream platforms. Homophobic slurs were so widespread as to be saturated throughout the platform. In addition, users on the dot win sites did not shy away from employing the (((echoes))) orthographic marker, which is long-established in white supremacist and conspiracy theory circles as a symbol for alleged, malicious Jewish influence in events. For example, in response to the December 19th tweet, one user posted a long, deeply nihilistic screed doubting whether a march on Jan. 6 would have any impact, which contained the following line:

> "The clock is ticking. And it seems the faggot deep state advisors around GEOTUS like (((Jared))) and others are trying to talk him out of military intervention or refusing to leaving the white house."

---

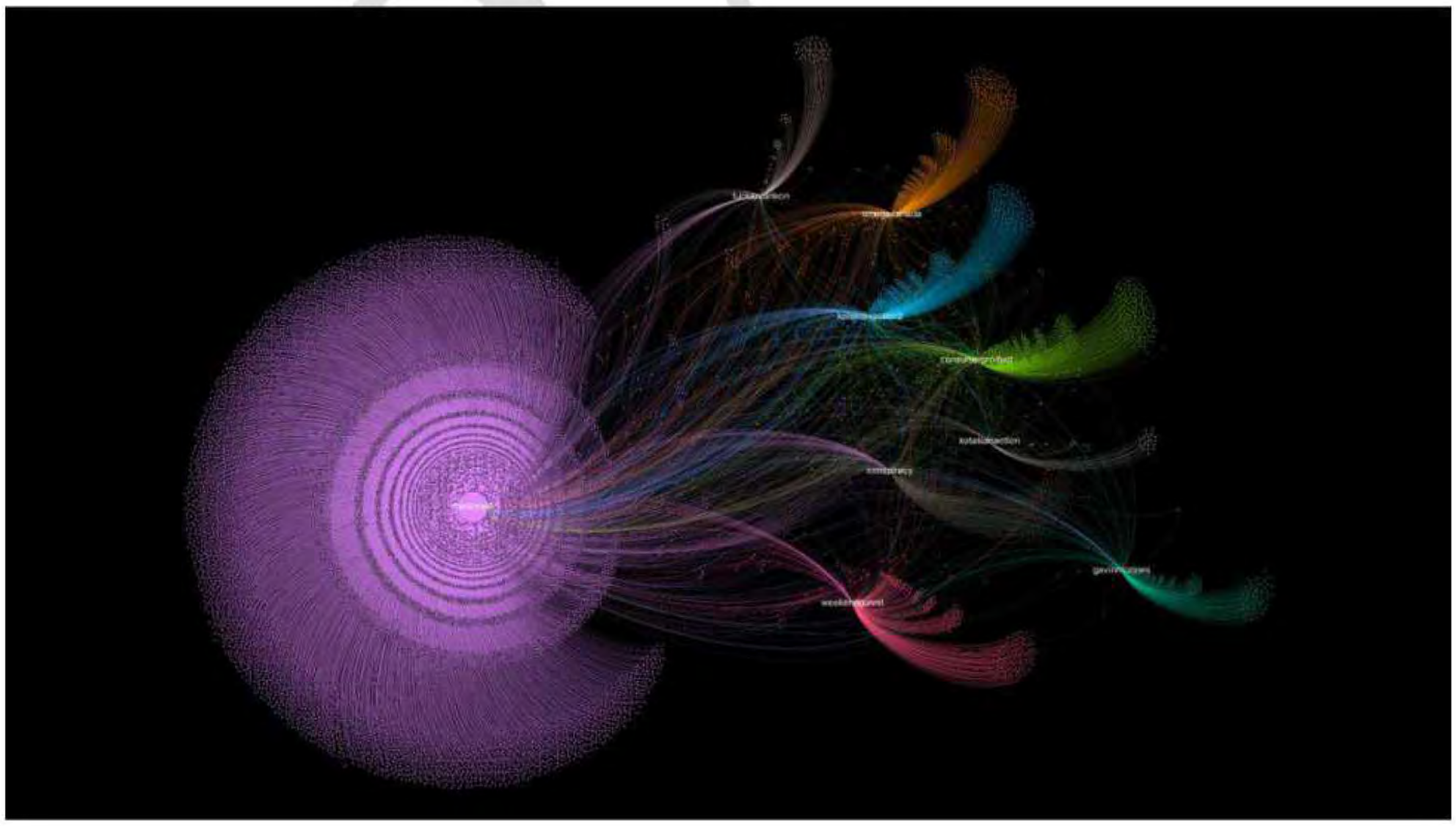
[441] Adrienne LaFrance, "Is Donald Trump a Secret Redditor?" *The Atlantic* (May 11, 2017). https://www.theatlantic.com/technology/archive/2017/05/is-donald-trump-reading-reddit/526425/.
[442] Jody Williams Deposition, 55:15-56:4.
[443] *Id.* at 72:3-11.
[444] James Watkins Deposition, 74:3-9, 76:22-24
[445] *Id.*

Anti-Black rhetoric was used frequently as well and often mirrored the type of language used on fringe Telegram chats and 4chan boards. These included racist terms like "chimpout" to refer to social justice or left-wing demonstrations, "jogger" as a stand-in for a racial slur, and discussions of falling white birthrates and black people "replacing" whites.

The presence of extremist content in TheDonald likely results in part from its proximity to more fringe forums in the dot win network of websites. WeekendGunnit, for instance, is a Boogaloo-oriented website that was created following enforcement against its predecessor subreddit. ConsumeProduct is a free-wheeling forum similar to 4chan's /pol/ board that houses users who routinely promote racial slurs and neo-Nazi ideologies. Although TheDonald served as the flagship website for the dot win network, its community nonetheless developed connections to these other boards where extreme content is more central.

The network topology below shows the status of these connections in the time period between November 3, 2020, and January 6th, 2021, based on a dataset of dot win data. While this data is likely not comprehensive, it provides an important view into the ways that the communities were sharing users.

The network topology below shows the status of these connections in the time period between November 3, 2020, and January 6th, 2021, based on a dataset of dot win data. While this data is likely not comprehensive, it provides an important view into the ways that the communities were sharing users.



In this network map, colors represent clusters of users for each dot win site (e.g., purple is TheDonald, green is ConsumeProduct). This graph shows that all dot win sites overlap significantly with TheDonald's community, which enables flows of extreme narratives and content.

IX. Other Platforms

Separate from platforms that were meant to cater almost exclusively to far-right extremists, other small- to mid-sized social media platforms—including Discord, Zello, Twitch, and MeWe—also contributed in varying degrees to the organization of the attack and the spread of violent content and disinformation. Like Parler, these platforms suffered from somewhat underdeveloped content moderation policies that made them ill-equipped to confront organized far-right extremist prior to January 6th. However, unlike Parler—let alone some of their larger peers—these companies have provided evidence of varying degrees of corrective action that they have taken since the attack itself.

The role of other, smaller platforms in the attack on January 6th was highly dependent on their unique functionalities, which caused some to be caught off-guard by extremist organizing. Just as far-right message boards like 8kun and TheDonald.win were the perfect vessel for particularly violent postings given their anonymous and unmoderated nature, niche platforms had often limited roles in the attack that fit their unique profile.

Discord

For example, Discord saw massive growth in its user base during 2020 due in part to its ability to give users private message channels based on discrete topics. Discord claims that it has devoted substantial resources to combatting violent extremism since it was used to help organize the Unite the Right rally in 2017; 15% of its employees are devoted to Trust & Safety and engage in both proactive searches for dangerous content and reactive reviews of user reports.[446]

There are some indications that Discord did attempt to take proactive actions against organizing extremists' attempts to leverage its platforms. An internal report on its Trust & Safety activities for the second half of 2020 shows that 1,504 servers were removed for violent extremism between July and December, which is a 93% increase from the first half of the year.[447] This statistic could reflect both increased extremist activity and migration of extremist users from larger platforms to places like Discord.

In a briefing, Discord told the Committee that it was aware of the risks of relying on relying too much on user moderation when the userbase may not have an interest in reporting problematic content, which informed some of the proactive monitoring it underwent during the weeks before January 6th. Bri Riggio, Discord's Policy Platform Monitor, explained that one important aspect of Discord's proactive monitoring was looking at TheDonald.win to see whether there were

---

[446] CTRL0000020826.
[447] CTRL0000020824.

indications that users on the platform were seeking to utilize Discord.[448] This kind of incorporation of off-platform signals to determine the level of risk has proven an important piece of managing threats from violent extremism.

Internal reviews show how important President Trump's messaging was to galvanizing actors prior to January 6th and the broad scope of far-right extremist organizing. An internal review of Discord's January 6th response showed at least some awareness of how the platform might be used by extreme groups, especially by advertising Discord servers on other forms such as TheDonald.win or 8kun.

Discord says that its tracking began in earnest after it saw an increase in potential calls to violence after President Trump's December 19 tweet, and its Anti-Extremism Team began actively reviewing servers that it identified as being potentially used to organize plans to go to D.C. The after-action report also admits that there was not a "proactive plan" for January 6th but had a team in place to review organized extremist content; additionally, it notes that it should expect that the attack is "serving to galvanize new militia organizing" and should result in proactive tools to track future militia servers.[449]

Discord's internal deliberations about the decision to potential ban some of servers related to violent extremist organizing also provide a window into how tech platforms consider the prospect of real-world harm in the lead-up to January 6th and shows the centrality of President Trump's messaging to that harm. For example, an internal review of DonaldsArmy.US server showed a spike in activity following the President's December 19 tweet, while another internal review of The Donald server indicated a similar pattern.[450] Although heavily redacted, these reports also show the intersection between Discord servers and other platforms, especially TheDonald.win, underscoring the network of alt-tech websites in enabling extremist organizing.

Discord explained that, following the attack, moderation staff at Discord began to see indications that membership in TheDonald.win was being used as a vetting process to enter The Donald server, which was experiencing an explosion of activity that made Discord concerned that it would be used for future violent acts.[451] Discord's Platform Policy Manager Bri Riggio said that there were no indications that The Donald server and TheDonald.win were so intimately connected before January 6; however, Jody Williams, former owner of the TheDonald.win, testified that he was in charge of managing the website's Discord server prior to January 6th, indicating that the pro-Trump forum was seeking to utilize Discord earlier than the platform itself realized.[452]

---

[448] Memo on July 29, 2022, Select Committee Briefing with Discord.
[449] CTRL0000028919.
[450] CTRL0000062532, CTRL0000034886.
[451] Memo on July 29, 2022, Select Committee Briefing with Discord.
[452] Jody Williams Deposition, 15:11-17.

Even more significantly, Riggio explained to the Committee that the internal review the DonaldsArmy.US showed that individuals on that served viewed the President's tweet as a "call to action," and that users on TheDonald.win were directing individuals to use that server as a place to organize their response to the President's call for them to come to Washington, D.C.[453] Although the server had existed for weeks prior to the President's tweet, Discord had not noticed significantly violations of site policy occurring until the President's call.

In the hours after the tweet, Discord saw very clear organization on the server, including users trying to connect with each other by region, plans to travel to D.C., and, most concerningly, discussion of how and whether to evade D.C. gun restrictions and bring firearms into the city.[454]

In addition to demonstrating the fast and furious reaction to the President's December 19 tweet, the DonaldsArmy.US server also shows how extreme corners of the internet—in that case, TheDonald.win—directed their followers to more sophisticated, mainstream platforms to take advantage of their technological capacity to respond to the President's orders. Discord banned DonaldsArmy.US just a few hours after the President's tweet.[455] Riggio explained that, following the President's tweet, there was a shift in other forums on Discord related to election misinformation.[456]

The DonaldsArmy.US server was the main instance of mobilization for January 6th that Discord detected on its platform, as well several other servers that were used in the preplanning stages of the attack. On the day of the attack itself, Riggio said that there were only isolated reports of organization and the glorification of violence on that day, indicating that the main utility of Discord came prior to the attack.[457] The weekslong interplay between TheDonald.win and various servers on Discord demonstrated the continuing role that non-extreme platforms played in fomenting the organizing efforts of far-right extremists even after major companies such as Reddit had taken action against them.

Proactive, dynamic enforcement measures are an essential component of making alt-tech and developing platforms less hospitable to extremist organizing in the future. Internal documents from Discord also explain how its team identifies proactive measures that its Trust & Safety Team can take to supplement reports from users themselves.[458] Discord also has instructed its employees on different layers of extremist organizing and established different courses of action for varying degrees of harmful activity.[459] This is a contrast to the absolutist First Amendment approach taken by platforms like 4chan, 8kun, and Parler, whose stances on moderation remain unchanged since January 6th.

---

[453] Memo on July 29, 2022, Select Committee Briefing with Discord.
[454] *Id.*
[455] *Id.*
[456] *Id.*
[457] *Id.*
[458] CTRL0000062533.
[459] CTRL0000062534.

Policy considerations for newer, smaller platforms are heavily influenced by practices of Big Tech giants and fears about overpromising on moderation. Twitch's policy proposals draw directly on the content moderation practices of larger networks, including monitoring the topics that Twitter and Facebook consider to be misinformation.[460] Discord's internal discussion of policies on Disinformation and Violent Organization in the wake of the election were also heavily influenced by the actions taken by larger platforms, including Twitter's Civic Integrity Unit, which places further pressure on those large platforms to be transparent leaders about their efforts to moderate far-right extremist organizing.[461]

Fig. 8. Discord's internal model for considering extremist content

Zello

On the other hand, Zello—an audio app that essentially functions as a walkie-talkie—was utilized as a live mode of communication for militias and other groups who participated in the attack on the Capitol.[462] The live nature of this usage made it more difficult for them to identify content.[463] As we have seen with other platforms, the increase in far-right extremist organizing affected Zello throughout the course of 2020. Internal emails show that in the weeks prior to the election, Zello was grappling with how to remove far-right channels in response to negative press. The emails indicate a rather scattershot approach; only 59 channels were blocked.[464] Another internal working document from Zello also shows the reactive nature of their counterextremism efforts prior to January 6th, as employees propose new search terms and moderation policies, including a "formal policy regarding armed militia use."[465]

In a briefing with the Select Committee, Zello CEO Bill Moore explained that January 6th was a "watershed moment" for the app that sped up its shift away from social networking components towards a more business-oriented walkie-talkie app.[466] However, the briefing also revealed major vulnerabilities in Zello's functionalities that would have made it easier for extremists to leverage the app prior to the attack on the Capitol. The most glaring of these was the inability of

---

[460] CTRL0000007925; see also CTRL0000010149.
[461] CTRL0000028918.
[462] Micah Loewinger and Hampton Stall, *Revealed: Walkie-Talkie App Zello Hosted Far-Right Groups Who Stormed Capitol*, The Guardian (Jan. 13, 2021), available at https://www.theguardian.com/us-news/2021/jan/13/zello-app-us-capitol-attack-far-right.
[463] CTRL0000007646.
[464] CTRL0000007648 at ZELLO 000078
[465] *Id.* at ZELLO 000082.
[466] Memo on May 11, 2022, Select Committee Briefing with Zello

Zello's moderators to access private channels; therefore, most of the channels that were being utilized by militia groups and other extremists prior to January 6th were closed off to Zello employees, leaving them overly reliant on user reports of policy violations.[467] While this still left Zello with the ability to take action against channels and users whose names were in violation of the policies, the practical impact of this lack of visibility into private channels is that the leadership of the platform did not have a complete view of the threat landscape that existed prior to January 6th. In contrast, Discord has the capacity to look at closed servers and have proactive reviews of servers that are likely to result in high-harm content, including violent extremism.[468]

During the briefing, Zello confirmed that the FBI had never reached out to them about any specific threats of violent extremism on the app and contended that the actions Zello took to ban certain channels before the election was more about internal anxiety about extremism than knowledge of actual threats.[469] When asked by Committee staff about specific actions that its moderators took during the period between Election Day and January 6th, Moore said that they did not block or ban any users or channels during that period, despite "looking for problems."[470] However, January 6th, Zello banned over 2,000 militia-related networks. This is a reactive and belated counterextremism measure given the platform's awareness of problems with extremist organizing in the months leading up to the election, but the sheer size of removed channels gives a vivid picture of the scope of militia and far-right organizing.[471]

As discussed, it appears Zello took a far less proactive approach in tackling known instances of extremist utilization of their service, which is starkly illustrated by the need to remove thousands of militia-related communication channels following January 6th. Zello stated that many of these channels were likely small or no longer in use, which was the case for many of the channels that were taken off the platform prior to the election, although Zello was not able to access many channels that were private.[472]

However, the Select Committee has collected evidence to show that these closed channels involved pre-planning and and coordination between different elements of the far-right in the run-up to January 6th. For example, MyMilitia owner Josh Ellis, who was present in D.C. on January 6th, told the Committee he was on Zello channels with Proud Boys, Oath Keepers, other militia members, and "regular patriots" in the leadup to January 6th and in response to President Trump's December 19th tweet.[473] He confirmed that these channels included the sharing of intelligence.[474]

---

[467] *Id.*
[468] Memo on July 29, 2022, Select Committee Briefing with Discord.
[469] *Id.*
[470] *Id.*
[471] CTRL0000007646.
[472] Memo on May 11, 2022, Select Committee Briefing with Zello
[473] Josh Ellis Deposition, 38:18.
[474] *Id.* at 41:16-20.

Twitch

The livestreaming app Twitch was also caught off-guard by users who were promoting violence on their streams. However, it does not appear that Twitch issued many enforcement actions as a result of the attack on the Capitol; only 11 users were disciplined for showing inappropriate content.[475] Twitch confirmed that it did not see a significant spike in activity on January 6th itself, and that most of the enforcement actions related entailed users posting violent cable news footage or some isolated instances where the violence was glorified.[476] Despite the relative lack of activity on Twitch itself, internal documents from following January 6th show the failings of Twitch's content moderation policies in the run-up to the attack. Most significantly, the suspension of r/DonaldTrump's stream on January 20, 2021, was attributed to the need to proactively deal with "rhetoric encouraging violence" even when it occurs outside of Twitch itself, and staff noted a need to update policies to reflect this.[477]

In a briefing with the Committee, Twitch General Counsel Steve Bene said that the decision to President Trump's channel was due to the ongoing risk that there would be inciteful content on the channel, but was not based on any content posted on Twitch nor on any nonpublic knowledge of his behavior.[478]

More broadly, Twitch—like Zello—was forced to rely on user reports of violative behavior in the run-up to January 6th. The automated tools that Twitch uses to proactive sweep for violations of its policies were not well-equipped for detecting things such as election denialism or calls to violence, and were more targeted towards preventing nudity or violence.[479] Twitch told the Committee that it did not perform any proactive sweeps of streams or comments for disinformation or violent content related to the election in prior to January 6th, meaning that it relied largely on user reports in order to detect relevant misconduct, which leaves open the possibility that there was traffic that went undetected by Twitch employees.[480]

Twitch also implemented several new policies in the wake of the attack, which sought to formalize its nascent content moderation strategy after it had been forced to react to President Trump's calls for violent action without a clearly applicable incitement policy.

Internal memos from Twitch show that the platform is now trying to craft broad language in a new Incitement to Violence Policy that will allow it to respond to influential accounts that are likely to result in real-world harm, including preemptively suspending accounts when the risk of

---

[475] CTRL0000007930.
[476] Memo on May 11, 2022, Select Committee Briefing with Zello.
[477] CTRL0000007942
[478] Memo on May 10, 2022, Select Committee Briefing with Twitch.
[479] *Id.*
[480] *Id.*

harm is high enough.[481] The internal explanation of a new proposed Misinformation Policy that focuses on evolving trends in misinformation and indirect societal harm. [482] While the Misinformation Policy was based on a number of factors, including the spread of COVID-19 misinformation, the Incitement to Violence Policy was a direct response to the events of January 6th and the inciting behavior of President Trump.[483]

In a briefing with the Committee, Twitch General Counsel Steve Bene explained that the after the attack, Twitch did become concerned that its incitement policy did not reach far enough. He mentioned that President Trump had streamed on Twitch on January 5, 2021, and used the "Stop the Steal" hashtag, but under existing policy Twitch required "explicit incitement" to take action."[484] Now, Twitch is able to preemptively suspend prominent accounts when there is a high likelihood of incitement by looking at a number of factors, including the individual's influence and the scale of the ongoing threat. Bene explained that this policy was crafted after January 6th as a way to deal with world leaders and other prominent figures who could utilize Twitch to incite violence.[485]

Expert analysis given to Twitch also warns about potential downsides weighing on platforms who are deciding whether to implement policies on misinformation. The experts wrote that "by releasing a policy, it does open up the platform to more questions and scrutiny over time" including about its precise definition of mis- and disinformation, the transparency of enforcement mechanisms, and how quickly the platform can detect and respond to misinformation narratives and defuse threats.[486]

Specifically, the problem of what to do misleading or aggressive behavior by candidates related to elections is discussed as a borderline case of misinformation that may pose a challenge to Twitch.[487] In the context of continuing claims of election fraud by former President Trump, this kind of internal analysis shows why social media platforms may choose to shy away from taking firm stances on misinformation rather than risk backlash by misapplying policies or making controversial decisions. This reinforces the need for industry leaders like Facebook and Twitter to be open about their efforts to tackle these issues and lean into content moderation debates meant to effectively defuse threats of violence and extremist organizing.

Additional Platforms

---

[481] CTRL0000007943
[482] CTRL0000007924; CTRL0000007925.
[483] Memo on May 10, 2022, Select Committee Briefing with Twitch.
[484] *Id.*
[485] *Id.*
[486] CTRL0000010151.
[487] *Id.*

The above analysis of various platforms and sites and their respective roles in cultivating the radicalization, mobilization, and coordination of those responsible for the attack on the Capitol on January 6, 2021, has zeroed in on the platforms we determined to have played the most significant roles.

That said, the online ecosystem is ever-evolving, and communities of users are ever-migrating. Other platforms and sites that played a part in hosting and/or spreading election disinformation and violent rhetoric leading up to January 6th include but are not limited to:

- AR15.com
- Bitchute
- DLive
- Eventbrite
- GoFundMe
- Instagram (part of Meta)
- KiwiFarms
- Periscope
- Pinterest
- RocketChat
- Snapchat
- Rumble
- Vimeo

X. Broader Conclusions & Recommendations

Below is an overview of some broader conclusions about the challenges facing the social media landscape based on the experience before, after, and during the January 6th attack on the Capitol.

- **Irresponsible speech by politicians matters.** Experts and social media professionals pointed to President Trump's statements across multiple platforms and points in time as key factors in the inflammation and mobilization of violent extremists. Social media companies must take explicit and implicit incitement to violence seriously.
- **Social media companies struggle to strike the right balance between false positives and false negatives.** The scale of content on the platform necessitates automated decision-making—but mistakes are inevitable because AI models take time to create and cannot flawlessly interpret the nuances of human expression. The cost of minimizing mistakes is a higher degree of legitimate speech being removed from the internet; the cost of preserving user voice is a larger amount of harmful content. Platforms are in the difficult position of making these decisions without democratic legitimacy. More societal understanding of and conversation about this tradeoff is crucial to the future of social media and free speech.
- **Soft interventions can be more subtle and more powerful than removal.** Most social media companies already use machine learning to shape the flow of content across their platform. They also use it to automatically surface, label, demote, or otherwise treat harmful content. Such actions can be less damaging to user "voice" than content takedowns or account suspensions and the consequences for false positives are less binary. They should be more transparently studied, developed, and refined.
- **It is crucial to policymaking and public confidence that platforms become more transparent and consistent.** The Select Committee's findings demonstrate how complex the social media landscape has become. Policymakers scrutinizing social media's political and societal impact are often working from incomplete information, as are scholars and policy experts working on related topics. Opacity from social media platforms inhibits policymaking; it also undermines public confidence in platform content moderation and leaves companies vulnerable to unsubstantiated accusations of bias and censorship.
- While social media platforms may contribute to polarization *generally*, January 6th was driven by the *radicalization* of a smaller subset of users. On Facebook, Stop the Steal content, like QAnon and militia content, is associated with a relatively small, homogenous group of users. During the election, content spread amongst those users was harder to detect than widespread viral content but contributed to tremendous offline harm. Stop the Steal had significant overlap in membership with QAnon and militia groups which had already been banned, and the movement behaved similarly to dangerous actors Facebook had responded to in the past.

- **Platforms increasingly look to the rest of the internet to anticipate threats and form strategy.** Narratives or operations which begin on one platform do not stay there. Fringe or encrypted spaces can give early warning signals for nascent efforts to incite or mislead the wider public. And actors who are deplatformed from mainstream platforms may seek haven in darker corners of the internet, where they contribute to growing radicalization. Most of the larger platforms investigated by the Select Committee understand this and have teams dedicated to offsite threats.

XI. Key Evidence

Twitter

- "Twitter Policies and Enforcement: A Timeline of Events Surrounding the 2020 Election and the January 6th, 2021, Attack on the Capitol," Prepared for the Select Committee to Investigate the January 6th Attack on the United States Capitol, March 3, 2022.
- "Twitter's Responses to Select Committee Staff Questions of March 16, 2022." Twitter (April 15, 2022).
- Deposition with J. Smith, a Twitter whistleblower. J. Smith conducted another session of her deposition on September 1, 2022, under her name: Anika Collier Navaroli. In both sessions, she
- Summary of interview with J. Johnson (prepared by Dean Jackson). See also Letter to Candyce Phoenix, May 19, 2022. These documents relate to a second Twitter whistleblower.
- **TWITTER00020545**: Election Threat Model. Shows Twitter's estimated level of preparedness against certain threats and the risk those threats posed. Threats from incitement to violence and policy violations by very-important-tweeters received low-to-medium preparedness scores.
- **TWITTER00019259** and **TWITTER00019229**—the Coded Incitement to Violence policy and the Post-Election Protest Guidance which replaced it days later.
- **TWITTER00000736** – and open letter from Twitter staff regarding the events of January 6th.
- "Briefing on Safety Policy and Violent Organizations Policy with Twitter Staff," Dean Jackson (March 23, 2022).

Facebook

- **FB-CAP-24827**: a full list of all "break glass" measures, the dates at which they were activated or inactivated, and their final activation status.
- "Stop the Steal and Patriot Party: The Growth and Mitigation of an Adversarial Harmful Movement." This report is available in a few places, but the most readable is a leaked version of available via Buzzfeed here: https://www.buzzfeednews.com/article/ryanmac/full-facebook-stop-the-steal-internal-report. It remains the most useful source of insight about how Stop the Steal grew despite Facebook's efforts to contain it, and why the company failed to do better.
- **FB-CAP-00013392**: Another internal retrospective on Stop the Steal, this document contains several useful graphics. It shows, among other things, that Facebook acted on fewer than a sixth of Stop the Steal groups in November 2020 and that the lack of a policy against election delegitimization inhibited the company's response.

- "Capitol Riots – BTG Response": Provided to the Select Committee here: https://drive.google.com/file/d/1y5JbfZRolQsRWAPkTny6FJUO-LoyRFti/view?usp=sharing. This document describes several "break the glass" measures employed before the election and rolled back afterward, then redeployed after January 6th.
- Transcribed Interview with Brian Fishman: Fishman was head of Facebook's Dangerous Organizations policy and gives insight into Facebook's preparations for the 6th and claims he advocated for stronger action against Stop the Steal groups in December.
- Summary memos from on-the-record interviews with Tom Cunningham and Dmitry Borodaenko: Cunningham and Borodaenko were data scientists at Facebook, both of whom left in 2020. Cunningham felt strongly that Facebook was contributing to polarization and causing harm in US politics, and believed the company failed to take objective measures to reduce the amount of low quality civic content out of fear of political blowback. Borodaenko was fired after publicizing evidence of interventions in favor of conservative publishers by Facebook executives.
- **FB-CAP-00009657**: Samidh Chakrabarti, head of Facebook's civic integrity team, delivered this feedback on VP for Integrity Guy Rosen to HR on January 27th, 2021. Chakrabarti claims he pushed Rosen to focus on election delegitimization. Rosen said he did not even want to study the problem for fear of then having to do something about it.
- **FB-CAP-00012605**: A request in early fall to VP of Integrity Guy Rosen to more closely monitor problematic Facebook groups, later a vector for Stop the Steal. Rosen is supportive but later in the chain he has a candid exchange about things the public policy side of Facebook may or may not approve.
- **FB-CAP-00010172**: An email to Facebook integrity VP Guy Rosen and other key personnel laying out concerns on the platform in the early fall. Contains useful information on some of the most powerful interventions Facebook designed and why they were not used effectively.
- **FB-CAP-00010376**: Update on Civic Groups and BTG measures in advance of the election. Contains information about how groups are categorized as "civic" and about a site error in October that found tens of thousands of groups were not receiving strikes for violent incitement.
- **FB-CAP-00004004** and **FB-CAP-00010376**: Groups updates from Ryan Burmeister discussing challenges with detecting violence and incitement. The accuracy of Facebook's AI classifiers was an ongoing challenge, as was the fact that a great deal of harmful content takes place in homogenous groups where users do not report one another—an important signal for detecting harm and one which can be used to train those AI systems.
- **FB-CAP-00014022**: This email chain contains a great deal of information about the BTG measures as they were rolled out before the election.
- Summary of Briefings with Ryan Beiermeister and Nathaniel Gleicher: Beiermeister is a director of civic product at Facebook who oversaw BTG measures related to groups during the election. She was defensive about the company's approach to civic integrity,

but email evidence shows her working to address a known vulnerability at the time. Gleicher is head of Security Policy and discussed policies rolled out in early 2021 which were informed directly by the company's failure to contain Stop the Steal.

Google

- **GOOG-HSCI-00000001**: A basic retrospective on YouTube's election response. Seems to be intended for external audiences. While useful it is not comprehensive or as forthright as an internal retrospective could be.
- **GOOG-HSCI-00000386**; **GOOG-HSCI-00001378**; **GOOG-HSCI-00006804**: Documents touching on changes to YouTube's recommendation algorithm. Some of these are responses to previous Congressional inquiries and mirror evidence given to the Select Committee.
- **GOOG-HSCI-00001370**: Description of how Google evaluates borderline YouTube content.
- Memo on May 16, 2022, Briefing from Google for Select Committee Staff: Representatives from YouTube briefed Committee staff on the company's content recommendation policies. There were several follow up questions from this about how the company treated borderline election fraud content, which was labeled but allowed to remain on-platform. One major takeaway from this briefing is that the policy against election denial did not extend backward and mostly applied to claims of voting irregularities, not other forms of delegitimization.

TikTok

- Memo on May 24, 2022, Briefing with TikTok's Head of Safety: A wide ranging and useful conversation about how TikTok's content policies have changed over time and the actions it takes to restrict prohibited content. Interestingly, TikTok's election misinformation policy applies globally at all times and it has clear policies against violent incitement—areas which contrast with YouTube, Facebook, and Twitter's actions during the election period.
- **TT16SC_0001749**: Information on TikTok's "grayzoning" approach to borderline content. Often this strategy was successful but there were occasional high-profile failures leading to millions of views.
- **TT16SC_0001398**: Document with detail's on TikTok's misinformation policy and its development.
- **TT16SC_00000717** and **TT16SC_00000609**: These documents contain information on TikTok's response to January 6th.
- **TT16SC_0001615**, **TT16SC_000179** and **TT16SC_0001666**: These documents contain information on TikTok's response to QAnon, militia groups, and other dangerous actors.

Reddit

- **JAN6_0285:** Reddit's internal after-action report following the 2020 election that includes areas of suggested improvement around centralizing content moderation and includes an analysis of the subreddits where disinformation was most active.
- **JAN6_0279:** Internal email chain from Reddit showing the nature of its deliberations during January 6th itself. While this email shows that Reddit was not seeing a massive uptick in violent content during the attack, it also shows how its user-mediated content moderation scheme was insufficient, and also contrasts with Twitter's stonewalling.
- **JAN6_0756:** Email between Ory Rinat, chief digital strategist at the White House, and Reddit asking about attempts to regulate hateful content on r/The_Donald, which underscores the Administration's interest in these far-right forums throughout Trump's tenure.

Parler

- **CTRL0000007435:** An internal email from a Parler employee to the FBI about concerning content on Jan. 2, 2021, saying "concerned about Wednesday."
- **CTRL0000007468 and CTRL0000007469:** Internal emails from Parler showing ongoing discussions with the Trump Organization (cc'ing Eric Trump) for the purchase of Parler in late December 2020, while the platform was simultaneously the site of coordinated calls for violence on Jan. 6.

Discord

- **CTRL0000028919:** Discord's after-action review that shows how its January 6th response benefited from a team that was already devoted to violent extremism and the ability to remove servers quickly, but that the attack still caught the platform off-guard. The retrospective also discusses how Trump's Dec. 19 tweet was a pivotal moment.
- **CTRL0000062532:** Case study of pro-Trump Discord server that was shut down after increased violent organizing following Trump's Dec. 19 tweet.

Zello

- **CTRL0000007646:** A list of over 2,000 militia-related channels removed by Zello after January 6th, which is an indicator of the vast use of the platform by far-right extremists.
- **CTRL0000062532 and CTRL0000034886:** Internal after-action reviews from Discord showing its decision to shut down pro-Trump servers after increases in violent organizing targeting January 6th in the wake of Trump's Dec. 19 tweet.

Twitch

- **CTRL0000007943:** Internal discussion of post-J6 Incitement to Violence Policy that is an explicit response to President Trump's calls for action in early 2021.
- **CTRL0000007942:** Internal emails from Twitch about their decision to suspend Donald Trump's stream and the gap his incitement revealed in their policies.